Optimal Search for Product Information*

Fernando Branco
(Universidade Católica Portuguesa)

Monic Sun
(Stanford University)

J. Miguel Villas-Boas
(University of California, Berkeley)

December, 2011

* Comments by Dmitri Kuksov, and seminar participants at the 2010 Choice Symposium, Hong Kong University of Science and Technology, University of British Columbia, University of California, Berkeley, University of North Carolina, University of Pittsburgh, and University of Rochester are gratefully appreciated. We also would like to thank Su Chen for excellent research assistance. E-mail: fbranco@ucp.pt, monic@stanford.edu, and villas@haas.berkeley.edu.

Optimal Search for Product Information

Abstract

Consumers often need to search for product information before making purchase decisions. We consider a tractable (continuous-time) model of gradual learning, in which consumers incur search costs to learn further product information, and update their expected utility of the product at each search occasion. We characterize the optimal stopping rules for either purchase, or no purchase, as a function of search costs and of the importance/informativeness of each attribute. The paper also characterizes how the likelihood of purchase changes with the ex-ante expected utility, search costs, and the importance/informativeness of each attribute. We discuss optimal pricing, the impact of consumer search on profits and social welfare, and how the seller chooses its price to strategically affect the extent of the consumers' search behavior. We show that lower search costs can hurt the consumer because the seller may then choose to charge higher prices. Discounting creates asymmetry in the purchase and no-purchase search thresholds, and may lead to lower prices if search occurs in equilibrium, or higher prices if there is no search in equilibrium. The paper also considers searching for signals of the value of the product, heterogeneous importance of attributes, endogenous intensity of search, and social learning.

1. Introduction

Most consumers search for product information before making their purchasing decisions. They examine the product, touch it, ask for information from other consumers, and maybe read about it. As they do this, they update their beliefs on how much they would enjoy the product. At a certain point, consumers decide not to search any more, and make a decision as to whether or not to purchase the product. A 2010 survey from Zillow Mortgage Marketplace, for example, studied the average amount of time spent by American consumers researching several types of purchases.[1] The survey found that the average amount of time spent researching a new purchase was 40 hours for a new home, 10 hours for a major home improvement, 10 hours for a car, 5 hours for a vacation or a mortgage, 4 hours for a computer, and 2 hours for a television set.

One distinct feature of information search is that consumers often gather a large amount of small pieces of product information, from potentially different sources, with gradual belief updating at each search occasion. Consider books for example. Suppose a consumer is trying to decide whether to purchase a fiction book, "The Devil Wears Prada." He would be able to find many small pieces of information on the book such as the name and introduction of the author, the publisher, the publication date, the number of pages, the price, the available delivery methods, potential discounts, various sellers' descriptions of the book, as well as thousands of consumer reviews,[2] and sometimes even further opinions on these reviews.[3] Consider now computer games, for a second example. Suppose a consumer is interested in the game, "The Legend of Zelda." There are many images, videos, descriptions and reviews of this game, all of which can help potential consumers get familiar with various aspects of the game. For another example, consider makeup products, say the BADgal Lash Mascara from Benefits. If a consumer searches for information on this product, again the consumer would encounter various pictures, descriptions and reviews that are related to many different aspects of the mascara.

---

[1] http://zillow.mediaroom.com/index.php?s=159&item=201, accessed in July 2010.

[2] See, e.g., http://www.amazon.com/Devil-Wears-Prada-Novel/dp/0767914767, accessed in December 2011.

[3] Throughout the paper we refer to the consumer as a 'he,' and to the seller as a 'she.'

In all of these examples, the focal product has many small attributes that are uncertain to consumers *a priori*, consumers obtain information on these attributes in a somewhat random order, and their learning and updating tends to be smooth and gradual. As consumers gather information, they continuously update their beliefs on how much they would enjoy the product. At a certain point, consumers decide not to search any more, and make a decision as to whether or not to purchase the product. That is, there are optimal stopping points in the search process, beyond which any additional product information gathered is no longer worth the cost.[4]

From the managers' perspective, it is important to understand how consumers optimally trade off the different forces in their search process, and where the optimal stopping points are for this process. Recent development of web analytics tools have made it easier for managers to track consumers' search behavior. By trying to understand the factors that affect information search, a manager could potentially influence the extent of search. *A priori*, it is not clear whether it benefits the seller when consumers spend more time searching for product information – the information obtained could either increase or decrease the consumer valuation of the product. As a result, the extent of search is closely linked with the price charged and profit obtained.

Given these considerations, we study the following questions in this paper. First, when would a consumer stop searching for more information and make a decision on whether to purchase a product? Second, given an understanding of the optimal stopping rule, how can a seller manipulate price to influence the extent of search? Third, what factors determine the consumer's search behavior and the firm's pricing decisions in this context?

In order to answer these questions we consider a framework in which a consumer could sequentially gather information about a focal product. As a consumer searches, he obtains information on one additional attribute in each step of the search process, and the valuation of the product is the sum of the utility from all the product attributes. The information contained in an attribute could either increase or decrease his expected valuation of the product. With an infinite number of attributes, each providing an infinitesimal amount of information, we can obtain that the expected valuation follows a Brownian motion as the consumer keeps searching.

Setting up the problem in this way we can then show that the information acquisition problem of the consumer can be solved by using the continuous-time general methodology for analyzing real options (e.g., Dixit and Pindyck, 1994) or of the theory of options in financial markets, with the specifics inherent to consumer search for information. This yields a rich, yet simple and tractable, description of a consumer's search and purchasing decision process. Furthermore, we set up the problem such that the optimal stopping policy is characterized in closed form, and we obtain the effects of endogenous pricing on search, consumer surplus, and profits. The model also allows us to investigate the effects of discounting, signals over product valuations, heterogeneous importance of attributes, endogenous search intensity, and social learning.

Some of the market forces of the consumer search for information problem considered here have been introduced in the literature (e.g., Chatterjee and Eliashberg 1990, Putsis and Srinivasan 1994) as discussed below. However, to the best of our knowledge, there are no results in the literature fully

[4]The development of the Internet may have lowered search costs, but consumers still end up making purchase decisions without full information, as search costs still include the information processing costs (see, e.g., Chen and Sudhir 2004). Consumers seem to be cautiously trading off their search cost with the likelihood that they will eventually purchase and enjoy the product.

solving such a problem with closed-form expressions, and considering the different effects analyzed here. We further discuss below the relation of the results presented here to the literature.

We show that the optimal stopping rule consists of an upper bound and a lower bound, which we refer to as the "purchase threshold" and "exit threshold," respectively, on the consumer's expected valuation. When the consumer's expected valuation of the product hits the purchase threshold, the consumer stops searching and purchases the product. When the expected valuation hits the exit threshold, the consumer stops searching and does not purchase the product. When the expected valuation is in between the two bounds, the consumer continues to search.

We derive closed-form expressions for these two optimal bounds. One can obtain how they are determined by the marginal cost of search and the importance of each product attribute, i.e., the amount of information contained in each product attribute. If it is more costly for the consumer to keep searching, consumers search less, and the purchase threshold becomes lower. This result suggests that the purchase threshold may actually become higher as the Internet drives down consumer's search costs. One implication of this is that the likelihood to purchase can decrease for some purchase occasions with the development of the Internet. Specifically, if the ex-ante expected utility is negative, increasing search costs makes the consumer less likely to buy the product, while if the ex-ante expected utility is positive, increasing search costs makes the consumer more likely to buy the product. These results suggest that, for example, while sellers of niche and less well-known brands may benefit from the increasing ease of disseminating product information through the Internet and other channels, established brands may in fact be potentially harmed by too much information on some product attributes. This implication is consistent with the long tail phenomenon in the recent literature of the impact of the Internet on sales distribution across different sellers (e.g., Brynjolfsson et al., 2011).

Moreover, the purchase threshold increases with the amount of information contained in each product attribute: If the consumer's expected valuation of the product is more volatile as he searches through different product attributes, he wants to be more confident and positive about the product before deciding to purchase, hence the purchase threshold is also higher. This result coincides with real-world observations that when the utility from a product is perceived to be highly uncertain, buyers may invest more time and effort in information gathering before making a purchase decision.

Depending on the ex-ante valuation, the seller either charges a low price to eliminate search for information or a high price to stimulate search. When the ex-ante expected valuation (prior to any search) is high, the seller maximizes her profit by charging a price such that the consumer buys the product right away. Stimulating search in this case has a big cost (driving down purchase likelihood) and a small benefit (potentially increasing the consumer's willingness to pay). When the ex-ante expected valuation is low, the seller encourages the consumer to search for more product information. In the latter case, the price would be too low if the seller insists on getting the consumer to purchase without any search. In equilibrium, the purchase likelihood depends on the ratio of the ex-ante valuation and the purchase threshold.

The profit always increases with the ex-ante expected valuation of the product. The search cost, however, enters into the profit equation in a more sophisticated manner. Essentially, lower search costs lead to more search, or in other words, to a higher purchase threshold. When this happens, the optimal price increases, but the purchase likelihood can decrease. As these two forces trade off against each other, the profit may either increase or decrease with the purchase threshold. Overall, the seller wants to encourage search at the margin, or increase the purchase likelihood, if the ex-ante expected valuation is low. This is because information search may then lead to both higher purchase likelihood and higher price. The seller wants to hinder search if the consumer's ex-ante expected valuation of the product is high, and have the consumer buy the product with little additional information. Besides using price, the seller may also be able to influence the purchase threshold by changing how much product information is made available and how such information is communicated to the consumer.

We investigate the effect of discounting by consumers and find that discounting leads consumers to search less, with a greater effect on the purchase threshold than on the exit threshold. When the consumer has beliefs about the expected utility of purchase that are close to the purchase threshold, the consumer is likely to end up purchasing the product, and due to discounting of the expected benefits, the consumer wants to purchase sooner. This leads to a lower purchase threshold. In terms of pricing, discounting then leads the firm to price lower than under no discounting if the firm decides to price such that search occurs in equilibrium, as the consumer has greater value of purchasing sooner. On the other hand, if the firm decides to price such that there is no search in equilibrium, then the optimal price is higher than under no discounting, as the reservation utility of the consumer is now lower. For the extreme case of discounting where there are zero search costs (that is, all the search costs are due to delay in purchase), then there is only a purchase threshold but there is no longer an exit threshold. Obviously, as search is costless, if the consumer does not buy, he will always search.

We consider a variety of extensions of the base model. We show that the model is qualitatively similar to one in which consumers try to infer the "true" valuation of the product from a series of independent signals. Technically, the latter model generalizes our baseline model in making the amount of information contained in each attribute vary through the search process. We also show how this extension can be seen as one with heterogeneous importance of attributes or with correlated attribute fit. We further discuss what happens when the mass of attributes that the consumer could check is finite. In another extension, we allow for the possibility of consumers choosing the search intensity per unit of time. In this case we show that consumers search with greater intensity when their valuation of the product is closer to the purchase threshold. Finally, we also consider the case of social learning where later consumers can additionally gain information from observing purchase decisions of previous consumers. Table 1 summarizes the qualitative takeaways from all the extensions of our model.

The current paper contributes to the literature on information gathering. Papers in this literature often study searching for the lowest price (e.g., Diamond 1971, Stahl 1994, Kuksov 2004), or searching for the best-matched alternative (e.g., Weitzman 1979, Wolinsky 1986, Bakos 1997, Anderson and Renault 1999, Armstrong et al. 2009) with complete learning of the value of an alternative when that alternative/price is searched. This literature is similar to this paper as costs are cumulative while rewards are uncertain. However, information in this literature is not cumulative: the reward in each box has a probability distribution that is independent of the other rewards, and only one reward can be obtained. On the other hand, in this paper, there is only one ultimate valuation of the product, and all search yields useful information for figuring out what that valuation is.[5] That is, our paper focuses on

[5]Hauser et al. (1993) consider the rules that a consumer can use on deciding the order to check different sources of information. See also Moorthy et al. (1997) for an investigation of the empirical implications from the results in Weitzman (1979). For competition with search costs see also, for example, Varian (1980), Narasimhan (1988), Kuksov (2006), Cachon et al. (2008). For the effect of search costs creating a potential hold-up problem see, for example, Lal and Matutes (1994), Wernerfelt (1994), Rao and Syam (2001), Villas-Boas (2009). Search costs can also have effects on consideration sets (e.g., Hauser and Wernerfelt 1990, Mehta et al. 2003), and channel structure (e.g., Gal-Or et al. 2008). Consumers can also obtain information by learning from other consumers (e.g., Ellison and Fudenberg 1995, Zhang 2010), or experimenting

the fundamental properties of gradual, cumulative private learning through search prior to the decision of whether to purchase a single product.

Through focusing on a single product, we present a tractable and rich model of consumer search for information that incorporates uncertainty around what information is obtained at each search stage, and optimal stopping rules for search that lead to either purchase or no purchase. Given the optimal stopping rules, we derive the optimal pricing decision by a firm, taking into account the effect of price on endogenous search for information.[6]

One related paper in terms of model set-up is Roberts and Weitzman (1981). That paper models Sequential Development Projects (SDP) that feature additive costs and uncertain benefits. While that paper focuses on how firms can choose to optimally decide whether to continue researching a sequential project, we focus on how a consumer gathers information on a product before deciding whether to purchase it. One distinct difference between our model and Roberts and Weitzman is that in that model information eventually runs out as the outcome of the Research and Development (R&D) becomes clear. Consequently, the boundaries of search in that model would always shrink as the firm moves toward the end of uncertainty. Our model, on the other hand, assumes that the consumer can always search for more information before he makes a purchase, motivated by the observation that online search engines now provide boundless opportunities for a consumer to research a product. This allows us to obtain sharper results on the consumer search thresholds and search behavior. In addition we also model in detail how consumer beliefs evolve through the search process given the ex-ante priors and expectations. The focus on the consumer, combined with the departure from the limited-information assumption, enables us to get closed-form results, provide insights for optimal consumer search behavior for product information, generate implications for pricing, and discuss the impact of discounting, choice of search intensity, and social learning. Two other related papers are Chatterjee and Eliashberg (1990) and Putsis and Srinivasan (1994). Chatterjee and Eliashberg present a model of consumer adoption of an innovation under gradual information acquisition, in which a consumer chooses to adopt when the expected utility of adoption exceeds the price of adoption. Putsis and Srinivasan consider the issue of information acquisition focussing on the duration of the consumer deliberation process. The model discussed generates hypothesis on the duration of the deliberation which are then tested using survey data on new car purchases.

The remainder of the paper is organized as follows. The next section presents the model, and Section 3 characterizes the optimal consumer search. Section 4 considers the optimal decisions of a firm. Section 5 looks at the effect of consumer discounting, and Section 6 discusses extensions of the model. Section 7 presents concluding remarks.

2. The Model

across multiple products (e.g., Bergemann and Välimäki 1996, Villas-Boas 2004). For empirical estimates of search costs see, for example, Honka (2010), De los Santos et al. (2010).

[6]Also related to this paper in terms of modeling of continuous learning of information are Bolton and Harris (1999) and Bergemann and Välimäki (2000). Both of these papers model the evolution of beliefs of agents about whether a product has high or low quality, while here we model beliefs for any level of quality, which is important to understand when a consumer stops searching to buy or stops searching not to buy a product. Neither of these papers looks at optimal consumer search for information, considering instead strategic experimentation in Bolton and Harris and experimentation with social learning in Bergemann and Välimäki.

Consider a consumer who gathers information sequentially on the attributes of a particular product to decide whether or not to buy it.[7] The total utility that he ultimately derives from the product is the sum of utilities obtained from each of the T product attributes. This can be written as the ex-ante expected utility prior to learning about any attribute, v, plus the sum over attributes of the difference between the realized utility from an attribute and his expected utility. For attribute i, we denote this difference between the realized utility from that attribute and the expected utility as x_i. The realized utility from a product is then $U = v + \sum_{i=1}^{T} x_i$.

The term v can also be seen as the stand-alone valuation of the product, the utility from the most basic function of the product, such as the "picture-taking" function of a camera, the "hunger-reducing" function of a meal, or the "getting-you-around" function of a car. Without searching for information, the only thing that the consumer knows is v. In this interpretation, besides the most basic function, there exist product attributes that additively affect the total utility.

Multiple attributes of a product are often considered before a purchase is made. For example, consumers can check the color, pattern, material, and size of a product. The realization of each attribute (e.g., a car being yellow) either increases or decreases the total expected utility. That is, x_i, the deviation from the expected utility of the attribute, can be either positive or negative. We assume that the x_i's are independent across i. One interpretation is that after checking t attributes, checking the next attribute i changes the expected utility by x_i with what is unexpected given the information on the prior t attributes already checked. Given that x_i is unexpected it is then independent of the previous x_i's that were checked.

Suppose that for any attribute i, x_i can take the value of z or $-z$ with equal probability, such that the expected value of x_i is zero. That is, attribute i can be either a fit or no-fit with the consumer's preferences, with each specific attribute-fit being independent across consumers. The value of z can be interpreted as the weight of a product attribute on the consumer's utility. That is, the change in utility from the attribute, x_i, is equal to $z \cdot y_i$, where y_i takes the value of 1 or -1 with equal probability. We consider that z is the same across all attributes, but in general, the analysis remains the same as long as consumers do not know the value of z prior to checking an attribute.[8] We assume that the number of attributes, T, is infinity. This allows us to get sharp results, and captures the idea that the number of attributes is large enough, such that the set of attributes is never fully searched.[9] Conceptually, one can think of each "attribute" in the continuous framework below as a small piece of product information that a consumer can process within a step of search. Some complicated aspects of the product, such as safety for a car, would hence be composed of many "attributes."

The consumer does not know *a priori* the value of x_i for any i. When examining an attribute i, a consumer pays a search cost c and learns the realization of x_i. Note also that the search cost is assumed to be the same for all attributes. The search cost could take many forms in reality: the time and effort it takes to use search engines to gather relevant information on the product attribute, the mental costs of

[7]Even in situations where information on multiple attributes is displayed simultaneously to the consumer, the consumer may have to process these attributes one after another. We do not model the seller's incentive to add or delete product attributes. See Ofek and Srinivasan (2002) for a discussion on that issue.

[8]Alternatively, one could have that at each search occasion the consumer gets an independent signal of the product quality. This would lead to similar general ideas, with less sharp results. In Section 6 below we explore the features of such a model. We also consider there the case in which the consumer knows ex-ante which attributes are more important (heterogeneous and known z prior to search).

[9]In Section 6 below we discuss what happens when the set of attributes that consumers can check is bounded.

processing information on an attribute (e.g., Shugan 1980), the travel cost to the store to get a feel for the product, or even the monetary cost paid to a consultant if the product is rather complicated. Upon checking attribute i, the consumer's expected utility goes either up or down by z. As it will become clear in the next section, the consumer essentially trades off the cost of search with the possibility of gathering positive information which eventually leads to a purchase.

After examining t attributes, the "current expected valuation" of the product becomes $u(t) = v + \sum_{i=1}^{t} x_i + \sum_{i=t+1}^{T} E(x_i|t) = v + \sum_{i=1}^{t} x_i$. In the presentation, when it is clear, we write u in place of $u(t)$ for ease of notation. One can see u as a state variable that changes stochastically with the number of attributes checked. Note that by the assumptions above, $E(x_i|t) = 0$ for all i.

In order to simplify the solution of the dynamic programming problem we work with a continuous rather than discrete process.[10] Imagine that there are many small-valued attributes, such that the change in utility from each attribute, z, gets infinitely small as the number of attributes goes to infinity. The u process then becomes continuous. In the limit, it becomes a Brownian motion (Cox et al., 1979). That is, $du = \sigma\, dw$, where w is a standardized Brownian motion and $\sigma > 0$ is the standard deviation of the Brownian motion. By definition, $z = \sigma\sqrt{dt}$. The parameter σ therefore indicates how informative the learning about an attribute is, or more generally, the amount of updating the consumer can do when checking each product attribute. Figure 1 gives an example of how u can change with the amount of attributes checked.

After checking t units of attributes, the consumer has to make two decisions. First, does he want to continue examining attributes, or stop? Second, if he stops gathering information, should he buy the product or not? The answer to these questions is provided in the next section.

3. Optimal Consumer Search

3.1. Expected Utility of the Search Problem

To analyze the consumer's search behavior it helps to characterize the expected utility function of the search problem under the consumer's optimal decision making, $V(u,t)$, given that the expected utility of the product if purchased without further search is u and the consumer has checked t units of attributes. This expected utility of continuing to search, after checking t units of attributes, can be written as

$$V(u,t) = -c\, dt + EV(u + du, t + dt). \tag{1}$$

By a Taylor expansion (see, e.g., Dixit 1993), we obtain

$$V(u,t) = -c\, dt + V(u,t) + V_u E(du) + V_t dt + \frac{1}{2} V_{uu} E[(du)^2] + V_{ut} E(du)\, dt, \tag{2}$$

where V_u is the partial derivative of $V(u,t)$ with respect to u, V_t is the partial derivative of $V(u,t)$ with respect to t, V_{uu} is the second derivative of $V(u,t)$ with respect to u, and V_{ut} is the cross derivative of $V(u,t)$ with respect to u and t.

[10]While in some cases product attributes are obviously discrete, in other cases the information gathering process can be rather continuous. For example, a consumer might be constantly updating his beliefs on the valuation of a novel or a movie when reading or watching a sample of it.

As $E(du) = 0$ and $E[(du)^2] = \sigma^2\, dt$ we have, dividing (2) by dt,

$$-c + V_t + \frac{\sigma^2}{2} V_{uu} = 0. \tag{3}$$

Given that, conditional on u, the number of attributes already checked does not affect the consumer's problem under the assumptions considered, we can obtain

$$V_{uu} = \frac{2c}{\sigma^2}. \tag{4}$$

Hence $V(u)$ is quadratic in u (see the Appendix for further explanations on this derivation).[11]

In addition, we also have two sets of boundary conditions, when u is large enough such that the consumer chooses to buy the product, and when u is low enough such that the consumer chooses to stop searching and not buy the product. When u is large enough such that the consumer is indifferent between continuing the search process, and stopping the search to purchase the product, we have

$$V(\overline{U}) = \overline{U} \text{ and } V'(\overline{U}) = 1, \tag{5}$$

where $\overline{U}$ is the upper bound such that when u reaches $\overline{U}$ the consumer stops searching and buys the product. Graphically, condition (5) states that the curve of $V(u)$ is tangent to the 45 degree line at $u = \overline{U}$ (see Figure 2).

The first part of condition (5) just means that when the expected utility of buying the product reaches $u = \overline{U}$ then the consumer chooses to purchase the product and gets the expected utility $\overline{U}$.[12] The second part of condition (5) is known as the "high contact" or "smooth pasting" condition (e.g., Dumas 1991, Dixit 1993). The condition is implied by the fact that the consumer maximizes $V(u,t)$ for all (u,t). The Appendix provides further intuition on this condition.

Next, we write down the second set of boundary conditions, when u is low enough such that when u reaches $\underline{U}$ the consumer stops searching and decides not to buy the product,

$$V(\underline{U}) = 0 \text{ and } V'(\underline{U}) = 0. \tag{6}$$

Condition (6) is similar to condition (5) for the case in which the consumer decides to stop searching without buying the product. The first part of (6) indicates that at $u = \underline{U}$, the consumer does not expect any positive utility from buying the product. The second part indicates that even if u starts to increase as the consumer continues to search, the change of $V(u)$ will be very slow. It means that the curve $V(u)$ is tangent to the x-axis at $u = \underline{U}$ (see Figure 2).

3.2. Optimal Stopping Rule

In order to derive the optimal stopping rule, consider first some intuition of why the consumer would stop searching when u is either high enough or low enough. When u is far above zero, the likelihood of u returning back to zero is low. Even if u comes back to zero, the consumer would have to pay substantial

[11] As $V(u,t)$ is independent of t we then just write $V(u,t)$ as $V(u)$.

[12] This is equivalent to exercising an American put option with $\overline{U}$ being the value of the strike price minus the stock price.

search costs before seeing it happen. Therefore, when u first hits $\overline{U}$, the consumer buys the product right away. Similarly, when u is near $\underline{U}$, the expected valuation is far below zero, and it is not likely to return above zero again. Rather than paying additional search costs for the low probability event, the consumer terminates the search process without buying the product.

Equations (4)-(6) fully determine $V(u)$:

$$V(u) = \frac{c}{\sigma^2}u^2 + \frac{1}{2}u + \frac{\sigma^2}{16c}. \tag{7}$$

From (5)-(7) we can then obtain

$$\overline{U} = -\underline{U} = \frac{\sigma^2}{4c}. \tag{8}$$

We have now derived the consumer's optimal stopping rule.

PROPOSITION 1: *The consumer searches for more product information when* $-\frac{\sigma^2}{4c} < u < \frac{\sigma^2}{4c}$. *The consumer stops searching and buys the product when* $u \geq \frac{\sigma^2}{4c}$. *The consumer stops searching without buying the product when* $u \leq -\frac{\sigma^2}{4c}$.

As implied by the proposition, if the stand-alone valuation (ex-ante valuation prior to any search) v is greater than $\overline{U}$, the consumer should buy the product without any search. If v is lower than $\underline{U}$, the consumer should exit the market also without any search. If v is between the two thresholds, the consumer should initiate search, in the hope of eventually making a purchase.

We discuss more properties of the optimal stopping rule by looking at a graphical presentation of $V(u)$ in Figure 2.

Two features of $V(u)$ are noteworthy. First, $V(0) > 0$: the consumer expects positive utility from the possibility of buying the product even when his current expected valuation of buying the product is zero. This reflects the option value of the search process: By searching, a consumer might find that the product provides a good fit and generates a positive expected utility, and if the product turns out to be a poor fit the consumer can always choose not to buy it.

Second, the upper and lower bounds turn out to be symmetric around zero. The upper bound being above zero and the lower bound being below zero confirms the intuition that the consumer trades off the likelihood of changing his purchase decision with the amount of search costs he needs to pay.[13] This intuition explains why the starting point v does not affect the boundaries. The intuition also explains why, for example, the upper bound increases with σ. When the attributes become more important, a high u can walk back to zero in fewer "steps" and the consumer has more incentive to continue searching. On the other hand, when the search cost, c, decreases, it is also easier for the consumer to gather enough information to change his purchase decision. He therefore has more incentive to continue searching, which leads to a higher $\overline{U}$.

[13]The symmetry around zero depends on the symmetric properties of the Brownian motion. As shown below in Section 5, with discounting we no longer have the symmetry of $\overline{U}$ and $\underline{U}$ around zero, as with discounting the consumer would prefer to purchase earlier and not delay the product benefit, leading to $\overline{U}$ being closer to zero than $\underline{U}$.

3.3. Purchase Likelihood

It would be interesting to understand how the parameters of the model affect the purchase likelihood. As a next step, we write out the purchase likelihood given any $u, Pr(u)$. (The proof is in the Appendix.)

PROPOSITION 2: *The consumer's purchase likelihood $Pr(u)$, for any $u \in [\underline{U}, \overline{U}]$, is*

$$Pr(u) = \frac{u - \underline{U}}{\overline{U} - \underline{U}} = \frac{1}{2}(1 + \frac{u}{\overline{U}}). \tag{9}$$

Proposition 2 states that when search occurs, the purchase likelihood is the ratio of the distance from u to the exit threshold, $u - \underline{U}$, and the total distance between the two thresholds, $\overline{U} - \underline{U}$. Therefore, prior to any search, the consumer buys the product with probability

$$Pr(v) = \frac{1}{2}(1 + \frac{v}{\overline{U}}) = \frac{1}{2} + \frac{2cv}{\sigma^2} \tag{10}$$

for intermediate values of v. If $v > \frac{\sigma^2}{4c}$ the consumer buys the product with probability one, without incurring any search costs. If $v < -\frac{\sigma^2}{4c}$ the consumer buys the product with probability zero and also does not incur any search costs.

One can make several interesting observations from equation (10). First, making search more informative, i.e., increasing σ, makes the consumer less likely to buy the product if $v > 0$, and more likely to buy the product if $v < 0$.

To gain intuition on this result, think of the extreme case of $\sigma \to 0$. In this case, the consumer has no incentive to search and immediately buys the product if $v > 0$, and does not buy the product if $v < 0$. As σ increases, the consumer is more likely to search. When σ is high, a positive initial valuation $v > 0$ has a lower impact on subsequent valuations, such that a subsequent valuation is more likely to fall below the exit threshold. Similarly, when σ is high, a negative initial valuation $v < 0$ also has a lower impact on subsequent valuations, such that a subsequent valuation is more likely to go above the purchase threshold. Therefore, when σ increases, the purchase likelihood would be reduced if $v > 0$, and increased if $v < 0$. If *a priori* a consumer has a good expectation about the value of a product, providing more information may only hurt the purchase likelihood. On the other hand, if *a priori* a consumer has a poor expectation of the value of a product, the only possibility for that consumer to potentially want to buy the product is if the consumer receives more information.[14]

This observation has the strategic implication that a seller who aims to maximize the purchase likelihood should increase the difficulty for the consumer to search if the product brings the consumer a positive ex-ante expected utility, and facilitate consumer search if the product brings the consumer a negative ex-ante expected utility prior to any search. The seller could, for example, decide on how much information to provide. This conclusion for the positive ex-ante utility condition can be seen as fitting real-world observations where a consumer starts researching a product, being excited and optimistic, and after quickly reading some online reviews, decides that the product does not really fit his preferences and hence foregoes the idea of buying it.[15] On the other hand, if consumers would not buy the product

[14]See Ottaviani and Prat (2001) and Johnson and Myatt (2006) for similar points without consumer search.

[15]Sun (2011) has a similar result, through a different mechanism, that a seller whose product has a high vertical quality should avoid disclosing the product's horizontal location in the consumers' taste space.

without further information (the case when the ex-ante utility is negative), the firm wants to provide information such that with some probability the information on fit gained by a consumer is positive and the consumer decides to buy the product.

Consider now the effect of search costs. Note that if the ex-ante expected utility is negative, increasing search costs makes the consumer less likely to buy the product, while if the ex-ante expected utility is positive, increasing search costs makes the consumer more likely to buy the product. The intuition is similar: if the consumer starts out with a negative ex-ante expected utility, as search becomes easier the consumer is more likely to gather positive information, hence more likely to buy the product. These two observations suggest that, as the Internet drives down the cost of gathering product information, the purchase rate may increase among consumers with a negative ex-ante expected utility and decrease among consumers with a positive ex-ante expected utility. These results suggest that, for example, while sellers of niche and less well-known brands may benefit from the increasing ease of disseminating product information through the Internet and other channels, established brands may in fact be harmed by too much information on some product attributes. This implication is consistent with the long tail phenomenon in the recent literature of the impact of Internet on sales distribution across different sellers (e.g., Brynjolfsson et al. 2011).

In terms of social welfare note that a consumer prefers lower search costs, c, and greater informativeness of the attributes checked, σ^2. To see this one can just differentiate the expected utility of the search process, $V(v)$, with respect to c and σ^2 for $v \in (\underline{U}, \overline{U})$ to obtain that $V(v)$ is decreasing in c and increasing in σ^2, and note that for v outside this interval the consumer can only benefit from lower search costs or a greater informativeness of the attributes checked.

3.4. Extent of Search

The number of attributes checked during the search process can vary depending on the results of the search.[16] It is interesting to investigate what is the expected number of attributes checked. To obtain this note that the expected utility of the search process, $V(u)$, is equal to the purchase likelihood times the expected utility of buying the product when the decision to buy the product is made, $\overline{U}$, minus the expected costs of search: $V(u) = Pr(u)\overline{U} - cE[\text{number of attributes searched}|u]$. Using the expressions above for $V(u)$ and $Pr(u)$, one can obtain that $E[\text{number of attributes searched}|u] = \frac{\sigma^2}{16c^2} - \frac{u^2}{\sigma^2}$.

Note that the expected number of attributes searched is increasing in the informativeness of each attribute, σ^2, and decreasing in the costs of search, c. Note also that the expected number of attributes searched is greater the further away the consumer is from both deciding to buy the product and stop the search process, and deciding not to buy the product, the case when u is close to zero.

4. Firm's Pricing Decision

Consider now that the seller chooses a price, p, observable by consumers before they decide whether or not to engage in the search process. The consumer's initial valuation, prior to any search, becomes $v - p$. Conceptually, it is $u - p$ that is moving now as a consumer searches. Recall from (8) that the starting point does not enter the expression of the optimal stopping boundaries. Therefore, the consumer buys the product when $u - p$ first hits $\overline{U}$, exits without buying when $u - p$ first hits $\underline{U}$, and

[16] Given the continuous framework, the number of attributes checked means the mass of attributes checked.

otherwise continues to search. In what follows we restrict attention to the case where v is homogeneous across consumers, but a similar analysis could be considered with heterogeneous v, and then in the expected profit one would have to integrate the profit across all v.[17]

Based on (10), the ex-ante likelihood of purchase becomes

$$Pr(v-p) = \begin{cases} 1, & \text{if } v-p \geq \overline{U}; \\ \frac{1}{2} + \frac{2c(v-p)}{\sigma^2}, & \text{if } \underline{U} < v-p < \overline{U}; \\ 0, & \text{if } v-p \leq \underline{U}. \end{cases} \tag{11}$$

Given (11), and a marginal cost g, the seller maximizes her expected profit, $(p-g)\cdot Pr(v-p)$, which leads to the following result.[18]

Proposition 3: *The seller can make a positive profit if $v \geq g + \underline{U}$. The optimal price is*

$$p^* = \begin{cases} v - \frac{\sigma^2}{4c} = v - \overline{U}, & \text{if } v \geq g + 3\overline{U}; \\ \frac{v+g}{2} + \frac{\sigma^2}{8c} = \frac{1}{2}(v+g+\overline{U}), & \text{if } g + \underline{U} \leq v < g + 3\overline{U}. \end{cases} \tag{12}$$

It is intuitive that when v is high, as in the first case, the seller would rather that the consumers buy the product without any search.

When v is not that high, as in the second case in Proposition 3, the optimal price turns out to be increasing in both v, the ex-ante expected utility, and $\overline{U}$, the final valuation when the consumer decides to purchase the product. Note that in this case the optimal price will move in the opposite direction with the search cost: if search were easier, the consumer has a higher purchase threshold which leads to a higher equilibrium price. Lower search costs lead to the possibility of consumers learning about their higher valuation, which then leads the firm to charge higher prices.[19]

The expression for the optimal price in the case with consumer search is noteworthy. If it were not possible for the consumer to search at all, the seller would simply charge v; if the seller can choose the price after the consumer decides to purchase the product, she would charge $\overline{U}$. The optimal price is in between the two scenarios above: the consumer is given the option to search, and the seller needs to commit to a price before the consumers starts searching. Correspondingly, the seller takes these two prices into account in her pricing. The intuition behind this result is based on standard monopoly pricing. The demand, or probability of buying, is proportional to $v + \overline{U} - p$, and the marginal cost is g. Therefore, profit is maximized when marginal revenue equals marginal cost, implying $p^* = \frac{1}{2}(v+g+\overline{U})$.

The ex-ante expected valuation of buying the product, or the starting point of search, now becomes

$$v - p^* = \begin{cases} \overline{U}, & \text{if } v \geq g + 3\overline{U}; \\ \frac{1}{2}(v-g-\overline{U}), & \text{if } g + \underline{U} \leq v < g + 3\overline{U}. \end{cases}$$

[17]Note that for heterogeneous v, a firm could potentially have the ability to condition the price on whether a consumer searches. This possibility is not considered here, but it would be interesting to investigate in future research (see also Armstrong and Zhou, 2010). Another related interesting possibility not considered here is to be able to price condition on the extent of consumer search.

[18]Note that since attribute-fit is assumed to be independent across consumers, there is no role for price signaling product quality. Note also that signaling would not occur if product attribute-fit were correlated across consumers but the firm were not informed about which product attributes would generate greater fit.

[19]Note that the result of higher prices with lower search costs can also be obtained if there is only one attribute to inspect, without the gradual search effects considered here. In that case lower search costs can lead more consumers to search, and the firm can then target the consumers who have found a greater product-fit with a higher price.

Consequently, the ex-ante purchase likelihood is

$$Pr(v-p^*) = \begin{cases} 1, & \text{if } v \geq g + 3\overline{U}; \\ \frac{1}{4}(1+\frac{v-g}{\overline{U}}), & \text{if } g + \underline{U} \leq v < g + 3\overline{U}. \end{cases} \tag{13}$$

Note that whenever v is not too high, the seller chooses the price so that the consumer would conduct some search before making the purchase decision. The equilibrium purchase likelihood depends on the level of $\frac{v-g}{\overline{U}}$. Based on (12) and (13), we can obtain the equilibrium profit:

$$\Pi(v) = \begin{cases} v - g - \overline{U}, & \text{if } v \geq g + 3\overline{U}; \\ \frac{(v-g+\overline{U})^2}{8\overline{U}}, & \text{if } g + \underline{U} \leq v < g + 3\overline{U}. \end{cases} \tag{14}$$

Naturally, profit always increases with the stand-alone valuation v. By taking the derivative of $\Pi(v)$ with respect to $\overline{U}$, one can also see that the expected profit increases with $\overline{U}$ if and only if $v < g + \overline{U}$. Intuitively, as $\overline{U}$ increases the optimal price goes up and the purchase likelihood goes up or down depending on whether $v < g$ or $v > g$, respectively. When $\overline{U}$ goes up by one unit, the price always increases by half a unit. The purchase likelihood, however, decreases faster when v is big: when v is near g, the purchase likelihood is near $1/4$ and is not affected much when $\overline{U}$ increases by a unit; when v is high, the purchase likelihood is almost one and can be driven down significantly when $\overline{U}$ increases by one unit.

Therefore, a profit-maximizing seller should aim to increase $\overline{U} = \frac{\sigma^2}{4c}$ when v is small, i.e., $v < g + \frac{\sigma^2}{4c}$, and decrease $\overline{U}$ when v is big.[20] In most cases, the seller cannot fully control the level of either the informativeness of search σ or the search cost c. At best, she can increase the informativeness of search by, for example, trying to provide better answers to consumers' questions or making specifications of product attributes, such as color and size, vastly different. Similarly, the seller can try to decrease the search costs by, for example, circulating online video introductions of the product. We state the results above in the following proposition.

PROPOSITION 4: *The profit is increasing (decreasing) in the search costs and decreasing (increasing) in the informativeness of search if* $v > (<) g + \frac{\sigma^2}{4c}$. *The highest profit is obtained when* $\frac{\sigma^2}{4c} \to \infty$, *and, given* v, *the lowest profit is obtained when* $\frac{\sigma^2}{4c} = \max[v-g, 0]$.

Now consider the effects of search costs and informativeness of search on social welfare, given that the firm chooses the price to charge. To see this, first consider the effect on consumer surplus, and let us focus the analysis on the case in which a firm chooses its price such that consumers decide to search.

An increase in search costs reduces the consumer surplus (the expected utility of the search process) for a given price, but in this case the firm may lower its price. To obtain the total effect, one can write the consumer surplus given the price charged by the firm as $V(v-p^*) = \frac{(v-g+\overline{U})^2}{16\overline{U}} = \frac{1}{2}\Pi(v)$.

[20] We consider a situation where it is increasingly costly for a seller to change the search costs or the informativeness of search, such that the change in $\overline{U}$ does not affect the region of the parameter space where we are in terms of the comparison of v with $g + \frac{\sigma^2}{4c}$. If it is costless to change the search costs or the informativeness of search, a seller would prefer to have $\overline{U}$ as high as possible (through lower search costs or greater informativeness of search, being in the region of v small), as in that case the expected profit converges to $\frac{\overline{U}}{8}$, while if the seller reduced $\overline{U}$ the highest profit she could get would be $\max[v-g, 0]$.

Therefore, the comparative statistics on social welfare and consumer surplus are the same as those on the seller's profit, which means that social welfare can potentially decrease with lower search costs and higher search informativeness. Note that when the seller charges an optimal price, consumer surplus no longer always increases with search informativeness and decreases with search costs. This is because more search could lead to a higher price that the consumer has to pay.

PROPOSITION 5: *With endogenous pricing, consumer surplus decreases with lower search costs if* $v \in (g + \frac{\sigma^2}{4c}, g + \frac{3\sigma^2}{4c})$.

We restricted attention to the case of a monopolist. As discussed below it would be interesting to consider the case of competition with learning about multiple alternatives. Note that the case of competition where only one alternative needs to be learned can be considered by the model above in a relatively straightforward way.

5. DISCOUNTING

In the model above, search is considered to be done relatively quickly so that issues of discounting the future payoffs because of the time duration of search were not important. However, in some search problems, the delay in obtaining the potential utility of purchasing the product could actually matter. For example, the consumer may want to purchase a fiction book. If he delays his purchase, with information arriving slowly, the potential utility of reading the book is delayed, and therefore less valued. For such problems it is important to discount the future potential payoff of purchasing the product. As a result, it also becomes important for a seller to understand how discounting would affect the purchase and exit thresholds, which directly affect her profit.

To incorporate discounting, equation (1) becomes now $V(u,t) = -c\,dt + e^{-r\,dt}EV(u + du, t + dt)$, where r is the continuous-time discount rate. Doing the Taylor expansion on this equation as above we have that equation (3) now becomes

$$rV = -c + V_t + \frac{\sigma^2}{2}V_{uu}. \tag{15}$$

Noting that the problem does not change with time, we again have $V_t = 0$. We can then obtain, by solving the differential equation (15), that $V(u)$ satisfies

$$V(u) = A_1 e^{\frac{\sqrt{2r}}{\sigma}u} + A_2 e^{-\frac{\sqrt{2r}}{\sigma}u} - \frac{c}{r}, \tag{16}$$

where A_1 and A_2 are two constants that are obtained together with $\overline{U}$ and $\underline{U}$ with conditions (5) and (6). In comparison with the no-discounting case, one can also obtain the following result (the proof is in the Appendix).

PROPOSITION 6: *With discounting, the stopping boundaries are*

$$\overline{U} = \sqrt{\frac{c^2}{r^2} + \frac{\sigma^2}{2r}} - \frac{c}{r}, \text{ and } \underline{U} = \left(\sqrt{\frac{c^2}{r^2} + \frac{\sigma^2}{2r}} - \frac{c}{r}\right) - \frac{\sigma}{\sqrt{2r}} \log\left(\sqrt{\frac{r\sigma^2}{2c^2}} + \sqrt{1 + \frac{r\sigma^2}{2c^2}}\right).$$

The purchase threshold ($\overline{U}$) *is smaller than the absolute value of the exit threshold* ($\underline{U}$).

That is, with discounting we have $\overline{U} < -\underline{U}$. The intuition is that with discounting the future benefits of searching are smaller, and therefore the consumer wants to stop search sooner: $\overline{U}$ is lower, and $\underline{U}$ is higher. This effect is more important when the consumer is more likely to buy the product, i.e., when $u > 0$, as there is a clear benefit of buying that is being discounted, and therefore the consumer does not need as much positive information about the product to decide to stop searching and buy the product. That is, with discounting $\overline{U}$ goes down by more than $\underline{U}$ goes up. While the consumer is always trading off search cost with the potential benefit from buying the product, the discount rate is applied to an immediate payoff $\overline{U}$ when u is near the purchase threshold, and it is applied to more distant payoffs when u is near the exit threshold. An interesting observation is that an ex-ante indifferent consumer with $v = 0$ buys the product with probability of .5 in the no-discounting case, but buys the product with a higher probability if he discounts the future. As an example, consider the case where $\sigma^2 = .5, c = .1$, and $r = .1$. Without discounting we have $\overline{U} = -\underline{U} = 1.25$. With discounting with the discount rate $r = .1$, we have $\overline{U} = .87$ and $\underline{U} = -1.09$.

It is interesting to explore the implications of consumer discounting for optimal pricing. Suppose for now that the seller is infinitely patient (that is, the seller does not do any discounting). Along the lines of the analysis above we can obtain the probability of purchase as $\frac{v-p-\underline{U}}{\overline{U}-\underline{U}}$, for $\underline{U} \leq v - p \leq \overline{U}$. The optimal price can be obtained as $p^* = \frac{v+g-\underline{U}}{2}$ for $v < g + 2\overline{U} - \underline{U}$ and $p^* = v - \overline{U}$ for $v \geq g + 2\overline{U} - \underline{U}$. Comparing with the no-discounting case we can then see that the optimal price is lower than in no discounting if there is search in equilibrium (if $v < g + 2\overline{U} - \underline{U}$), and is greater than in no discounting if there is no search in the equilibrium under no discounting (if $v > g + 3\frac{\sigma^2}{4c}$).[21] Figure 3 illustrates the comparison between optimal prices with and without discounting. When there is search in equilibrium, the seller has to lower her price as consumers, discounting the future benefits, end up having less overall expected value for the product. When there is no search in equilibrium, the seller can charge a higher price, as the purchase threshold of the consumer ($\overline{U}$) is now lower.

When the seller also discounts future payoffs (such that she prefers the consumer to purchase earlier), she charges an even lower price when there is search in equilibrium, and the price remains unchanged when there is no search in equilibrium. When there is search in equilibrium, the seller chooses to lower her price further in order to induce the consumer to purchase earlier.

Since both discounting and a positive search cost motivate the consumer to make a purchase decision early, it is interesting to consider how these two forces work differently. For this purpose, let us consider the extreme case in which the search cost is zero, but there is discounting. Note that this extreme case and our baseline model are motivated by different search situations in the real world. The extreme case ($c = 0, r > 0$) is suitable for products for which information flows freely (no information processing costs) but takes time to arrive, while in the baseline model ($c > 0, r = 0$), the consumer needs to actively invest resources in information acquisition.[22]

Letting the search cost c approach zero in Proposition 6, one can see that in the extreme case of $c = 0$ and $r > 0$, the stopping boundaries become $\overline{U} = \frac{\sigma}{\sqrt{2r}}$ and $\underline{U} = -\infty$. Similar to the baseline model, the purchase threshold increases with search informativeness. Instead of decreasing with the search cost, the purchase threshold now decreases with the discount factor, capturing the intuition that when the consumer is less patient (greater r), he has a lower standard for purchase. A striking feature of the new

[21]The overall condition is that the optimal price is lower than under no discounting if and only if $v < g + 2\overline{U} + \frac{\sigma^2}{4c}$.
[22]We are grateful to a reviewer for pointing us to this comparison.

stopping boundaries is that the exit threshold is nonexistent: as search is costless, the consumer would always continue to search if he has not decided to buy the product. In this case, if the seller does not discount future profits, she should charge an extremely (infinitely) high price, as a higher price does not lower the probability of purchase. When the seller discounts the future, the seller has now to worry about when the consumer makes the purchase. One can compute for this extreme case of $c = 0$ that the probability of purchase before time T for $v \leq \overline{U}$ is (see Appendix for the derivation)

$$\Pr(v, T) = \int_0^T \frac{\overline{U} - v}{\sqrt{2\pi\sigma^2 t^3}} e^{-\frac{(\overline{U}-v)^2}{2\sigma^2 t}} \, dt. \tag{17}$$

Figure 4 illustrates how $\Pr(v, T)$ evolves with v and T.

A seller with discount rate r_S, maximizes then $\max_p (p-g) \int_0^\infty e^{-r_S t} \frac{\overline{U}-v+p}{\sqrt{2\pi\sigma^2 t^3}} e^{-\frac{(\overline{U}-v+p)^2}{2\sigma^2 t}} \, dt$. An increase in the seller's discount rate, r_S, leads the seller to reduce the price, as it cares more about selling sooner, which can be facilitated by increasing $v - p$. Note that if the consumer's discount rate r is connected to the seller's discount rate, then an increase in the discount rate reduces $\overline{U}$, which increases the probability of purchase before any time period T. This counteracts the effect to decrease prices when the seller's discount rate increases. This effect can, however, be seen as smaller than the effect of r_S, as for any $r_S > 0$, the optimal price is bounded for any $r \geq 0$, while for any $r > 0$, the optimal price is unbounded for $r_S \to 0$. As a result we may expect that if $r = r_S$ the pressure to sell sooner can be dominant when the discount rate increases, and hence the seller would tend to lower the price. For the particular model considered we find this to hold for the numerical examples examined, which are presented in Figure 5. The effects of discounting and of the extensions discussed below are summarized in Table 1.

6. Extensions

6.1. Signals on the Value of a Product

An alternative way to model the search for information is to have the consumer receive, at each search stage, an independent signal of the value of the product to him. That is, one could have that at any search opportunity i the consumer receives a signal S_i which is equal to the true valuation of the product for that consumer, which we can denote as U, plus an error $\varepsilon_i, S_i = U + \varepsilon_i$. The consumer does not know U but tries to infer U from the sequence of signals received and his priors on U. Suppose that the consumer, before obtaining any signal, has a normal prior on U with mean v and variance e^2, and that the errors ε_i are normally distributed with mean zero and variance s^2.

We can then write the expected utility of buying the product, u, after observing t signals as,

$$u(t) = \frac{\frac{v}{e^2} + \sum_{i=1}^{t} \frac{S_i}{s^2}}{\frac{1}{e^2} + \frac{t}{s^2}}. \tag{18}$$

In comparison to the baseline model where $u(t) - u(t-1) = x_t$, we now have

$$u(t) - u(t-1) = x_t = (S_t - u(t-1)) \frac{z_{t-1}^2}{z_{t-1}^2 + s^2}, \tag{19}$$

where z_{t-1}^2 is the variance of the posterior beliefs about U given the information obtained after checking $t-1$ signals. One can compute $z_t^2 = 1/(\frac{1}{e^2} + \frac{t}{s^2})$, and $z_t^2 = \frac{z_{t-1}^2 s^2}{z_{t-1}^2 + s^2}$ as the equation of motion for z_t^2 for all t. As above, the expected value of $u(t) - u(t-1)$ is zero, but the variance of $u(t) - u(t-1)$, given the information obtained until $t-1$, can now be shown to be $\sigma_t^2 = \frac{z_{t-1}^4}{z_{t-1}^2 + s^2}$, which is decreasing in the number of signals already checked, t.

In order to obtain a continuous version of this model, we can make the precision of the signals (the inverse of the variance of the signals) go to zero, such that one unit of signals allow the consumer to update his expected utility of buying the product, u, with the statistical properties of the discrete case described above. Suppose that a signal received of size dt has a variance s^2/dt, such that the update of the variance of the posterior beliefs over one unit of signals is equivalent to the update due to one signal of variance s^2. Therefore, the variance of the consumer's posterior beliefs about U after obtaining t units of signals continues to be $z_t^2 = 1/(\frac{1}{e^2} + \frac{t}{s^2})$. Moreover, we can obtain that the variance of the change in the expected utility of buying the product with dt signals, the variance of $u(t+dt) - u(t)$, is equal to $\sigma_t^2 \, dt = \frac{z_t^4}{z_t^2 + s^2/(dt)} = \frac{z_t^4}{s^2} \, dt$, for dt converging to zero.

One can then obtain in the continuous version that the change in expected utility of buying the product can be described as the Brownian motion $du = \sigma_t \, dw$, where the variance of the Brownian motion decreases now with the number of signals received.

Applying the analysis above we can then obtain that the expected utility of the search process given the expected utility of buying the product, u, and after observing t signals, has to satisfy equation (3). Note that now $V_t \neq 0$, that is, the expected utility of the search process depends on the number of signals checked, t, in addition to the current expected utility of buying the product, u.

Note that now the boundary conditions also depend on the number of signals checked. That is, the upper and lower levels of u that trigger either a purchase or a no-purchase decision now depend on the number of signals checked, $\overline{U}(t)$ and $\underline{U}(t)$. In particular, the boundary conditions are now represented by

$$V(\overline{U}(t), t) = \overline{U}(t) \quad \text{and} \quad V_u(\overline{U}(t), t) = 1; \tag{20}$$

$$V(\underline{U}(t), t) = 0 \quad \text{and} \quad V_u(\underline{U}(t), t) = 0. \tag{21}$$

For this case, one can show that, in general, the optimal search process is characterized by the consumer continuing to search as long as the expected utility of buying the product after observing t signals is in between $\underline{U}(t)$ and $\overline{U}(t)$. Furthermore, one can show that $\overline{U}(t)$ is decreasing, and $\underline{U}(t)$ is increasing, in the number of signals checked, t. The intuition is that as the consumer checks more signals, he knows that the new signal is going to be less informative on the expected utility of buying the product and, therefore, the consumer requires less in terms of the expected utility of buying the product in order to either stop searching and buy the product, or stop searching and decide not to buy the product. Signals become less informative as t increases. As the early signals are the most informative, the consumer starts out being hopeful, and loses hope quickly if the signals turn out to be unfavorable. As t approaches infinity, the signals are hardly informative, and the consumer makes a decision even if the expected utility of buying is relatively close to zero. Furthermore, as the variance of the posterior beliefs approaches zero at a decreasing rate, we can obtain that the purchase and exit thresholds also approach zero at a decreasing rate. That is, the purchase threshold, $\overline{U}(t)$, is convex, and the exit threshold, $\underline{U}(t)$, is concave, both approaching zero as t goes to infinity. Figure 6 presents the boundaries $\overline{U}(t)$ and $\underline{U}(t)$ that determine the optimal search behavior for an example with $c = 0.01$ and $s = e = 1$.

6.2. Heterogeneous Importance of Attributes

In the base case all attributes were considered to have the same importance, or the consumers would not know the importance of an attribute before checking it. Consider now that the consumers know which attributes are more important before starting to search. The consumers in this case would check first the more important attributes (more information per search cost incurred), and the change in expected utility of buying the product can be described as the Brownian motion $du = \sigma_t dw$, where the variance of the Brownian motion decreases now with the number of attributes checked. One can then apply the analysis of subsection 6.1 above, where the expected utility of buying the product has to satisfy (3), and boundary conditions (20)-(21).

Similarly to the extension above, one can show that the purchasing threshold, $\overline{U}(t)$, is decreasing in the number of attributes checked, and the exit threshold, $\underline{U}(t)$, is increasing in the number of attributes checked. The intuition is that after checking a certain number of attributes, the remaining attributes have less importance, and if the consumer has not chosen to purchase or stop searching without buying before, then he must be willing to do so with lower thresholds in absolute value if less information is available.

Another related case is when the realization of the fit of the attributes is correlated across attributes. As argued in the base case above, the contribution of each attribute checked to the change in expected utility of buying the product has to be independent from the previous changes resulting from previous attributes checked as information in the attributes represents "news" to the consumer. "News" are by definition independent from previous news. In the base case above this was presented as the fit of all attributes being independent from each other. Consider now the case in which every attribute fit contributes equally to the total utility of buying the products, but in which the attribute fit is correlated across attributes. Then, in this case, when checking the first attribute the consumer, because of its correlation with the unchecked attributes, learns also something about the fit of those unchecked attributes. That is, the contribution of the checked attribute to the change in expected utility of buying the product is greater than that attribute fit alone, and includes part of the unchecked attributes. When the consumer goes on to check further attributes, the contribution of those attributes to the change in the expected utility of buying the product has now to be smaller (in absolute value). That is, the change on the expected utility of buying decreases with the number of attributes checked, and we are in the same situation as above with σ_t decreasing in t.

In terms of pricing by the seller, note that as now the thresholds are more extreme at the beginning of the search process, it is more costly for the firm to price such that the consumer does not search and purchases the product immediately. The optimal price is hence more likely to involve the consumer doing some search. As above, if the firm chooses a low price the consumer is more likely to purchase after a small amount of search. If the firm chooses a high price the consumer is less likely to purchase. Furthermore, we now have that extreme prices (too high or too low) lead to even less search (with a low price leading to purchase, and a high price leading to exit without purchase), while intermediate prices lead to more search.

6.3. Finite Mass of Attributes

In the baseline model above it was also considered that there was an infinite mass of attributes, such that the consumer, after observing t attributes, was in the same conditions in terms of attributes that he could check as if he had observed t' attributes for any t and t'. This assumption could be possibly justified by the fact that for many product categories there is a huge amount of information available online and a consumer never needs to worry about running out of attributes to check. For products that are relatively simple, it may be true that they have relatively fewer attributes. To incorporate such situations we extend our baseline model in this section to have a finite number of attributes.[23]

The analysis of this situation is similar to the one considered above leading to the partial differential equation (3). However, in this case we do not have $V_t = 0$, as the situation of the consumer changes when the number of attributes checked is different. That is, after checking more attributes, there are fewer attributes that the consumer can check. Let T represent the total mass of attributes that the consumer can check.

To solve this problem we have again that the boundary conditions on u are functions of the number of attributes checked, and satisfy (20)-(21), as in the subsection 6.1 above. Furthermore, we have an additional boundary condition that

$$V(u, T) = \max[u, 0]. \tag{22}$$

One can obtain that $\overline{U}(t)$ is decreasing, and $\underline{U}(t)$ is increasing, in the number of attributes checked, t. That is, as the number of checked attributes increases, the consumer is less demanding on the expected utility of buying the product, u, to decide to either stop searching and buy the product, or stop searching and not buy the product. Furthermore, one can obtain that $\overline{U}(T) = \underline{U}(T) = 0$. That is, after checking all the attributes the consumer decides to buy the product even if the utility of buying the product is barely above zero. Note also that each attribute checked affects the expected utility of buying the product in the same way, such that the purchase and exit thresholds do not change too much for t small, but start changing more as t approaches T. That is, one can show that in this case the purchase threshold, $\overline{U}(t)$, is concave, and the exit threshold, $\underline{U}(t)$, is convex. Figure 7 presents the boundary conditions that characterize the optimal consumer search behavior for an example with $\sigma = T = 1$ and with $c = .01$ or $c = 0.1$. The figure also illustrates that the search region enlarges as search costs decrease.

6.4. Choosing the Search Intensity

When the search for information takes time, a consumer may often be able to choose how many attributes to search per unit of time, with more attributes being searched per unit of time being more costly. For example, a consumer could decide not to search actively, and simply obtain information that comes to him without effort or, alternatively, to search intensively for information. It is interesting to understand what factors influence the consumer's choice of search intensity, which naturally affects the stopping boundaries and hence the corresponding price and profit levels.

In terms of the model above one can then endogenize σ^2, the informativeness of the search process at each point in time. Greater informativeness, greater σ^2, would mean that the consumer incurs higher search costs, c, which we now write as $c(\sigma^2)$, as the search costs depend on σ^2, with $c'(\sigma^2), c''(\sigma^2) > 0$. Consider first the case with no discounting.

[23]See also the two-sided case in Roberts and Weitzman (1981) for the case with direct specific assumptions on how the posteriors change through search in a setting of research and development for a project.

From (2) and (3), maximizing $V(u)$ with respect to σ^2 leads to

$$c'(\sigma^2) = \frac{1}{2}V_{uu}. \tag{23}$$

Putting this equation together with (4) one obtains the condition that determines σ^2 as

$$\sigma^2 c'(\sigma^2) = c(\sigma^2). \tag{24}$$

Note that for this no-discounting case we obtain that the optimal intensity of search, σ^2, is independent of the expected valuation of buying the product, u. For example, for the case where $c(\sigma^2) = a_0 + \frac{a_1}{2}\sigma^4$, we obtain the optimal $\sigma^2 = \sqrt{\frac{2a_0}{a_1}}$.

With discounting, we have $\sigma^2 V_{uu}$ increasing in u, by (15), and given that $V(u)$ is increasing in u. Then, by (15) and (23) we have that the optimal σ^2 is increasing in u. That is, with discounting, a consumer invests more in search intensity the closer he is to buying the product, which occurs when u is greater. The closer the consumer is to buying the product, the more he invests in learning about the product, as the benefits of purchase are discounted. Note that in terms of pricing, a low price increases the expected utility of buying the product, which leads to more intense search, while a high price may lead to less intense search. This endogeneity of search can therefore be a force towards lower prices, pushing the consumer to search more intensively.

6.5. Social Learning

Another interesting possibility to consider is the case where consumers arriving later to the market can obtain information not only from their own information acquisition but also from observing what previous consumers have done (e.g., Bikhchandani et al. 1992, Zhang 2010). This can be seen as generating an integrated treatment of information acquisition and (potential) herding behavior.[24] We discuss these effects under the two separate cases of sequential and simultaneous entry of consumers in the market.

Consider a variation of the baseline model in which there is a sequence of consumers acquiring information about whether to purchase a product, only one consumer searching at a time, and later consumers can see the purchase or no-purchase decision of previous consumers. In order for this possibility to affect the results above suppose that a fraction $q \in (0, 1)$ of the attributes affect all consumers equally. Note first that the purchase and exit thresholds for each consumer do not change from consumer to consumer, as the process for each consumer starts from his expected utility of buying the product, and evolves as in the base case above. Suppose also that, in order to simplify the analysis, the expected utility of buying the product is zero if a consumer has not observed the decision-making of any other consumer.[25]

Consider the first consumer. The probability of that consumer purchasing the product is 1/2. Consider now the second consumer. If the first consumer bought the product, the second consumer knows that the first consumer reached an expected utility of buying the product of $\overline{U}$. As a fraction of q attributes are shared across consumers, the second consumer knows that his expected utility, if

[24] We thank a reviewer for the suggestion to explore this extension.

[25] It would be interesting to explore what happens if consumers have different starting valuations for the product. If the consumers do not know each other's evaluations, the main effects in this section should still hold.

the product were purchased immediately without search, would be $q\overline{U}$. This consumer needs now fewer positive signals to buy the product than the first consumer.[26] The probability of the second consumer buying the product after search, after the first consumer bought the product, would be $\frac{1+q}{2} > \frac{1}{2}$. Similarly, the probability of the second consumer buying the product, after the first consumer having decided not to buy the product, can be computed to be $\frac{1-q}{2}$.

Consider now the third consumer. If the first two consumers bought the product, the third consumer knows that the first consumer obtained $\overline{U}$ worth of positive attributes minus negative attributes, and that the second consumer obtained $(1-q)\overline{U}$ worth of positive attributes minus negative attributes. The expected utility of buying the product without further search is then $q(2-q)\overline{U}$. The probability of the third consumer purchasing the product is then $\frac{1+2q-q^2}{2} > \frac{1+q}{2}$. More generally, if the first n consumers decided to purchase the product, the probability of the $(n+1)$-th consumer deciding to buy the product is $1 - \frac{(1-q)^n}{2}$, which is increasing in n. Note that this leads to a probabilistic herding behavior: The greater the number of consecutive consumers who buy the product, the less demanding is the next consumer in terms of attributes with good fit to buy the product.

Consider now the situation of the third consumer after the first consumer bought the product, and the second consumer decided not to buy the product. The third consumer knows that the first consumer obtained $\overline{U}$ worth of positive attributes minus negative attributes, and that the second consumer obtained $(1+q)\overline{U}$ worth of negative attributes minus positive attributes. The expected utility of buying the product without further search is then $-q^2\overline{U}$. The probability of the consumer purchasing the product is then $\frac{1-q^2}{2} < \frac{1}{2}$. That is, even though the third consumer observes one consumer buying the product, and another consumer not buying the product, the consumer is less likely to buy the product than if he did not observe the purchasing behavior of any other consumer. The intuition is that the second consumer had to receive a larger amount of negative news in order to decide not to buy the product than the amount of positive news received by the first consumer.

Similarly, one can obtain that if the first consumer does not buy the product and the second consumer buys the product, then the probability that the third consumer buys the product is $\frac{1+q^2}{2} > \frac{1}{2}$. Meanwhile, if neither of the first two consumers buy the product, then the third consumer buys the product with probability $\frac{(1-q)^2}{2} < \frac{1}{2}$. For completeness, we can derive the probability distribution recursively for the net amount of positive attributes that the n-th consumer can infer from the cumulative purchase decisions of the earlier consumers. Denote this net amount by $a_n\overline{U}$. We have

$$a_n = \begin{cases} (1-q)a_{n-1} + 1, & \text{with probability } \frac{1+qa_{n-1}}{2}; \\ (1-q)a_{n-1} - 1, & \text{with probability } \frac{1-qa_{n-1}}{2}. \end{cases}$$

The first line in the equation above corresponds to the case when the $(n-1)$-th consumer made a purchase, and the second line corresponds to the case when that consumer exited without a purchase. Using this recursive formula, and the assumption that $a_0 = 0$, one can write out the probability distribution of a_n for any n. Note that there would be 2^n possible values for a_n, with the probability for each of these values obtainable also from the equation above. Importantly, notice that for any given a_{n-1}, we have $E(a_n) = a_{n-1}$. Applying this property we can obtain that $E(a_n) = a_0$ for any given n. That is, in expectation, the existence of sequential search does not change the expected probability

[26]Note that if all consumers value all attributes equally, $q = 1$, then all consumers after the first consumer do what the first consumer had done, without search. This also holds if there is a known heterogeneity of attribute importance.

of purchase for any consumer, although it does create a distribution of potentially inefficient purchase and exit scenarios. As a result, if the seller needs to commit to a price before the first consumer starts to search for information, and if the optimal price is such that there is always positive search by all consumers in equilibrium, then the optimal price she would set should equal to that in the baseline model without social learning, as the probability of buying is linear in a_n in that case. On the other hand, if the seller has the ability to adjust the price over time, then she would adjust the price to reflect the change in a_n. That is, she would increase the price when many consumers decide to buy the product, and decrease the price otherwise.

Consider now the other extreme case where all consumers search simultaneously. The basic knowledge structure is the same: a fraction q of attributes affect all consumers equally, and each consumer observes when any other consumer stops searching, and whether he buys or not. This framework allows us to discuss two further issues: complete herding behavior and free riding.

Suppose that n consumers start searching simultaneously. Consider the moment where one of them stops searching. Suppose that he buys. Then, all other consumers discretely increase their current expected utility by $q\overline{U}$. Thus, all consumers, if any, whose current expected utility is at least $(1-q)\overline{U}$, buy immediately.

But, this may trigger a cascade. If m consumers buy in the first wave, all remaining consumers discretely increase their expected utility by $q(1+(1-q)m)\overline{U}$, and some additional consumers may decide to buy. In particular, if $m \geq \frac{2-q}{q(1-q)}$, all other consumers will buy in this second wave, regardless of their private information.[27] This is complete herding. As above, there may be a benefit of the seller adjusting the price over time to reflect changes in the net amount of the revealed positive information and also capture further consequences on herding.

Another possibility opened up by social learning with simultaneous search is that some consumers may decide not to search, simply free riding on others' information. A full analysis of free riding is beyond the scope of this paper, but the intuition should be clear. Suppose that the number of consumers is very large. Then, it is likely that herding will follow, as soon as one consumer decides either to buy or not to buy. Moreover, with a large number of consumers, the first such decision will come sooner. Then, any potential consumer may consider the alternative of free riding: not searching and simply observing others' decisions. Free riding helps him save on the search costs. Obviously, if each consumer has the possibility to decide on whether or not to search, a multiplicity of equilibria will arise, but all will require a certain number of consumers to search and the others to free ride.

7. Concluding Remarks

We have studied two central topics in this paper. First, how does a consumer optimally search for product information when his valuation of the product is uncertain? Second, given the consumer's optimal search strategy, what can the seller do to maximize her expected profit? We provide a parsimonious model of consumer search for gradual information that captures the benefits and costs of search, determining optimally the thresholds when the consumer decides to stop searching.

In summary, the consumer searches for further information whenever his valuation of the product is in between two boundaries. When his valuation first hits the upper bound, he stops searching and

[27]The expected utility of these other consumers was at least $\underline{U}$, as they had not left yet.

buys the product. When his valuation first hits the lower bound, he stops searching without buying the product. The two bounds are determined by the informativeness and marginal cost of search. In particular, when search becomes more informative and less costly, the purchase threshold, or the upper bound, increases.

As for the seller, when the stand-alone valuation of the product is not too high, the optimal price that she should set takes into account both the stand-alone valuation (expected utility before search) and the upper-bound valuation (expected utility after search given that the consumer purchases). For a high stand-alone expected utility, the seller chooses the price such that the consumer's net utility prior to search is exactly at the purchase threshold, hence the consumer buys the product immediately.

While providing an understanding of consumer behavior and implications for a firm's marketing strategies, the paper presents a tractable foundation to understand information-search behavior in a stochastic environment. From this perspective, one can extend the model in several directions. We discuss extensions to incorporate discounting, signals of the overall value of the product, endogenous search intensity, and social learning.

It would also be interesting to investigate the seller's decision on how much information to provide. Another interesting issue to consider is that the consumer's search prior to purchase may help him in the post-purchase usage. This possibility would require modeling the consumer's learning process during product usage, which is not considered here. This extra-benefit of search may potentially lead to more search prior to purchase. Other dimensions that could potentially be interesting to study would be the case of competition, and of a seller selling multiple products.

APPENDIX

Derivation of Equation (7): Equation (2) must hold for all sufficiently small dt (using Itô's Lemma). Since the mass of attributes available to check is unbounded by assumption, the mass of attributes already checked does not matter:

$$V_t = 0. \tag{i}$$

From (3) and (i) we can then obtain (4). In addition, we also have a set of boundary conditions (5) and (6) as described in the text.

To get some intuition on the high-contact condition (for example, at $\overline{U}$, $V'(\overline{U}) = 1$), note that by searching extra dt attributes when reaching $u = \overline{U}$, a consumer pays a cost $c\,dt$ and gets

$$\frac{1}{2}(\overline{U} + E[du|du \geq 0]) + \frac{1}{2}E[V(\overline{U} + du)|du < 0].$$

By a Taylor approximation, and $V(\overline{U}) = \overline{U}$, we have $V(\overline{U} + du) = \overline{U} + V'(\overline{U}^-)\;du$ for $du < 0$, where $V'(\overline{U}^-)$ represents the first derivative of $V(u)$ just below $\overline{U}$. Then, using the fact that the distribution of du is symmetric, and $E[du|du > 0] = \sigma\frac{\sqrt{dt}}{\sqrt{2\pi}}$, we have that the benefit over $\overline{U}$ of a consumer searching dt more attributes is $\frac{1}{2}[1 - V'(\overline{U}^-)]\sigma\frac{\sqrt{dt}}{\sqrt{2\pi}} - c\,dt$. That is, for dt close to zero, it is beneficial for a consumer to keep on searching as long as $V'(\overline{U}^-) < 1$. Since for $u < \overline{U}$, we have necessarily $V(u) > u$ (as otherwise it would be better to stop searching) and consequently, $V'(\overline{U}^-) \leq 1$, it must be that $V'(\overline{U}^-) = 1$. Essentially, the argument states that when the consumer's expected utility, u, walks away from $\overline{U}$, in order for him to be indifferent between continuing to search and purchasing the product right away, the marginal decrease in $V(u)$ when u walks to the left would have to equal the marginal increase in $V(u)$ when u walks to the right. Equations (4)-(6) fully determine equation (7).

PROOF OF PROPOSITION 2: To obtain (9), we can see that

$$\begin{aligned} Pr(u) &= Pr(u \text{ hits } \overline{U} \text{ first } |u) \\ &= Pr(u \text{ hits } \overline{U} \text{ first } |u+z') \cdot \frac{1}{2} + Pr(u \text{ hits } \overline{U} \text{ first } |u-z') \cdot \frac{1}{2} \\ &= \frac{1}{2}Pr(u+z') + \frac{1}{2}Pr(u-z'), \end{aligned} \tag{ii}$$

where $z' = E(du|du>0) = \sigma\frac{\sqrt{dt}}{2\pi}$. From (ii), $Pr(u+z') - Pr(u) = Pr(u) - Pr(u-z')$. Therefore, $Pr(u)$ is linear in u. Since $Pr(\underline{U}) = 0$ and $Pr(\overline{U}) = 1$, $Pr(u) = \frac{u-\underline{U}}{\overline{U}-\underline{U}}$.

PROOF OF PROPOSITION 6: To obtain $A_1, A_2, \overline{U}$, and $\underline{U}$ for the discounting case we can write conditions (5) and (6) as

$$A_1 e^{\frac{\sqrt{2r}}{\sigma}\overline{U}} + A_2 e^{-\frac{\sqrt{2r}}{\sigma}\overline{U}} - \frac{c}{r} = \overline{U}; \tag{iii}$$

$$A_1 \frac{\sqrt{2r}}{\sigma} e^{\frac{\sqrt{2r}}{\sigma}\overline{U}} - A_2 \frac{\sqrt{2r}}{\sigma} e^{-\frac{\sqrt{2r}}{\sigma}\overline{U}} = 1; \tag{iv}$$

$$A_1 e^{\frac{\sqrt{2r}}{\sigma}\underline{U}} + A_2 e^{-\frac{\sqrt{2r}}{\sigma}\underline{U}} - \frac{c}{r} = 0; \tag{v}$$

$$A_1 \frac{\sqrt{2r}}{\sigma} e^{\frac{\sqrt{2r}}{\sigma}\underline{U}} - A_2 \frac{\sqrt{2r}}{\sigma} e^{-\frac{\sqrt{2r}}{\sigma}\underline{U}} = 0. \tag{vi}$$

Solving (iii)-(vi) together gives the solution to $A_1, A_2, \overline{U}$, and $\underline{U}$. Denoting

$$\overline{X} = e^{\frac{\sqrt{2r}}{\sigma}\overline{U}} \text{ and } \underline{X} = e^{\frac{\sqrt{2r}}{\sigma}\underline{U}}, \tag{vii}$$

from (v) and (vi) one can then obtain $A_1 = \frac{c}{2r\underline{X}}$ and $A_2 = \frac{c\underline{X}}{2r}$. Using these in (iv), and denoting $Z = \frac{\overline{X}}{\underline{X}}$, one can then obtain

$$Z = \sqrt{\frac{\sigma^2 r}{2c^2}} + \sqrt{1 + \frac{\sigma^2 r}{2c^2}}. \tag{viii}$$

By the definition of $\overline{X}$: $\overline{U} = \frac{\sigma}{\sqrt{2r}} \log \overline{X}$. Therefore, using (iii),[28] $\overline{U} = \sqrt{\frac{c^2}{r^2} + \frac{\sigma^2}{2r}} - \frac{c}{r}$, and

$$\overline{X} = e^{\sqrt{1+\frac{2c^2}{r\sigma^2}} - \sqrt{\frac{2c^2}{r\sigma^2}}}. \tag{ix}$$

Using this previous condition and the solution for Z one gets:

$$\underline{X} = \frac{e^{\sqrt{1+\frac{2c^2}{r\sigma^2}} - \sqrt{\frac{2c^2}{r\sigma^2}}}}{\sqrt{\frac{r\sigma^2}{2c^2}} + \sqrt{1 + \frac{r\sigma^2}{2c^2}}} \tag{x}$$

[28] It is worth noting that $\lim_{r\to 0} \sqrt{\frac{c^2}{r^2} + \frac{\sigma^2}{2r}} - \frac{c}{r} = \frac{\sigma^2}{4c}$ which is $\overline{U}$ in the no-discounting case.

and[29] $\underline{U} = \left(\sqrt{\frac{c^2}{r^2} + \frac{\sigma^2}{2r}} - \frac{c}{r}\right) - \frac{\sigma}{\sqrt{2r}} \log\left(\sqrt{\frac{r\sigma^2}{2c^2}} + \sqrt{1 + \frac{r\sigma^2}{2c^2}}\right)$. Finally,

$$A_1 = \frac{1}{2r} \frac{\sqrt{\frac{r}{2}}\sigma + \sqrt{c^2 + \frac{r\sigma^2}{2}}}{e^{\sqrt{1+\frac{2c^2}{r\sigma^2}} - \sqrt{\frac{2c^2}{r\sigma^2}}}}, \text{ and } A_2 = \frac{c}{2r} \frac{e^{\sqrt{1+\frac{2c^2}{r\sigma^2}} - \sqrt{\frac{2c^2}{r\sigma^2}}}}{\sqrt{\frac{r\sigma^2}{2c^2}} + \sqrt{1 + \frac{r\sigma^2}{2c^2}}}.$$

The derivation of $V(u)$ would follow accordingly. To show that $\overline{U} < -\underline{U}$ we can just show that $Z > \overline{X}^2$, where $Z = \frac{\overline{X}}{\underline{X}}$ is defined above, and $\overline{X}$ and $\underline{X}$ are defined in (vii). Given that Z is computed to be (viii) and $\overline{X}$ is obtained in equilibrium in (ix), this inequality can be written as

$$\omega + \sqrt{1 + \omega^2} > e^{\frac{2\omega}{1+\sqrt{1+\omega^2}}}, \tag{xi}$$

where $\omega = \sqrt{\frac{r\sigma^2}{2c^2}}$. Taking the logarithm (an increasing function) on both sides of (xi) one obtains

$$\log(\omega + \sqrt{1 + \omega^2}) > \frac{2\omega}{1 + \sqrt{1 + \omega^2}}. \tag{xii}$$

Noting now that the left-hand side of (xii) is equal to the right-hand side of (xii) for $\omega = 0$ (also the result that $\overline{U} = -\underline{U}$ for $r = 0$), and that the derivative of the left hand is greater than the derivative of the right hand side for $\omega > 0$, we have that $\overline{U} < -\underline{U}$ for all $r > 0$.

Derivation of the Probability of Purchase before T under Discounting in the Extreme Case $c = 0$: For the discounting case with $c = 0$ let us compute the probability of consumer starting with expected utility v buying the product before time T. Let us denote this probability as $\Pr(v, T)$.

For $u < \overline{U}$ we have

$$\Pr(u, T) = E \Pr(u + du, T - dt). \tag{xiii}$$

Taking a Taylor approximation on the right-hand side we obtain the partial differential equation

$$\Pr_T = \frac{\sigma^2}{2} \Pr_{uu}, \tag{xiv}$$

where $\Pr_T$ is the partial derivative of $\Pr(u, T)$ with respect to T, and $\Pr_{uu}$ is the second derivative of $\Pr(u, T)$ with respect to u. As boundary conditions we have: (1) $\Pr(u, 0) = 0, \forall u < \overline{U}$, the probability of buying immediately is zero; (2) $\Pr(\overline{U}, T) = 1, \forall T$, when u reaches $\overline{U}$ the probability of buying before T is one; and (3) $\lim_{T\to\infty} \Pr(u, T) = 1, \forall u$ the consumer buys the product with probability one. In order to solve this partial differential equation, we can follow Evans (2010, pp. 44-65). We start with the fundamental solution of (xiv) given that we know $\Pr(u, 0) = 0$, $\Phi(u, T) = \frac{1}{\sqrt{2\pi\sigma^2 T}} e^{-\frac{(\overline{U}-u)^2}{2\sigma^2 T}}$. Define now the function $\psi(u, T) = \sigma^2 \Phi_u$. One can show that $\psi(u, T)$ and $\int_0^T \psi(u, t)\, dt$, which is a convolution with respect to the variable T of $\psi(u, T)$ and the identity function, are also solutions to (xiv). Defining $\Pr(u, t) = \int_0^T \psi(u, t)\, dt$ one obtains (17), and one can check that $\Pr(u, 0) = 0$, $\Pr(\overline{U}, T) = 1$, and $\lim_{T\to\infty} \Pr(u, T) = 1$. The proof of the last equality uses the fact that $\int_0^\infty e^{-ax^2}\, dx = \frac{1}{2}\sqrt{\frac{\pi}{a}}$.

[29]It is worth noting that $\lim_{r\to 0} \frac{\sigma}{\sqrt{2r}} \log\left(\sqrt{\frac{r\sigma^2}{2c^2}} + \sqrt{1 + \frac{r\sigma^2}{2c^2}}\right) = \frac{\sigma^2}{2c}$, so that $\lim_{r\to 0} \left(\left(\sqrt{\frac{c^2}{r^2} + \frac{\sigma^2}{2r}} - \frac{c}{r}\right) - \frac{\sigma}{\sqrt{2r}} \log\left(\sqrt{\frac{r\sigma^2}{2c^2}} + \sqrt{1 + \frac{r\sigma^2}{2c^2}}\right)\right) = -\frac{\sigma^2}{4c}$ which is $\underline{U}$ in the no-discounting case.

REFERENCES

Armstrong M., J. Vickers, and J. Zhou (2009), "Prominence and Consumer Search," *RAND Journal of Economics,* **40**, 209-233.

Armstrong, M., and J. Zhou (2010), "Exploding Offers and Buy-Now Discounts," *working paper,* University College London.

Anderson, S.P., and R. Renault (2006), "Advertising Content," *American Economic Review,* **96**, 93-113.

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

Bakos, J. Y. (1997), "Reducing Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science,* **43**, 1676-1692.

Bergemann, D., and J. Välimäki (1996), "Learning and Strategic Pricing," *Econometrica,* **64**, 1125-1149.

Bergemann, D., and J. Välimäki (2000), "Experimentation in Markets," *Review of Economic Studies,* **67**, 213-234.

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy,* **100**, 992-1026.

Bolton, P., and C. Harris (1999), "Strategic Experimentation," *Econometrica,* **67**, 349-374.

Brynjolfsson E, Y. Hu, and D. Simester (2011), "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales," *Management Science,* **57**, 1373-1386.

Cachon, G., C. Terwiesch and Y. Xu (2008), "On the Effects of Consumer Search and Firm Entry in a Multiproduct Competitive Market," *Marketing Science,* **27**, 461-473.

Chatterjee, R., and J. Eliashberg (1990), "The Innovation Diffusion Process in a Heterogeneous Population: A Micromodeling Approach," *Management Science,* **36**, 1057-1079.

Chen, Y., and K. Sudhir (2004), "When Shopbots Meet Emails: Implications for Price Competition on the internet," *Quantitative Marketing and Economics,* **2**, 233-255.

Cox, J.C., S.A. Ross, and M. Rubinstein (1979), "Option Pricing: A Simplified Approach," *Journal of Financial Economics,* **7**, 229-263.

De los Santos, B., A. Hortacsu, and M.R. Wildenbeest (2010), "Testing Models of Consumer Search using Data on Web Browsing and Purchasing Behavior," *working paper,* University of Chicago.

Diamond, P. (1971), "A Model of Price Adjustment," *Journal of Economic Theory,* **3**, 156-168.

Dixit, A. (1993), *The Art of Smooth Pasting,* Harwood Academic Publishers: Chur, Switzerland.

Dixit, A. and R. Pindyck (1994), *Investment under Uncertainty,* Princeton University Press: Princeton, NJ.

Dumas, B. (1991), "Super Contact and Related Optimality Conditions," *Journal of Economics, Dynamics, and Control,* **15**, 675-685.

Ellison, G., and D. Fudenberg (1995), "Word-of-mouth Communication and Social Learning," *Quarterly Journal of Economics,* **110**, 93-125.

Evans, L.C. (2010), *Partial Differential Equations,* American Mathematical Society: Providence, Rhode Island.

Gal-Or, E., T. Geylani, and A. Dukes (2008), "Information Sharing in a Channel with Partially Informed Retailers," *Marketing Science,* **27**, 642-658.

Hauser, J.R., G.L. Urban, and B.D. Weinberg (1993), "How Consumers Allocate Their Time When Searching for Information," *Journal of Marketing Research,* **30**, 452-466.

Hauser, J.R., and B. Wernerfelt (1990), "An Evaluation Model of Consideration Sets," *Journal of Consumer Research,* **16**, 393-408.

Honka, E. (2010), "Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry," *working paper,* University of Chicago.

Johnson, J.P., and D.P. Myatt (2006), "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review,* **96**, 756-783.

Kuksov, D. (2004), "Buyer Search Costs and Endogenous Product Design," *Marketing Science,* **23**, 490-499.

Kuksov, D. (2006), "Search, Common Knowledge, and Competition," *Journal of Economic Theory,* **130**, 95-108.

Lal, R., and C. Matutes (1994), "Retail Pricing and Advertising Strategies," *Journal of Business,* **67**, 345-370.

Mallik, A.K. (1987), "A Note on the Buyer's Problem," *Annals of Statistics,* **15(3)**, 1329-1331.

Mehta, N., S. Rajiv, and K. Srinivasan (2003), "Price Uncertainty and Consumer Search: A Structural Model of Consideration Set Formation," *Marketing Science,* **22**, 58-84.

Merton, R.C. (1990), *Continuous-Time Finance,* Basil Blackwell: Oxford, U.K.

Moorthy, S., B.T. Ratchford, and D. Talukdar (1997), "Consumer Information Search Revisited: Theory and Empirical Analysis," *Journal of Consumer Research,* **23**, 263-277.

Narasimhan, C. (1988), "Competitive Promotional Strategies," *Journal of Business,* **61**, 427-449.

Ofek, E., and V. Srinivasan (2002), "How Much Does the Market Value an Improvement in a Product Attribute?," *Marketing Science,* **21**, 398-411.

Ottaviani, M, and A. Prat (2001), "The Value of Public Information in Monopoly," *Econometrica,* **69**, 1673-1683.

Putsis, W.P., and N. Srinivasan (1994), "Buying or Just Browsing? The Duration of Purchase Deliberation," *Journal of Marketing Research,* **31**, 393-402.

Rao, R., and N. Syam (2001), "Equilibrium Price Communication and Unadvertised Specials by Competing Supermarkets," *Marketing Science,* **20**, 61-81.

Roberts, K., and M.L. Weitzman (1981), "Funding Criteria for Research, Development, and Exploration Projects," *Econometrica,* **49**, 1261-1288.

Shugan, S.M. (1980), "The Cost of Thinking," *Journal of Consumer Research,* **7**, 99-111.

Stahl II, D. (1989), "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review,* **79**, 700-712.

Sun, M. (2011), "Disclosing Multiple Product Attributes," *Journal of Economics and Management Strategy,* **20**, 195-224.

Varian, H. (1980), "A Model of Sales," *American Economic Review,* **70**, 651-659.

Villas-Boas, J.M. (2004), "Consumer Learning, Brand Loyalty, and Competition," *Marketing Science,* **23**, 134-145.

Villas-Boas, J.M. (2009), "Product Variety and Endogenous Prices with Evaluation Costs," *Management Science,* **55**, 1338-1346.

Weitzman, M.L. (1979), "Optimal Search for the Best Alternative," *Econometrica,* **47**, 641-654.

Wernerfelt, B. (1994), "Selling Formats for Search Goods," *Marketing Science,* **13**, 298-309.

Wolinsky, A. (1986), "True Monopolistic Competition as a Result of Imperfect Information," *The Quarterly Journal of Economics,,* **101**, 493-512.

Zhang, J. (2010), "The Sound of Silence: Observational Learning in the U.S. Kidney Market," *Marketing Science,* **29**, 315-335.

Table 1: Qualitative Takeaways from the Extensions

Extension	Qualitative Takeaways
Discounting	(1) The purchase threshold decreases as the consumers discount more, and it becomes smaller than the exit threshold in absolute value. Consequently, consumers who are a-priori indifferent are more likely to make a purchase. (2) The seller charges a lower (higher) price than in the no-discounting baseline if there is (no) search in equilibrium. (3) When the search cost approaches zero, discounting becomes the major force that incentivizes consumers to make a purchase, and the seller would charge a higher price if it is sufficiently patient.
Independent Signals / Heterogeneous Attributes	(1) The purchase (exit) threshold decreases (increases) with the number of signals/attributes searched. The purchase threshold is convex, the exit threshold is concave, and both thresholds approach zero as the number of signals/attributes searched goes to infinity. (2) The optimal price is more likely to be at an intermediate level, inducing the consumer to do some search, as opposed to sufficiently low to lead to immediate purchase.
Finite Mass of Attributes	(1) The purchase (exit) threshold decreases (increases) with the number of attributes searched. The purchase threshold is concave, the exit threshold is convex and both equal zero as the number of attributes checked reaches its maximum. (2) Hence, after checking all the attributes, the consumer decides to buy the product even if the expected utility is barely above zero.
Choosing the Search Intensity	(1) With no discounting, the intensity of search is independent of the consumer's current expected valuation of the product. (2) With discounting, a consumer searches more intensively when his valuation of the product is higher (closer to the purchase threshold). The seller in this case may want to lower the price to facilitate intensive search.
Social Learning	(1) In sequential search, the greater the number of consecutive consumers that choose to buy the product, the less demanding is the next consumer on the amount of positive information he has to obtain to make a purchase. Also, the more recent decisions may carry larger weight than more remote decisions when influencing the current consumer's decision. (2) In simultaneous search, a single purchase decision may trigger a cascade and lead to complete herding. There may also be incentive for free-riding: a proportion of consumers may choose not to search and simply follow the others' decisions.

Figure 1: An Example of *u(t)*

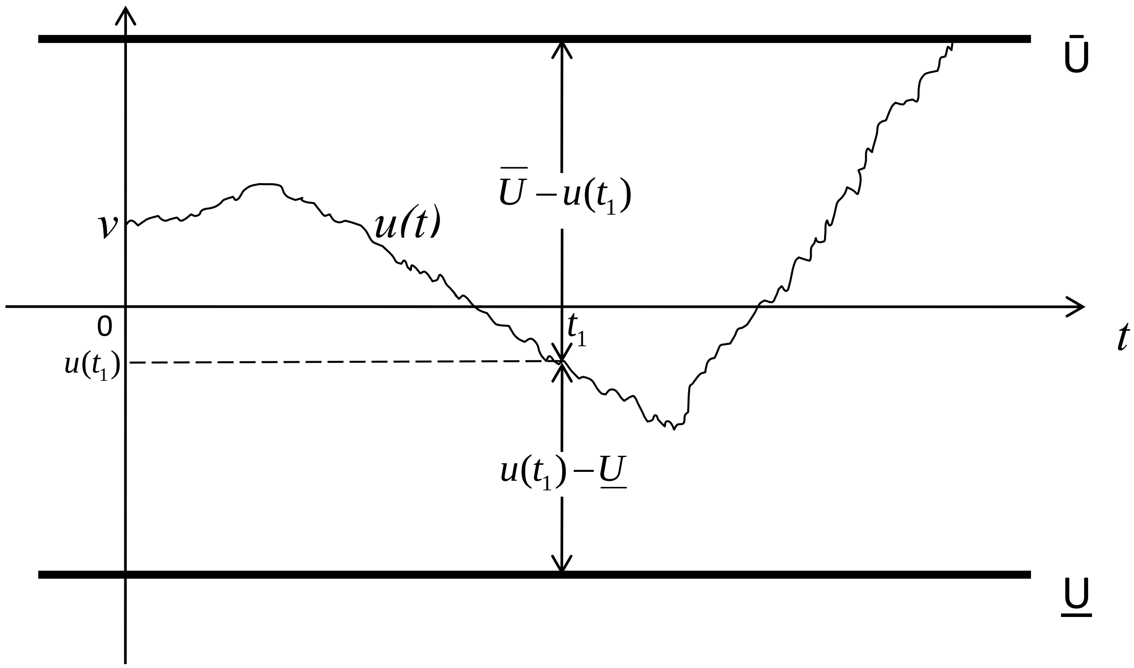

Figure 2: Graph of *V(u)*

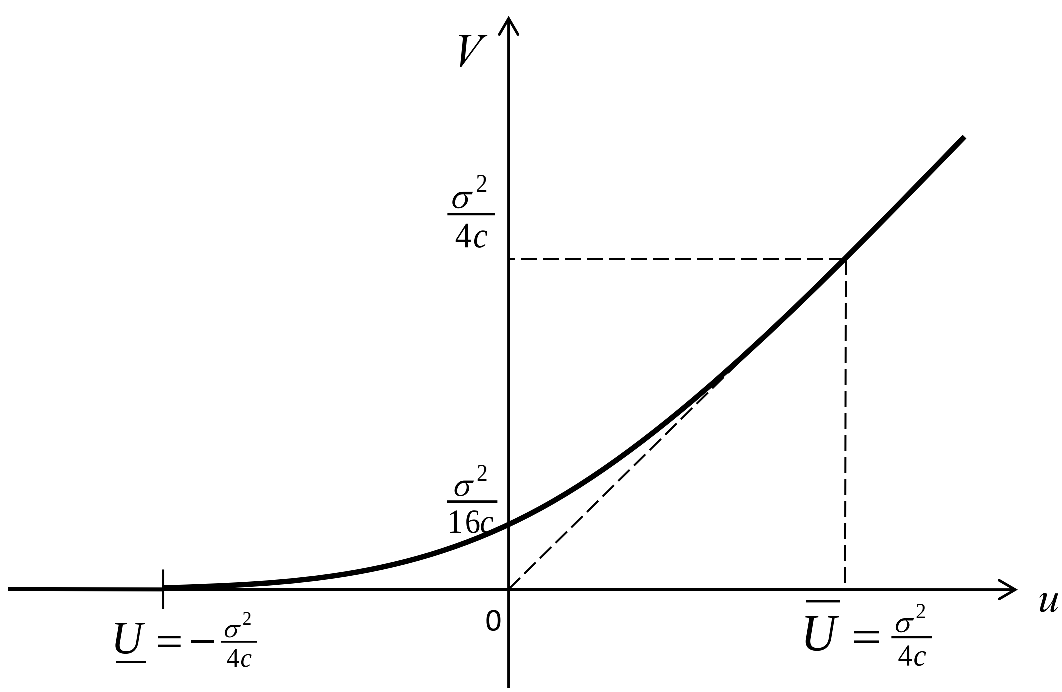

Figure 3: Equilibrium Prices with and without Consumer Discounting

In this figure, c=0.1, σ^2=0.5, r=0.1, and g=0. The dashed line is the equilibrium price without discounting, and the solid line with discounting.

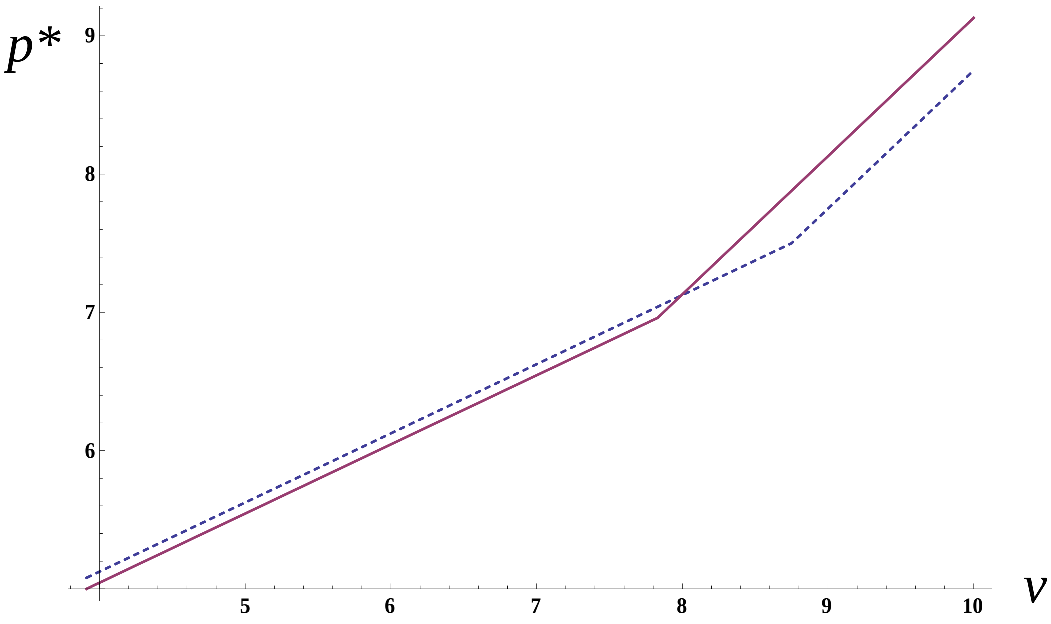

Figure 4: Purchase Probability *Pr(v,T)* with Discounting

In this figure, c=0, σ^2=0.5, r=0.1, and $\overline{U}$=1.58. The four curves (from the right to the left) correspond to, respectively, T=1, 10, 100, and 1000.

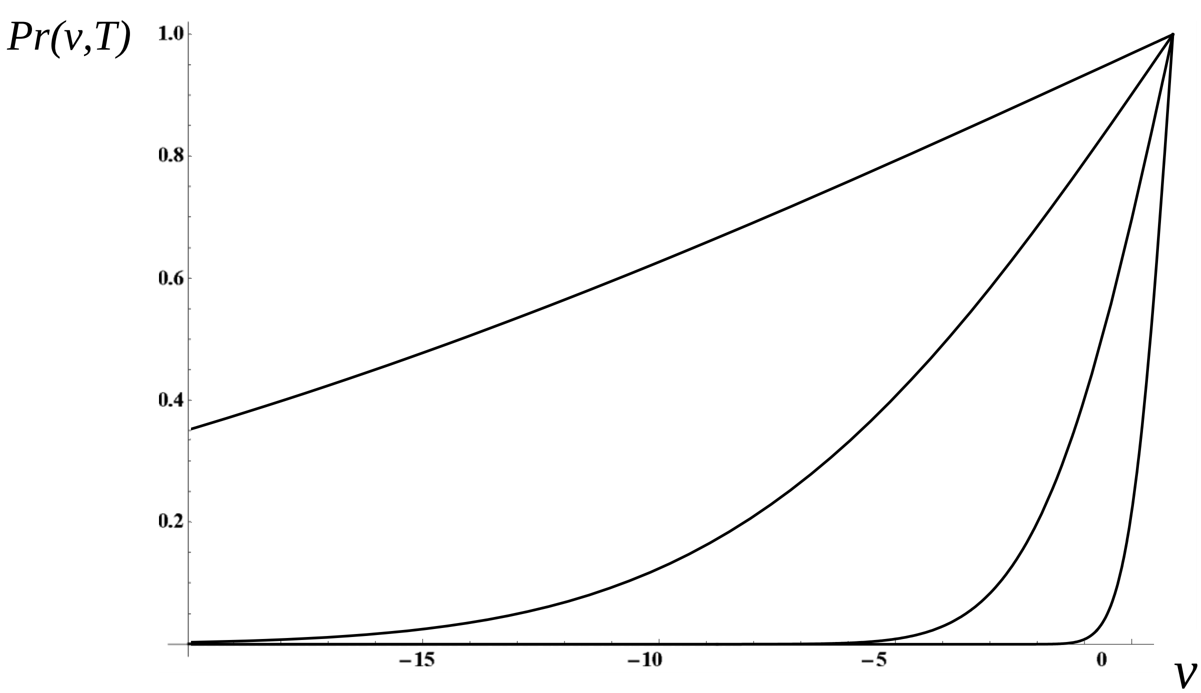

Figure 5: The Optimal Price for Both Consumer and Seller Discounting

In this figure, v=0, c=0, g=0, and r=r_S. On the x-axis we have the discount rate ranging from 0.001 to 1 and on the y-axis we have the optimal price ranging from 0 to 10. The curves (from the bottom to the top) correspond to, respectively, σ = 0.1, 0.5, 1, 3, 5, 10.

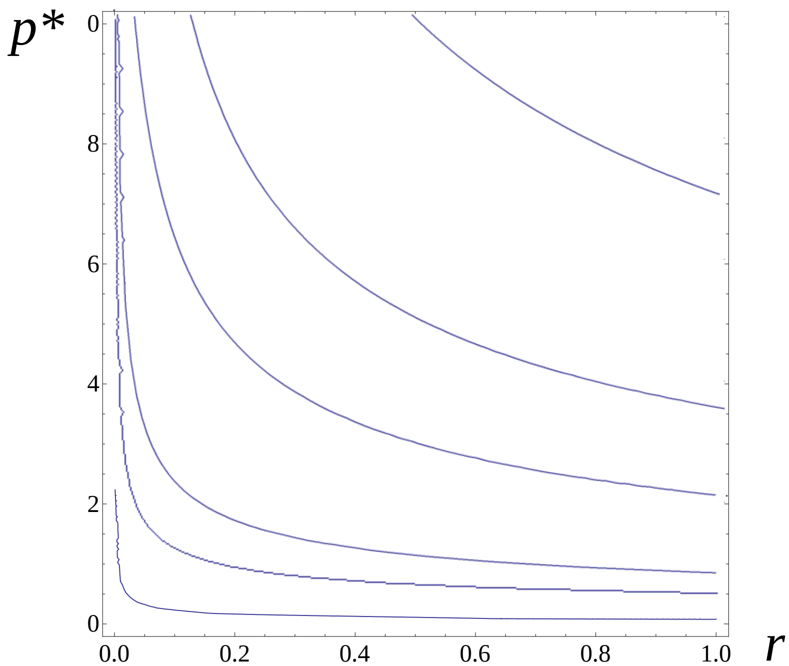

Figure 6: Stopping Boundaries with Independent Signals

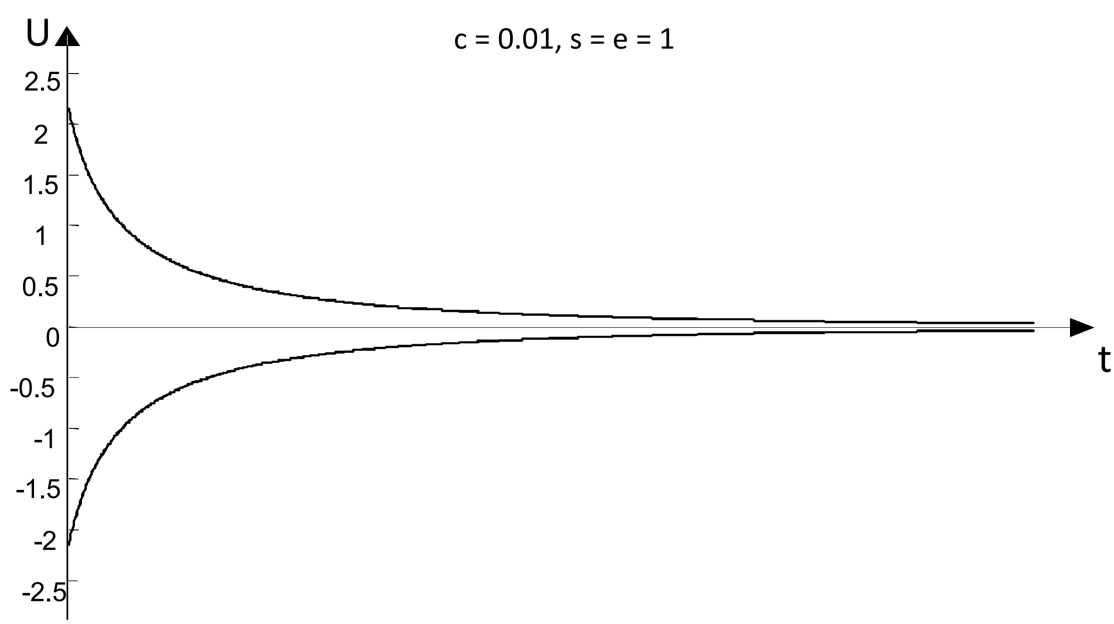

Figure 7: Stopping Boundaries with A Finite Mass of Attributes