Search and Categorization

Fershtman, Chaim and Fishman, Arthur and Zhou, Jidong

Tel Aviv University, Bar-Ilan University, NYU Stern

December 2013

Search and Categorization*

Chaim Fershtman
Department of Economics
Tel Aviv University and CEPR

Arthur Fishman
Department of Economics
Bar-Ilan University

Jidong Zhou
Stern School of Business
New York University

December 2013

Abstract

The internet has not only reduced consumer search costs, but has also enabled more efficient and sophisticated search procedures. For example, online consumers can streamline their search process if appropriately defined categories of products and services are available. This paper proposes a search model with product categories where consumers choose which categories to search and firms respond to such more targeted search by strategically choosing the categories in which to list their products. The analysis focuses on the relationship between category architecture and the type of information which can be credibly disclosed by firms' category choices to consumers.

*We are grateful to Simon Anderson, Mark Armstrong and audiences in various seminars for their comments.

1 Introduction

The classical search paradigm is designed to capture the situation in which consumers visit different stores before settling on a store or brand which best matches their taste and budget. But, in the internet era, actual consumer search is cheaper, more sophisticated and more efficient. Previous research regarding the effect of the internet on consumer search has focused on the fact that online search reduces search costs.[1] But contemporary consumers can take advantage of online resources that list sellers under various categories to narrow down the list of potential sellers from which to further refine their search. Thus the internet has not merely reduced consumers' search costs, but has also changed the way that consumers search. Firms respond to these more sophisticated search procedures and more effectively target the customers they wish to attract, by strategically choosing the sites and product categories under which they are listed.[2]

The structure of the different possible product categories under which firms may be listed - which we refer to as the category architecture - is typically determined by the search intermediaries themselves. For example Yelp determines the different categories of restaurants which are available on its site. The actual choice of category under which the firm is listed may be made either by the search intermediary or by the firm itself. For example martial art schools may list themselves as either Kung fu, Karate, Jiu-Jitsu etc., or architectural firms may describe their service as interior design, commercial planning, residential planning or city planning.[3]

[1] For example, Brown and Goolsbee (2002) provides evidence that the internet reduces search frictions and makes the life insurance market more competitive. Bar-Isaac, Caruana and Cunat (2012) study how the reduction of search costs induces more firms to choose niche product designs and so changes the composition of product types in the market. Goldmanis, Hortascu, Syverson and Emre (2010) study how the reduction of search costs reallocates market shares from high-cost to low-costs producers.

[2] Search with keywords represents an alternative internet search procedure whereby consumers, by looking for certain "keywords", reveal personal information which enables sellers to target consumers more effectively. See, e.g., de Corniere (2013) for a model which studies this type of internet search.

[3] Clearly there are several possible combinations of the two procedures in which the firms apply for certain categories and the search intermediaries need to approve or reject the application.

The paper develops a simple framework to analyze the 'search with categories' setting in which consumers choose the category in which to search and their stopping rule for each category while firms choose, beside product characteristics and prices, the categories in which they are listed. The focus of the paper is on the relationship between the category architecture and the type of information which may be credibly disclosed by firms to searching consumers. In our setting both firms and consumers actively try to overcome informational asymmetries: firms by choosing a specific category in which to list their product, and consumers by choosing a category in which to search. The firms' category choice can be viewed as a form of active information disclosure. However, in contrast to conventional information disclosure, in our setting the information which actually reaches consumers depends on the latter's active participation. That is, the information which is disclosed by firms' choice of category is only revealed to consumers who actually choose to search in that category.

We consider a setting in which each firm produces one type of product and products are differentiated both horizontally and vertically. There are two types of products, A and B, and each product is available in low or high quality. Consumers differ with respect to their preferences between the product types but they all prefer high quality over low quality. Consumers know the distribution of firms' characteristics but must incur search costs to find specific product attributes. While in the standard search setup consumers sample the entire population of firms, here we assume that there are different categories of products in which consumers may search. The categories may be in terms of firms' horizontal characteristics only (A or B), vertical characteristics only (high or low qualities), or both types of characteristics. Firms do not control the categorization structure and are unable to create new categories but can only decide in which categories to be listed. The availability of exogenous product categories may enable firms to direct consumer search and promote more efficient matches between products and consumers.

We begin our analysis by analyzing the firms' choice of categories with respect to product characteristics under different category architectures when prices are exogenous.[4] In section 3, we consider the case in which only horizontal categories, A, B and

[4] We incorporate pricing into the model in Section 6.

AB, are available, where category AB provides no explicit information about product type. Given this category structure firms decide the category in which to list their products while consumers choose a category in which to search. In this case, there exists a product-type revealing equilibrium where horizontal characteristics are perfectly revealed (i.e., all A products list in category A and all B products list in category B). Under certain conditions, there also exists a quality-revealing equilibrium in which the low-quality firms list according to their horizontal characteristics, while high-quality firms list in the anonymous AB category. In this equilibrium, a firm implicitly discloses that it is of high quality by *not* disclosing its product type. Thus in the former equilibrium consumers have perfect product type information, while in the latter they have perfect quality information.

In section 4, we consider vertical (quality) categories. If only vertical categories are available, we show that there are no equilibria in which firms fully reveal their quality if vertical categories are not *verifiable* - that is, if firms are able to list under any category they wish, even if their products do not match the category description. Thus, in our setting, firms are able to reveal their quality when the available categories describe only horizontal characteristics but not when the available categories explicitly refer to quality.[5]

In section 5, we endogenize the product quality distribution by opening the market to free entry. This allows us to examine how the category structure feeds back on and determines the equilibrium distribution of product qualities. Focusing on horizontal categorization we show that when search costs are relatively small both the quality revealing equilibrium and the product-type revealing equilibrium exist but the quality revealing equilibrium induces a higher fraction of high-quality firms and higher consumer surplus.

In section 6, we allow firms to choose prices as well as categories. We extend the framework presented in Section 2 and demonstrate that our main results regarding categorization and information revelation continue to hold in the model with pricing.

[5] We also show that when there are both vertical and horizontal categories (but the latter ones are not verifiable), firms cannot reveal more information than when there are only horizontal categories are available.

There is a vast literature on consumer search. For example, Diamond (1971), Burdett and Judd (1983), and Stahl (1989) study consumer search models with homogenous products where consumers search for low prices. Wolinsky (1986), and Anderson and Renault (1999) study consumer search models with horizontally differentiated products where consumers search for both low prices and products matching their taste. Our model is more closely related to the latter branch of the literature, but features both horizontal and vertical product differentiation. More importantly, in our setting, the introduction of product categories changes the way in which consumers search. In standard search models, firms are usually *ex ante* identical and so consumers sample firms in a random order. But in our setting, some product information can be revealed through firms' category choices, enabling consumers to search from among more relevant products. In this sense, our paper is also related to Weitzman (1979) which studies the optimal stopping rule when options are *ex ante* asymmetric, and the more recent papers on prominence (caused by online paid placement, for instance) and non-random search (see, for example, Athey and Ellison, 2011, Armstrong, Vickers and Zhou, 2009, Armstrong and Zhou, 2011, and Chen and He, 2011).

Category choice in our model plays a role similar to advertising product information. This relates our paper to the literature on search and advertising. For example, Robert and Stahl (1993) and Janssen and Non (2008) study price advertising in a search model where consumers can gather price information through a combination of advertising and their own search. Anderson and Renault (2006) considers advertising and search in a monopoly setting but in their model the firm can advertise either price information or match utility information or both. It is also interesting to note that in a different setting, Mayzlin and Shin (2011) derive an equilibrium similar to our quality revealing equilibrium. Their model assumes a monopolistic market structure and that products have two vertical attributes. The firm is able to disclose only one attribute at most but consumers can learn about both attributes through costly search. They show that a signalling equilibrium can exist where the high quality firm signals that it is of high quality in the second dimension by not disclosing its type in the first dimension.

2 A Model of Search with Categories

Consider a market with a continuum of firms whose measure is normalized to 1. Firms' products are differentiated both horizontally and vertically. There are two product types, A and B (e.g., A is Japanese food and B is Chinese food). Half of the firms produce product A and the other half produce product B. In each group of firms, a fraction α produce a high-quality product (denoted H), and a fraction $1-\alpha$ produce a low-quality product (denoted L). A firm's type is denoted as $t_f \in T_f = \{AH, AL, BH, BL\}$, where, for instance, AH indicates product A of high quality. We assume that firms have constant marginal cost, which is assumed to be zero. In the basic model we keep the number of firms of each quality type fixed. In section 5 we consider free entry of firms and endogenize the fraction of firms of each quality type.

There is a continuum of consumers of measure m. Consumers have heterogenous preferences with respect to the product type (A or B) and with respect to the product quality (H or L). Specifically, A and B are located at the two ends 0 and 1 of a Hotelling line of length one. Consumers are distributed uniformly along this line, and a consumer's location is denoted by $x \in [0,1]$. Let γ be the Hotelling unit "transportation cost". All consumers prefer high quality to low quality but differ in their valuations for quality, which is indexed by q. We assume that q is also uniformly distributed on $[0,1]$. Thus, a consumer's type is denoted by $t_c = (x,q) \in T_c \equiv [0,1]^2$. The valuation of a type (x,q) consumer for the low-quality A product and the high-quality A product are respectively,

$$U_{AL}(x,q) = v - \gamma x \quad \text{and} \quad U_{AH}(x,q) = v + q - \gamma x \ .$$

Similarly, her valuations for the low-quality B product and the high-quality B product are respectively,

$$U_{BL}(x,q) = v - \gamma(1-x) \quad \text{and} \quad U_{BH}(x,q) = v + q - \gamma(1-x) \ .$$

We assume that the basic valuation v is large enough that the market is fully covered.

In our setting products differ along two dimensions: product type and quality. The former represents horizontal variation as some consumers prefer product A while others prefer product B. The latter represents vertical variation as all consumers prefer H

over L.[6]

We assume that *ex ante* consumers know neither the product type nor the quality of any firm, but can learn both through a sequential search process. Whenever a consumer investigates a firm, she learns its type (both its product type and quality). Following convention, we assume that it is costless to investigate the first firm but after that it costs s to investigate each additional firm. We further assume that search is not too costly such that $s < \min\{\frac{1}{2}, \frac{\gamma}{2}\}$. After each search, a consumer learns the firm's type and then decides whether to buy the product or to continue to search. We do not consider prices explicitly in the basic model. Thus, a consumer's surplus from buying product $i \in \{A, B\}$ of quality $j \in \{H, L\}$ after searching n times is $U_{ij}(x, q) - (n-1)s$. In section 6 we extend the model to include price competition.

In conventional search models, there are no product categories (or, equivalently, there is only one category), and consumers search by sampling firms randomly as firms are *ex ante* identical. Here we depart from this and implicitly suppose that there is an information intermediary (e.g., a search web site) that provides product information in categories. The set of all possible categories is

$$C^A \equiv \{A, B, AB, AH, BH, AL, BL, H, L, HL\}\ .$$

A *category structure* $C \subseteq C^A$ specifies the available categories. For example if there are categories only with respect to the horizontal dimension then $C = \{A, B, AB\}$. When more than one category exists, each firm needs to choose in which category to list their products[7] and consumers decide in which category to search. Once a consumer chooses a category, she inspect firms sequentially within this category but may switch categories if she wishes to do so, where firms within the category are sampled in a random order. Unless otherwise stated, we will assume that a firm can only list itself under one category. (See section 7 for a discussion of the possibility that a firm can list under multiple categories.) Note, however, that since we assume a continuum of firms,

[6] More broadly speaking, the vertical dimension does not have to be quality. It can be two different colors, say, red and blue, provided that all people prefer one color over the other.

[7] We assume that firms determine the category in which they are listed. Even if the firm's listing is chosen directly by the intermediary, the firm can indirectly affect where it is listed by the way that it describes its service or product.

a consumer will never search more than one category even if she can - if it was initially optimal to search in a specific category, it remains optimal to search that category after having sampled a finite number of firms in that category.[8]

A potentially important distinction is between *verifiable categories* and *non-verifiable categories*. When categories are verifiable firms cannot join a category that is different from the type of product they sell. This might be because the information intermediary can verify the type of product that firms produce and can make sure that a firm's product actually matches the category in which it is listed.[9] For example, if categories are verifiable and $C = \{A, B, AB\}$, then firms that produce product A may only be listed under categories A or AB and firms which produce product B may only be listed under categories B or AB. Similarly, if $C = \{H, L\}$ and categories are verifiable, all high quality firms must choose H while low quality firms must choose L. By contrast, if categories are not verifiable, then a firm can list itself in any category.

Formally, we define "*a search problem with categories*" as a search problem with a given set of categories C such that (i) the strategy of a firm is its choice of category $S_f : T_f \to C$ in which it is listed; (ii) all consumers have the same beliefs about the distribution of product types in each category, denoted as $B(C)$; (iii) given these beliefs consumers choose the category in which to search and their acceptance set in that category, i.e., the set of product types that they are willing to accept in that category without further search.[10] That is, consumers' strategy set is $S_c : T_c \times B(C) \to C \times AC$ where AC is consumer's acceptance set (which is a subset of the product types in the chosen category). Let s_{t_f} denote the strategy of a firm of type $t_f \in T_f$, and let $s_f = \{s_{t_f}\}_{t_f \in T_f}$ be a profile of all firms' strategies. Similarly, let s_{t_c} denote the strategy of a consumer of type $t_c \in T_c$, and let $s_c = \{s_{t_c}\}_{t_c \in T_c}$ be a profile of all consumers' strategies.

---

[8] In a model with a finite number of firms the search strategy would need to specify a sequence of categories which consumers search through.

[9] Or consumers' behavior is such that whenever they observe such a contradiction they do not buy from such firms.

[10] Formally consumers' strategy is a history dependent search rule that specifies for every history the category of the next search and the acceptance set of the next search. But in order to simplify our discussion we use standard equilibrium condition to simplify the definition of consumers' strategy set.

Denote by $\pi_{t_f}(s_{t_f} \mid C, s_c, s_f)$ the expected profit of a firm of type t_f when its category choice strategy is s_{t_f} given the category structure C and the strategies of all the other firms and consumers (s_f and s_c respectively). Denote by $u_{t_c}(s_{t_c} \mid C, B(C))$ the expected surplus of a consumer of type t_c when she chooses strategy s_{t_c} given her beliefs about the distribution of product types in each category.

Definition 1 *For a given category structure C a search with categories equilibrium is a triple $\{s_c^*, s_f^*, B^*(C)\}$ such that:*

- For each consumer of type $t_c \in T_c$, $s_{t_c}^*$ maximizes the expected surplus $u_{t_c}(s_{t_c} \mid C, B^*(C))$ given their belief $B^*(C)$.
- For each firm of type $t_f \in T_f$, $s_{t_f}^*$ maximizes the expected profit $\pi_{t_f}(s_{t_f} \mid C, s_c^*, s_f^*)$ given consumer search strategies s_c^* and other firms' strategies s_f^*.
- The consumer belief $B^*(C)$ is consistent with the firm strategies s_f^*.

Note that the above definition applies whether or not firms can list in more than one category and whether or not categories are verifiable. If firms are able to list in more than one category, a firms' strategy is a choice of a subset of C. When categories are verifiable, verifiability constrains firms to list only in categories which match their type.

For any category structure C there are sets of strategies s_f under which some categories are empty.

Definition 2 *An "empty category" is a category in which the measure of firms is zero.*

The specification of consumer search behavior when there are empty categories is an important ingredient of the consumers' search strategy. When considering a putative equilibrium with one or more empty categories, it is necessary to specify consumers' beliefs and behavior when a firm deviates and chooses to list under an empty category (which is "supposed" to be empty). We will adopt the following simple behavioral rule:[11]

[11] An alternative approach would be to follow the signalling literature and specify as part of the equilibrium construction consumers' off-equilibrium beliefs if they observe a deviation to an empty category.

Assumption 1 *Consumers do not search empty categories.*

Under this assumption, a firm which lists in an empty category has no customers. Therefore, if in some equilibrium one or more categories are empty, it can never be profitable for a firm to deviate by listing itself under such a category. Note that since we assume continuum of firms a deviation of a firm to an empty category is not going to change the fact that there is still a measure zero of firms in this category.[12]

3 Horizontal Categorization

We first consider the case in which there are only horizontal categories, that is $C = \{A, B, AB\}$. In this case there are several possible equilibria. The first possibility is that categorization provides no information about product type or quality. This occurs if all firms list under the same category, say, AB. This case is essentially equivalent to a setting without any categories. A second possibility is that categories provide complete information about product type. That is, all type A firms list under one subset of categories while all type B firms list under a disjoint subset of categories. In this case consumers can perfectly infer a firm's product type from the category under which it is listed, but no quality information is revealed at all. The third possibility is that low-quality firms and high-quality firms choose to list under different categories so that even if the horizontal categories do not *explicitly* provide quality information, consumers are able to infer it from firms' category choices. For expositional convenience only, in the remainder of this section we assume that horizontal categories are verifiable, though this assumption is not crucial for our analysis. When horizontal categories are not verifiable, corresponding to the equilibria derived below there also exist identical equilibria in which the names of the categories are permuted.

[12] Note that in our equilibrium construction, we will actually only need to deal with the situation where an "empty category" has no firms at all or only one firm which deviated from a non-empty category. So the zero-measure part in the definition of empty categories can be weakened.

3.1 Horizontal categories reveal no information

Under horizontal categorization, there is always an uninformative equilibrium where all firms choose to list under the same category and so categories do not reveal any information at all. This equilibrium is equivalent to the case where there is no categorization at all.

Proposition 1 ***(Pooling equilibrium)*** *If $C = \{A, B, AB\}$, there is always an equilibrium in which all firms list in category AB and consumers search only in category AB. The characterization of consumer search behavior in category AB is depicted in Figure A1.*

Proof. See the Appendix. ■

It is easy to see that firms have no incentive to deviate from the equilibrium category, given that consumers do not search empty categories. The characterization of consumers search within this category is, however, not trivial[13] and is depicted in Figure A1 in the Appendix.

3.2 Horizontal categories reveal product type information

The second type of equilibrium with horizontal categorization is such that firms' choice of categories actually reveal their product type information.

Proposition 2 ***(Product-type revealing equilibrium)*** *If $C = \{A, B, AB\}$, there is an equilibrium in which all A firms list in category A and all B firms list in category*

[13] Note that even for a consumer with $x < \frac{1}{2}$ the ranking of the four possible products depends of his type (x, q):

$$\begin{array}{ll} BL \prec AL \preceq BH \prec AH & \text{if } q \geq \gamma(1-2x) \\ BL \prec BH \prec AL \prec AH & \text{if } q < \gamma(1-2x) \end{array}.$$

A consumer with a very large q and a very small x will search until she find the perfect match AH while on the other hand a consumer with a low q and with an x close to 0.5 will accept the first product that she samples without any further search. Similarly a consumer with a high q but x close to 0.5 (a consumer who cares about quality but not too much about product type A or B) will search for a high quality regardless of the type of product while consumers with a low q but with x close to zero or to 1 will search for the "right" product type regardless of its quality.

B, independent of quality, and consumers search either in category A or B and follow the search strategy described in Figure 1 below.

Proof. Given that consumers do not search in empty categories (Assumption 1), firms have no incentive to deviate and list in category AB.[14] Given firms' listing strategies, consumers with $x < \frac{1}{2}$ search category A while consumers with $x \geq \frac{1}{2}$ search category B. In each category, a consumer searches until she finds a high-quality product if and only if $q \geq \min\{1, \frac{s}{\alpha}\}$. To see that, consider, for instance, category A. If a consumer accepts the first sampled product, her expected surplus is $v - \gamma x + \alpha q$. (Recall that the first search is costless.) If she searches until she finds a high-quality product (which needs $\frac{1}{\alpha}$ searches on average), her expected surplus is $v - \gamma x + q - (\frac{1}{\alpha} - 1)s$. The latter is greater if and only if $q > \frac{s}{\alpha}$ (and notice that q cannot exceed 1 in our model). This optimal search strategy is illustrated in Figure 2 below. ∎

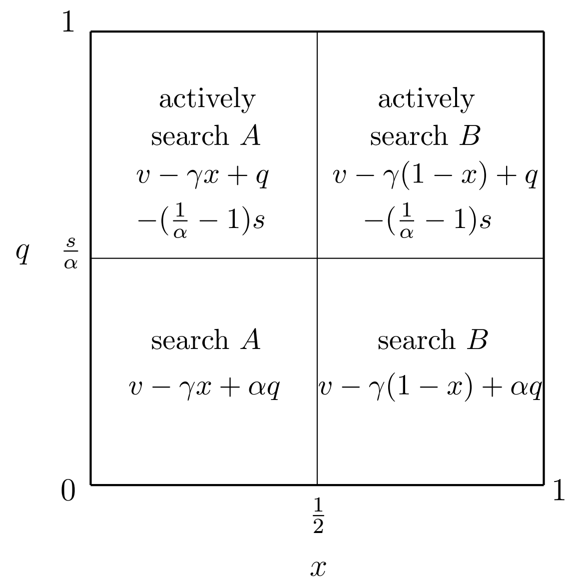

Figure 1: Pattern of demand when horizontal categories reveal product type information

[14] If consumers cannot observe the deviation (e.g., category headings do not indicate the number of firms in that category), it is natural that consumers do not search an empty category (as part of the equilibrium construction). If consumers can observe the deviation, the presumption that listing under an empty category leads to zero profit can be justified if consumers hold the (out of equilibrium) belief that a firm that lists in category AB is low quality. Given that belief, it is a dominant strategy for consumers never to search category AB.

Consumers are better off in this equilibrium than in the pooling equilibrium since they can at least get their preferred product type without search. However, to get a high-quality product they, on average, have to search.

3.3 Horizontal categories reveal quality information

We now turn to the less obvious (and perhaps also more interesting) equilibrium where horizontal categories provide information about the vertical attribute. In this equilibrium high-quality firms of either type list in category AB while low-quality firms of type i list in category i. Then the firms' category choices fully reveal product quality but only partially reveal their product type. This case is illustrated in Figure 2 below

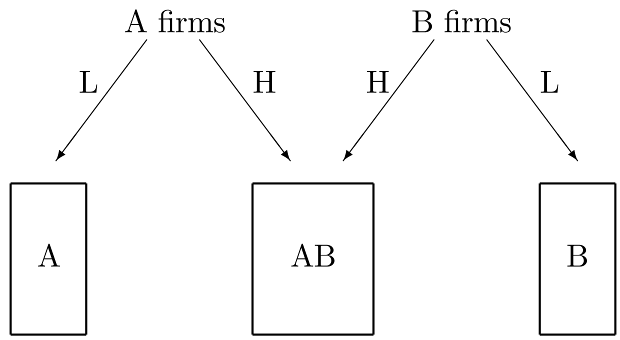

Figure 2: Horizontal categorization reveals quality information

Proposition 3 *(**Quality revealing equilibrium**) When*

$$\frac{2+s/\gamma}{2(1-s/\gamma)}\frac{1-s+s^2/\gamma}{1-s+2s/\gamma} \leq \frac{\alpha}{1-\alpha} \leq \frac{1-s+s^2/\gamma}{s-s^2/\gamma}\ , \tag{1}$$

there is an equilibrium in which all high-quality firms list in category AB, all low-quality i firms list in category i, $i = A, B$, and consumers follow the search strategy described in Figure 3.

Proof. See the Appendix. ■

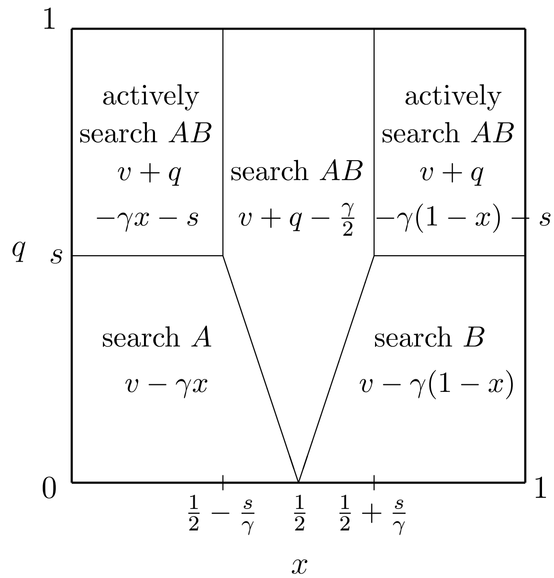

Figure 3: Pattern of demand when horizontal categories reveal quality information

Consumer search behavior in this quality revealing equilibrium is described in Figure 3. Consumers who do not care too much about quality (i.e., with $q < s$) but care about product type search category A or B, depending on their locations, even if they expect to get a low-quality product for sure. These consumers distribute on the regions of "search A" and "search B" in Figure 3. For a consumer in the region of "search A", her expected surplus, when her horizontal location is x, is $v - \gamma x$ as indicated in the figure, since she will buy the first product she samples and the first search is costless. Similarly, $v - \gamma(1-x)$ is the expected surplus of a consumer in the region of "search B" when her location is $1-x$. Consumers who care about both quality and product type search category AB (because all products in that category are of high quality) and, within that category, search until they find the right product type. They distribute on the two regions of "actively search AB". These consumers search twice on average and their expected search cost is thus s. Their surplus is indicated in each region. Finally, consumers who do not care too much about product type search category AB for high-quality products, but accept the first product they encounter, whether it is type A or B. They distribute on the central region of "search AB".

The interesting feature of the quality revealing equilibrium is that although the categories provide no explicit information about quality, in equilibrium consumers are able to perfectly infer quality from firms' category choices: the firms in category A or B supply low-quality products while those in category AB supply high-quality products. In particular, the apparently uninformative category AB now endogenously conveys quality information, and the partially informative categories A and B are now fully informative. However, since category AB is not informative about product type, consumers who are sensitive to product quality have to "pay" for high-quality, either by not buying their preferred product type (the region of "search AB") or by engaging in costly search for the right type (the regions of "actively search AB").

However, the quality revealing equilibrium exists only when condition (1) holds. For example, when $\gamma = 1$, it simplifies to $\frac{(2+s)(1-s+s^2)}{4-s-s^2+s^3} \leq \alpha \leq 1 - s + s^2$. The set of (s, α) which satisfy this condition is the region between the two solid curves in Figure 4 below. Intuitively, for a given s, the fraction of high-quality firms α cannot be too high or too low. If α is too high, then the market for high-quality products is too crowded and each high-quality firm faces only low demand. Therefore those firms would switch to category $i \in \{A, B\}$ and compete with relatively few low-quality firms. By contrast, if α is too low, the market for low-quality products is too crowded and low-quality firm would want to list in category AB and compete with high-quality firms.[15]

One may wonder if it is also possible that low-quality firms, either of type A or type B, list in category AB while firms that produce the high-quality i product list in category i, $i = A, B$. It is not difficult to see that this cannot be an equilibrium. Suppose, in contrast, there were such an equilibrium. Then categories A and B would identify both product types and all firms listed there would be high-quality, while category AB would only list low-quality firms and would reveal no information about product type. Therefore, no consumers would ever want to search in category AB, and thus the firms in that category could profitably deviate.[16]

[15]It would have some demand because those consumers who do not value quality highly would buy from it when it is encountered first.

[16]For example, when a low-quality A firm deviates and lists in category A, those consumers who search in that category and have a sufficiently low valuation for quality would buy this low-quality product without further search when it is encountered first. This ensures a positive deviation profit.

3.4 Horizontal categories and consumer welfare

We now turn to the effects of horizontal categorization on consumer welfare. The uninformative pooling equilibrium is clearly the worst for consumers. But different consumers may have different preferences between the product-type revealing equilibrium and the quality revealing equilibrium. From Figures 1 and 3 we can compare each type of consumer's surplus under the two equilibria. For example, consumers with relatively low valuation of quality but who are choosy about product type (i.e., those located on the southeast and southwest corners), prefer the product-type revealing equilibrium. In both equilibria, they get the right product type but in the product-type revealing equilibrium they also get high quality with probability α. While the consumers for whom both quality and product type are important (i.e., those located on the northeast and northwest corners), prefer the quality revealing equilibrium if s (the search cost needed to find the right product type in the quality revealing equilibrium) is less than $(\frac{1}{\alpha} - 1)s$ (the search cost needed to find a high-quality product in the product-type revealing equilibrium), i.e., if the fraction of high-quality firms $\alpha < \frac{1}{2}$.

The following result compares total consumer surplus between the two equilibria.

Proposition 4 *The quality revealing equilibrium gives rise to higher total consumer surplus than the product-type revealing equilibrium if and only if*

$$\frac{2}{3\gamma}s^2 - \left(\frac{1}{2} + \frac{1}{\gamma}\right) s + 1 + \left(\frac{s}{2\alpha} - 1\right)\left(\frac{1}{\alpha} - 1\right) < 0 \ . \tag{2}$$

Proof. See the Appendix. ■

For example, when $\gamma = 1$, the region between the two dashed curves in Figure 4 below describes the set of (s, α) which satisfy condition (2). Recall that the quality revealing equilibrium exists only when (s, α) is between the two solid curves. Therefore, only in the region in the middle does the quality revealing equilibrium exists and also generates higher consumer surplus. Beyond this region, either the quality revealing equilibrium does not exist or it is dominated by the product-type revealing equilibrium in terms of consumer welfare.

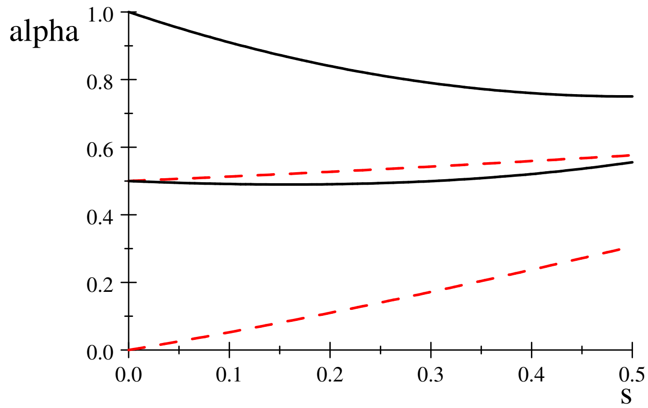

Figure 4: Quality revealing equilibrium and its welfare impacts

4 Vertical Categorization

Vertical categories are very common in many websites. Examples of such categories include the five-star rating system for hotels, rating of airlines etc. Moreover, many online information intermediaries rate sellers according to customers' quality reviews.

We start by considering the case in which there are only vertical categories: $C = \{H, L, HL\}$. As before, there is always a trivial equilibrium in which all firms list in the same category, say HL, and therefore categorization provides no information. This equilibrium can again be sustained by the assumption that consumers do not search empty categories. The interesting question, however, is whether there is a separating equilibrium where high-quality firms list in category H and low-quality firms list in category L. When quality categories are verifiable, such a separating equilibrium simply exists because each firm *must* list itself according to its actual quality. (Given that consumers do not search empty categories, an L firm cannot gain by deviating and listing itself in HL.) The outcome is different if vertical categories are not verifiable as the following proposition shows.

Proposition 5 *Consider the category structure $C = \{H, L, HL\}$. If vertical categories are not verifiable, there is no equilibrium in which high-quality firms list in category H*

and low-quality firms list in category L.

Proof. Suppose instead that all H firms list in category H and all L firms list in category L and category HL is empty. Then, since each list contains the same proportion of A and B firms, every consumer gets higher utility by searching category H and therefore L firms will have no demand and make zero profit. Suppose an L firm of type i deviates and lists in category H. Then consumers who have relatively low valuations for quality and prefer product type i will buy it if it samples it first. Thus the deviation is profitable for this L firm. ■

Thus in our model, a category structure with unverifiable vertical categories cannot fully disclose vertical information, but a category structure with only horizontal categories, even if they are unverifiable, can lead to an equilibrium in which vertical information is fully disclosed.

Now let us consider a more "complete" category structure $C = \{AH, AL, BH, BL\}$. If both horizontal and vertical categories are verifiable, there again exists an equilibrium in which each type of firms lists in the right category and both product-type and quality information are revealed. But if only horizontal categories but not vertical categories are verifiable, then for a similar reason as above, there is no equilibrium in which quality information is revealed.

Proposition 6 *Consider category structure $C = \{AH, AL, BH, BL\}$. If vertical categories are not verifiable quality information cannot be revealed in equilibrium.*

If neither horizontal nor vertical categories are verifiable, the following type of equilibrium can be sustained under certain conditions: all AL firms list in category AL, all BL firms list in category BL, all high-quality firms (independent of their product type) list in category AH, and category BH remains empty. Consumers who do not care about quality too much will search in either AL or BL to find the right product type. Consumers who care enough about quality will search in AH even though they may end up buying the wrong product type. No firms want to deviate and list in empty category BH because consumers do not search empty categories. In fact, this equilibrium is effectively identical to the quality-revealing equilibrium when the category structure is $C = \{A, B, AB\}$.

5 Free Entry and Endogenous Product Quality

Thus far we have assumed that the distribution of product quality is exogenously given. We now extend our model allowing for free entry focusing on the effect of categorization on the percentage of high-quality firms in the industry. Firm can enter the industry by paying an entry cost, where the entry cost for a high-quality firm is F_H and the entry cost of a low-quality firm is $F_L < F_H$. We focus on the horizontal category structure $C = \{A, B, AB\}$ since considering (unverifiable) vertical categories cannot lead to more informative new equilibria. We assume that firms enter the industry as long as their profits cover their entry costs. Since the firms' profits depend on the categorization equilibrium, the percentage of firms of each type depends on the category structure and the type of equilibrium which obtains. Our main result is presented in the following proposition. Its proof and the properties of the resultant market structure are derived with the aid of three lemmas that are presented after the proposition.

Proposition 7 *In a free-entry market with a sufficiently small search cost s, both the product-type revealing equilibrium and the quality revealing equilibrium exist, and the quality revealing equilibrium induces a higher fraction of high-quality firms and leads to greater consumer surplus.*

In order to prove this proposition we first investigate the conditions under which the quality revealing equilibrium in Proposition 3 exists in a free-entry setting. It turns out that when search costs are sufficiently small this equilibrium always exists in a free-entry environment.

Let n be the total measure of firms in the free-entry equilibrium, and α be the fraction of high-quality firms as before.

Lemma 1 *In a free-entry market, if the condition*

$$\frac{s(2 + s/\gamma)}{2(1 - s + 2s/\gamma)} \leq \frac{F_L}{F_H} \tag{3}$$

holds, there exists a quality revealing equilibrium where the fraction of high-quality firms is

$$\alpha = \left(1 + \frac{F_H}{F_L}\frac{s - s^2/\gamma}{1 - s + s^2/\gamma}\right)^{-1} . \tag{4}$$

In particular this equilibrium always exists for sufficiently small search costs.

Proof. See the Appendix. ■

To illustrate condition (3), let us consider the example with $\gamma = 1$.[17] The left-hand side of (3) increases from 0 to $\frac{5}{12}$, and so the quality revealing equilibrium exists if $\frac{F_L}{F_H} \geq \frac{5}{12}$.

We then turn to the product-type revealing equilibrium.

Lemma 2 *In a free-entry market, if $s \leq \frac{F_L}{F_H}$, there exists a product-type revealing equilibrium where the fraction of high-quality firms is*

$$\alpha = \left(1 + \frac{1}{s} - \frac{F_H}{F_L}\right)^{-1} \in (s, 1) \ . \tag{5}$$

Proof. See the Appendix. ■

The above two lemmas indicate that both the quality revealing equilibrium and the product-type revealing equilibrium exist if both $s < \frac{F_L}{F_H}$ and (3) are satisfied, or equivalently if

$$\max\left\{s, \frac{s(2 + s/\gamma)}{2(1 - s + 2s/\gamma)}\right\} \leq \frac{F_L}{F_H} \ . \tag{6}$$

In the following Lemma we compare the fraction of high-quality firms and consumer welfare between these two equilibria:

Lemma 3 *(i) If the condition*

$$\frac{s}{1 - s + s^2/\gamma} < \frac{F_L}{F_H} \tag{7}$$

holds, both equilibria exist and the quality-revealing equilibrium induces a higher proportion of high-quality firms than the product-type revealing equilibrium.
(ii) If the conditions (6) and

$$\frac{2}{3\gamma}s^2 - \left(\frac{1}{2} + \frac{1}{\gamma}\right)s + 1 + \frac{s}{2}\left(1 - \frac{1}{s} - \frac{F_H}{F_L}\right)\left(\frac{1}{s} - \frac{F_H}{F_L}\right) < 0 \tag{8}$$

hold, both equilibria exist and the quality-revealing equilibrium gives rise to higher consumer surplus than the product-type revealing equilibrium.

[17] So $s < \frac{1}{2}$ from our assumption that $s < \min\{\frac{1}{2}, \frac{\gamma}{2}\}$.

Proof. See the Appendix. ■

To prove Proposition 7 it suffices to note that when the search cost s is close to zero, both conditions (7) and (8) are satisfied. Therefore there is a higher proportion of high-quality firms in the quality revealing equilibrium and consumers are also better off under this equilibrium.

6 Search Categories and Pricing

So far we have only considered differences between product characteristics without formally considering price competition. In this section we extend our analysis to include price competition. Our aim is to show that the equilibria with information disclosure derived in the base model can still exist when prices are endogenous.

There are two ways to incorporate pricing into our setting of search with categories. One is to assume that there are "pricing categories". The second, more conventional approach is to allow for endogenous prices within each category. We believe that both approaches are realistic and applicable to different market settings.

The first approach is a direct extension of our previous analysis in which the vertical categories represent two different price levels. Since all consumers prefer low prices over high prices, L now stands for high prices and H now stands for low prices. Note that indeed many search intermediaries use discrete price categories such as $ and $$ symbols for restaurants etc. We thus can view prices as a special case of vertical differentiation such that the H and L categories represent two levels of prices.[18]

The second approach, which is the focus of this section, assumes that firms may choose any price they wish. In this case firm's strategy, s_{t_f}, is a choice of a category and a price. We denote by $P(C)$ the price distribution in each category in the category structure C.[19] We assume that all consumers have the same beliefs about the distribution of product types and prices in each category: $\{B(C), P(C)\}$. Consumer's strategy, s_{t_c}, is a choice of category in which to search and an acceptance set which is a set of product types and prices that they are willing to accept. We can thus modify

[18] One can extend our setup to include several price levels.

[19] Note that $P(C)$ is a vector of distribution of prices.

our previous definition of search with categories equilibrium in the following way.

Definition 3 *For a given category structure C a search with categories equilibrium is a quadruple $\{s_c^*, s_f^*, B^*(C), P^*(C)\}$ such that:*

- For each type of consumers, their search strategy $s_{t_c}^*$ is optimal given their beliefs $\{B^*(C), P^*(C)\}$.
- For each type of firms, their strategy $s_{t_f}^*$ maximizes their profits given consumer strategy s_c^* and other firms' strategies s_f^*.
- The consumer belief $\{B^*(C), P^*(C)\}$ is consistent with the firms' strategies s_f^*.

To simplify, we modify our basic setup. We assume that there are only two types of consumers in terms of their horizontal preferences who are located at the two ends of the Hotelling line. Half the consumers prefers product A: if they consume a low-quality A product they get utility v, and if they consume a low-quality B product they get utility $v - \gamma$. The other half of consumers prefers product B. We further assume that γ is sufficiently large such that consumers have strong horizontal preferences (e.g., a consumer who prefers product A will never want to buy a B product). We also assume that $v > s$ such that products are sufficiently valuable. The additional utility from higher quality remains the same as before, and we focus on the category structure $C = \{A, B, AB\}$. As before, for convenience we assume that horizontal categories are verifiable.

6.1 Product-type revealing equilibrium

We first investigate the consider the product-type revealing equilibrium where all $i \in \{A, B\}$ firms list in category $i \in \{A, B\}$, independent of their quality types. We seek to characterize an equilibrium where all the low-quality firms charge $p_L = v$ while all the high-quality firms charge $p_H \geq v$. Consumers that prefer product $i \in \{A, B\}$ search only category $i \in \{A, B\}$. For all consumers, the low-quality product at the price $p_L = v$ provides a zero surplus. But whether the high-quality product provides a positive surplus or not depends on a consumer's type q.

Given the consumers' (correct) beliefs about the distribution of prices and qualities in each category their optimal search behavior is as follows:

(i) For a consumer with $q < p_H - v$, the high-quality product provides a negative surplus and so she searches only once (given the first search is free). If the product she samples first is of low-quality she buys it immediately. If the first sampled product is of high quality, she leaves the market without purchasing anything.

(ii) For a consumer with $p_H - v \leq q \leq p_H - v + \frac{s}{\alpha}$, the high-quality product provides a positive surplus, but the surplus is too small to be worth searching for. So this type of consumer searches only once and buys whatever good she samples at the first search.

(iii) For a consumer with $q > p_H - v + \frac{s}{\alpha}$, the high-quality product is sufficiently attractive that she searches until she finds it.[20]

Regarding the firms' category choices, it is clear that firms cannot benefit by deviating and listing themselves in AB since consumers do not search empty categories.[21] So it remains to ensure that firms have no incentive to change their prices given the above consumer search behavior. The following proposition provides the conditions.

Proposition 8 *Suppose $C = \{A, B, AB\}$, and the condition*

$$(1 + \alpha + \frac{1}{\alpha})v - 1 \leq 2s \leq \alpha(1 + v) \tag{9}$$

holds. Then there exists a product-type revealing equilibrium where low-quality firms charge a price $p_L = v$ and high-quality firms charge a price $p_H = \frac{1}{1+\alpha}\left(1 + s + v - \frac{s}{\alpha}\right) \geq v$.

Proof. See the Appendix. ■

[20] Note that in order for a search equilibrium with all three types of consumers to exist, we need that $p_H - v + \frac{s}{\alpha} < 1$.

[21] Given that categories are verifiable, an A firm, say, cannot list itself in B category. But even if categories are unverifiable, firms have no incentive to deviate because consumers choose to search categories according to their own preferences and buy only the type of product that they like (under our assumption that γ is sufficiently large).

6.2 Quality revealing equilibrium

We now consider the quality revealing equilibrium where all the low-quality firms list in category A or B according to their type while all the high-quality firms list in category AB. We focus on an equilibrium in which all the low-quality firms charge the price p_L and all the high-quality firms charge p_H. Since all low-quality firms are identical, the Diamond paradox result implies that in equilibrium $p_L = v$. In this equilibrium, consumers' optimal search rule is characterized by two cutoffs, q_1 and $q_2 > q_1$:

(i) A consumer with $q < q_1$ searches in category A or B (depending on what type of product she likes) and buys the low-quality product at any price no greater than v.

(ii) A consumer with $q_1 \leq q \leq q_2$ searches in category AB and buys if and only if the product she finds at her first search is the type of product she prefers. Otherwise she leaves the market without buying.

(iii) A consumer with $q > q_2$ searches in category AB until she finds the right product.

Given $p_L = v$, the surplus from searching in the low-quality product category is zero. Then we must have

$$q_1 = p_H - v \tag{10}$$

such that a consumer of type q_1 is indifferent between searching in the low-quality product category and searching in the high-quality one. A consumer who chooses to search in category AB will not stop searching until she finds the right product if $\frac{1}{2}(v + q - p_H) > s$ or equivalently $q > 2s + p_H - v$. Hence, we have

$$q_2 = 2s + p_H - v = 2s + q_1 \ . \tag{11}$$

To sustain the proposed equilibrium, we need to find q_1 or p_H such that firms have no incentive to change their category choices or prices. The following proposition claims that this requirement is satisfied if p_H satisfies the following three conditions:[22]

$$p_H \leq 2v \ , \tag{12}$$

[22] One can check that there is a range of parameters that satisfies these conditions. For example, if $s \to 0$ and $v = 1$, then all the conditions are satisfied if

$$\frac{4}{3} \leq p_H \leq 2$$

$$\max\{v, \frac{2}{3}(1+v-s)\} \leq p_H \leq 1+v-2s \ , \tag{13}$$

$$\frac{1}{2}(v+\frac{s}{2})^2 \leq \frac{\alpha}{1-\alpha}v(p_H - v) \leq p_H(1-p_H+v-s) \ . \tag{14}$$

Proposition 9 *Suppose $C = \{A, B, AB\}$, and there is a price p_H that satisfies conditions (12)-(14). Then there exists a quality revealing equilibrium in which the low-quality firms of type i choose category $i \in \{A, B\}$ and charge a price $p_L = v$, and all high-quality firms choose category AB and charge a price p_H.*

Proof. See the Appendix. ■

In order for it to be optimal for high-quality firms to list in AB, the number of consumers which search there must be sufficiently large, which requires that p_H not be too large. It must also be optimal for high-quality firms in AB to charge p_H rather than deviate to some other price. The conditions in Proposition 9 guarantee that the high-quality firms as well as the low-quality firms cannot benefit from a unilateral deviation from p_H and p_L and from their choice of categories. Notice that Proposition 9 identifies a range of p_H that can be part of equilibrium behavior, so the quality revealing equilibrium is not unique.

Discussion: vertical categories. We saw in section 3 that when all products are priced the same, the quality revealing equilibrium can exist only under the horizontal category architecture but not under the vertical category architecture (if vertical categories are not verifiable). This result is modified when we introduce endogenous pricing. Specifically, suppose the available categories are H and L and prices are endogenous. Then for appropriate parameter values we can construct an analogous equilibrium to the one presented in Proposition 9 in which low-quality firms of both types list in L and charge the price v,[23] and high-quality firms of both types list in H and charge some $p_H > v$. Consumers with low q values search in L and buy at their first search if they

---

and

$$\frac{1}{2} \leq \frac{\alpha}{1-\alpha}(p_H - 1) \leq p_H(2-p_H) \ .$$

This is non-empty. For example, $p_H = \frac{4}{3}$ and $\alpha \in [\frac{3}{5}, \frac{8}{11}]$ satisfy all conditions.

[23] We assume that consumers have sufficiently strong type preferences that it is not profitable for low quality firms to reduce the price in order to sell to both types of consumers.

find the right product type and otherwise leave the market without buying. Consumers with higher q search in the H category, where consumers with intermediate values of q buy only if they find the right type of product at their first search and consumers with sufficiently high values of q search for their preferred product type. The details of the construction are similar to the above and are omitted.

7 Discussion and Concluding Comments

We have considered a very specific setup of search with categories. Clearly there are different aspects of search markets with categories that are important and deserve a more careful analysis. In our concluding section we discuss some of these issues and their potential effect on market analysis.

Multiple listings. In our basic setup each firm chooses one category. When the model is interpreted in terms of advertising or positioning by firms, it is natural to assume that a firm can list under one category only. But when categorization is implemented by an intermediary such as a search engine, it may be necessary to extend the model to allow for the possibility that firms may list under multiple categories. One possibility is that firms choose a subset of the provided categories in which to list where the cost of listing is determined by specific market arrangements. Formally, our model can handle such situations by modifying the firms' strategy choice from a single category to a subset of categories. Another possibility is that multiple listings are automatically implemented by the search engine, so that when a firm chooses to list in a narrow category, it is automatically also listed in a more general category. For example, if a firm lists under category A its also automatically appears in category AB. We provide a brief analysis of the latter case in the following paragraph in order to demonstrate the possible effects of multiple listing.

Consider the horizontal category structure $C = \{A, B, AB\}$ and suppose that if a type-i firm chooses to list in category $i \in \{A, B\}$, it will automatically appear also in the more general category AB. Consider a possible product-type revealing equilibrium first. Suppose all A firms choose to list in category A and all B firms choose to list in category B. Then all firms will also appear in category AB. Given that consumers will only search in either category A or category B, no firms want to list in category AB only. Therefore, a similar product-type revealing equilibrium always exists even if we allow multiple listings.[24]

Now consider a possible quality revealing equilibrium. Suppose AL type firms choose to list in category A and BL firms choose to list in category B, and all high-quality firms choose to list in category AB. Then the automatic multi-listing implies that all firms will actually appear in category AB. Compared to the case with single listing, quality information is not totally revealed and the expected quality of category AB becomes lower. But this cannot be sustained as an equilibrium because a high-quality firm can always do better by listing in category i. In fact, given the automatic enrollment into category AB, listing in category A or B weakly dominates listing in category AB. This result implies that whenever intermediaries wish to design the rules of category affiliation they need to take into account the fact that multi-listing may destroy the possibility of using categories to signal the firms' quality.

Endogenous consumer participation. Our analysis has assumed full-market coverage, i.e., all consumers buy the product. But search with categories may have interesting implications regarding the number of consumers participating in the market. When there is no categorization (or if the uninformative pooling equilibrium prevails), some consumers may opt out of the market because they anticipate an inefficient search process. For example, assume that most of the firms are type A and there are very few type B firms and consider a consumer with a very strong preference for type B product. Searching for type B may be very costly and therefore the consumer is better off not entering the market. Horizontal categorization may solve the problem by making it easier for the consumer to find the product he likes and therefore induce him to participate in the market.

Platform design. Generally the category structure is chosen by the information platform. So in addition to the fee structure that is often discussed in the literature, how to design the category structure is an important decision for the platform. It will affect both consumers' willingness to use the information service and firms' listing

[24] It is also easy to see that it is still an equilibrium that all firms list in category AB given that consumers do not search empty categories.

strategies and their willingness to list in the platform.

Prominence and Categorization. The standard search setup assumes that all the objects (the firms in our case) are randomly sampled, each with the same probability. The search and prominence literature assumes that objects may be sampled with different probabilities, such that a more prominent object is sampled with a higher probability. The objects' prominence can be either exogenously given or endogenously determined by the firms' activities. An interesting extension would be to introduce prominence into our model of search with categories. Specifically, firms may have different prominence in different categories (when prominence is determined exogenously). Thus the firms' category choice may depend also on its prominence in the different categories and not just on the categories' signalling value. When prominence is endogenously determined, say by advertising, it may be that the cost of achieving prominence is different for different categories which again may affect the firms' category choice. Such an extension is beyond the focus of this paper and may hold promise for interesting future research.

8 Appendix

Consumer search behavior in the pooling equilibrium in Proposition 1. Consider a consumer at $x < \frac{1}{2}$. (The case for $x > \frac{1}{2}$ is symmetric.) She values and ranks the four possible products as follows:

$$\begin{array}{cc} BL \prec AL \preceq BH \prec AH & \\ \Leftrightarrow v - \gamma(1-x) < v - \gamma x \leq v + q - \gamma(1-x) < v + q - \gamma x & \text{if } q \geq \gamma(1-2x) \\ & . \\ BL \prec BH \prec AL \prec AH & \\ \Leftrightarrow v - \gamma(1-x) < v + q - \gamma(1-x) < v - \gamma x < v + q - \gamma x & \text{if } q < \gamma(1-2x) \end{array}$$

In particular, if this consumer has a relatively high valuation for quality, she prefers BH to AL though the former is not her ideal product type. By contrast, if she has a relatively low valuation for quality, she prefers AL to BH.

Suppose $q \geq \gamma(1-2x)$. If the consumer buys the first product she samples, her expected surplus is $v - \frac{\gamma}{2} + \alpha q$. If she buys products no worse than AL, she needs to search $\frac{2}{1+\alpha}$ times on average and so her expected surplus will be

$$\frac{1-\alpha}{1+\alpha}(v-\gamma x)+\frac{2\alpha}{1+\alpha}\left(v+q-\frac{\gamma}{2}\right)-\left(\frac{2}{1+\alpha}-1\right)s\ .$$

If she buys high-quality products only (i.e., product BH or AH), she needs to search $\frac{1}{\alpha}$ times and so her expected surplus will be

$$v+q-\frac{\gamma}{2}-\left(\frac{1}{\alpha}-1\right)s\ .$$

Finally, if she buys the ideal product AH only, she needs to search $\frac{2}{\alpha}$ times and her expected surplus will be

$$v+q-\gamma x-\left(\frac{2}{\alpha}-1\right)s\ .$$

Comparing these four options reveals the optimal search strategy when $q \geq \gamma(1-2x)$. The case of $q < \gamma(1-2x)$ can be dealt with similarly. The optimal consumer search behavior is described in Figure A1 below. There, for example, "AH" indicates that consumers on that region stop searching only if they find a product no worse than AH, and "BH/AL" indicates that the threshold product for consumers on that region is the worse one between BH and AL (depending on $q \geq \gamma(1-2x)$ or not). ■

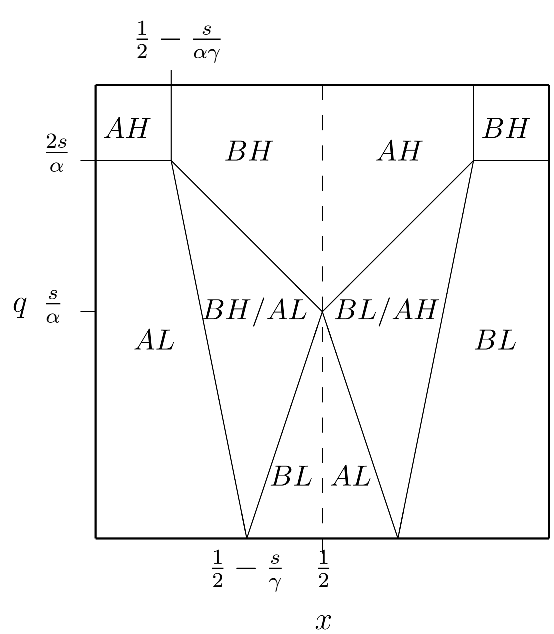

Figure A1: Pattern of demand when categorization reveals no information

Proof of Proposition 3. The proof consists of three steps.

Step 1: consumer search behavior in equilibrium. Suppose that indeed in equilibrium high-quality firms list in AB, low-quality i firms list in category i and consumers believe that firms list in this manner. Consider a consumer at $x < \frac{1}{2}$ (the case with $x > \frac{1}{2}$ is symmetric). She has three relevant search options. The first is to search category A and get a low-quality A product; the second is to search category AB and buy the first product she encounters; and the third option is to search category AB until she finds an AH product. The optimal search behavior can be derived by comparing these three options. (Notice that searching category B is dominated by searching category A for a consumer at $x < \frac{1}{2}$.)

If a consumer searches category A then, given her belief about the distribution of qualities she will buy the first product she samples. Thus her expected surplus will be $v - \gamma x$. Suppose the consumer searches category AB. If she does not actively search in AB and buys the first product she samples, her expected surplus will be

$$\frac{1}{2}(v + q - \gamma x) + \frac{1}{2}(v + q - \gamma(1 - x)) = v + q - \frac{\gamma}{2} \ .$$

If she searches sequentially until finding an A product then she needs to sample two products on average. Since the first sampling is costless, the (expected) search cost is only s. Consequently her expected surplus will be

$$v + q - \gamma x - s \ .$$

By comparing these three options, one can readily check that given our assumption $s < \min\{\frac{1}{2}, \frac{\gamma}{2}\}$ the optimal consumer search behavior is described as in Figure 3.

Step 2: A high-quality firm has no incentive to deviate and list in category A or B. Without loss of generality, consider an AH firm. Given consumer search behavior, in the proposed equilibrium, an AH firm's demand is

$$Q_H \equiv \left(1 - s + \frac{s^2}{\gamma}\right) \frac{m}{\alpha} \ , \tag{15}$$

where $(1 - s + \frac{s^2}{\gamma})m$ is the measure of consumers who chooses category AB, and α is the measure of high-quality firms. Notice that due to symmetry, an AH firm has the same demand as a BH firm, and thus each high-quality firm's demand is simply the number of consumers that search in the AB category divided by the number of firms listing in this category.

Suppose then an AH firm deviates and chooses to list in category A. Then the consumers who search this category and encounter it will buy its product without further search. So this AH firm's demand will be identical to any AL firm's demand in category A. To calculate this demand notice that in the proposed equilibrium, $\frac{s}{2}(1-\frac{s}{\gamma})m$ consumers choose to search in category A, and there are $\frac{1-\alpha}{2}$ AL firms in this category. Therefore,

$$Q_L \equiv \frac{\frac{s}{2}(1-\frac{s}{\gamma})}{\frac{1-\alpha}{2}}m = \left(s - \frac{s^2}{\gamma}\right)\frac{m}{1-\alpha}\ . \tag{16}$$

Consequently, a high-quality firm has no incentive to deviate and list in category A or B if

$$Q_L \leq Q_H\ . \tag{17}$$

Step 3: A low-quality firm has no incentive to deviate and list in category AB. Consider an AL firm. Its equilibrium demand is Q_L in (16). Suppose now that this firm deviates and lists in category AB. To calculate the deviation demand we need to figure out how a consumer who chooses to search category AB will behave if she encounters this deviation firm. We only need to consider those consumers on the left region of "actively search AB" and the region of "search AB" in Figure 3. (Those consumers on the right region of "actively search AB" will never buy from this AL firm since they do not even buy from an AH firm.)

If a consumer buys from this AL firm, her surplus is $v-\gamma x$. If she searches once more and buys at the next firm (which must supply a high-quality product), her expected surplus is $v+q-\frac{\gamma}{2}-s$. If she searches until finding an AH product, her expected surplus is $v+q-\gamma x-2s$. (Notice that the consumer needs to search twice on average in order to find an AH product.) The consumer's optimal behavior can be derived by comparing these three options. Given the assumption of $s<\frac{1}{2}$, the consumer will buy from this deviation firm if she locates on $[0,\frac{1}{2}-\frac{s}{\gamma}]\times[s,2s]$ or on the region of "search AB" below the line $q=\gamma(\frac{1}{2}-x)+s$. One can verify that the area of this whole region is $\frac{s}{2}+\frac{s^2}{4\gamma}$.

Notice that for an AH product, those consumers on the left region of "actively search AB" or on the region of "search AB" will buy it immediately once they sample it. The area of the whole region is $\frac{1-s}{2} + \frac{s}{\gamma}$. But the purchasing area for an AL product in the deviation case is a subset of it. Thus, the deviation firm's demand is a proportion of the equilibrium demand for an AH firm:

$$\frac{\frac{s}{2} + \frac{s^2}{4\gamma}}{\frac{1-s}{2} + \frac{s}{\gamma}} Q_H = \frac{\gamma s + s^2/2}{\gamma(1-s) + 2s} Q_H \ .$$

Therefore, an AL firm has no incentive to deviate if

$$\frac{\gamma s + \frac{s^2}{2}}{\gamma(1-s) + 2s} Q_H \leq Q_L \ . \tag{18}$$

Combining (17) and (18), we can see that the proposed equilibrium can be sustained if and only if

$$\frac{\gamma s + \frac{s^2}{2}}{\gamma(1-s) + 2s} Q_H \leq Q_L \leq Q_H \ ,$$

which is equal to (1) by using (15) and (16). ■

Proof of Proposition 4. From Figure 1, one can derive total consumer surplus in the product-type revealing equilibrium:

$$v + \frac{1}{2} - \frac{\gamma}{4} - \left(\frac{1}{\alpha} - 1\right)\left(s - \frac{s^2}{2\alpha}\right) \ .$$

From Figure 3, one can derive total consumer surplus in the quality revealing equilibrium:

$$v + \frac{1}{2} - \frac{\gamma}{4} - s + \left(\frac{1}{2} + \frac{1}{\gamma}\right) s^2 - \frac{2}{3\gamma} s^3 \ .$$

The latter is larger than the former if (2) holds. ■

Proof of Lemma 1. If a quality revealing equilibrium exists, from (15) and (16) we know that the profit of a high-quality firm and the profit of a low-quality firm (without considering the entry cost) are respectively

$$\pi_H(n, \alpha) = \left(1 - s + \frac{s^2}{\gamma}\right) \frac{m}{n\alpha} \ ; \ \pi_L(n, \alpha) = \left(s - \frac{s^2}{\gamma}\right) \frac{m}{n(1-\alpha)} \ .$$

Then the free-entry conditions are

$$\left(1 - s + \frac{s^2}{\gamma}\right) \frac{m}{n\alpha} = F_H \ ; \ \left(s - \frac{s^2}{\gamma}\right) \frac{m}{n(1-\alpha)} = F_L \ .$$

They determine n and α. In particular, one can solve

$$\alpha = \frac{1}{1+\frac{F_H}{F_L}\frac{s-s^2/\gamma}{1-s+s^2/\gamma}} \Leftrightarrow \frac{\alpha}{1-\alpha} = \frac{F_L}{F_H}\frac{1-s+s^2/\gamma}{s-s^2/\gamma} .$$

Recall that the condition for the quality revealing equilibrium is (1):

$$\frac{2+s/\gamma}{2(1-s/\gamma)}\frac{1-s+s^2/\gamma}{1-s+2s/\gamma} \leq \frac{\alpha}{1-\alpha} \leq \frac{1-s+s^2/\gamma}{s-s^2/\gamma} .$$

The second half of this condition holds given $F_L < F_H$. One can check that the first half of the condition also holds if and only if (3) is satisfied. ■

Proof of Lemma 2. If the product-type revealing equilibrium exists, a high-quality firm's profit is

$$\pi_H(n,\alpha) = \left(1-\min\{1,\frac{s}{\alpha}\}\right)\frac{m}{n\alpha} + \min\{1,\frac{s}{\alpha}\}\frac{m}{n} .$$

For those consumers with $q > \frac{s}{\alpha}$, a high-quality firm is competing only with other high-quality firms. But for those with $q < \frac{s}{\alpha}$, it is competing with all firms. (Note that we need to take into account the possibility that $\frac{s}{\alpha} > 1$.) For a low-quality firm, only those consumers with $q < \frac{s}{\alpha}$ may patronize it and it is competing with all other firms. Hence, a low-quality firm's profit is

$$\pi_L(n,\alpha) = \min\{1,\frac{s}{\alpha}\}\frac{m}{n} .$$

The free-entry conditions are then:

$$\left(1-\min\{1,\frac{s}{\alpha}\}\right)\frac{m}{n\alpha} + \min\{1,\frac{s}{\alpha}\}\frac{m}{n} \leq F_H \ ; \ \min\{1,\frac{s}{\alpha}\}\frac{m}{n} \leq F_L .$$

(We allow weak inequalities because corner solutions may exist in this case.) Then one can show that the equilibrium described in the lemma exist when $s < \frac{F_L}{F_H}$.[25] ■

Proof of Lemma 3. (i) It is ready to derive (7) by comparing (4) and (5). One can also check that under the assumption of $s < \min\{\frac{1}{2},\frac{\gamma}{2}\}$, (7) implies (6) and so both equilibria exist.

[25] There are also two product-type revealing equilibria with corner solutions: (i) There always exists a free-entry equilibrium with $\alpha = 0$ (i.e., only low-quality firms enter the market) in which each firm earns $\frac{m}{n} = F_L$. (ii) When $s < \frac{F_L}{F_H}$, there exists a free-entry equilibrium with $\alpha = 1$ (i.e., only high-quality firms enter the market) in which each firm earns $\frac{m}{n} = F_H$.

(ii) From the proof of Proposition 4, we can see that consumer welfare in the quality revealing equilibrium does not depend on α. Therefore, the condition for the quality revealing equilibrium to generate higher consumer welfare is the same as before:

$$\frac{2}{3\gamma}s^2 - \left(\frac{1}{2} + \frac{1}{\gamma}\right)s + 1 + \left(\frac{s}{2\alpha} - 1\right)\left(\frac{1}{\alpha} - 1\right) < 0 \ ,$$

except that α is now given in (5). Substituting (5) into this inequality yields (8). ■

Proof of Proposition 8. In the proposed equilibrium, the low-quality firms charge the price $p_L = v$. Clearly these firms cannot raise their price. On the other hand lowering their price may increase the number of units that they sell. Specifically, given the consumers' search strategy, a low-quality firm sells only to consumers who sample it at their first search and who do not search for a high-quality product. (No consumers will search for a low-quality product beyond the first visited firm given the (expected) price $p_L = v$.) Therefore, in equilibrium the demand for each low-quality product is $m\left(p_H - v + \frac{s}{\alpha}\right)$.[26] Suppose a low-quality firm slightly reduces its price to $v - \varepsilon$. The (first-order) loss of doing so is $m\left(p_H - v + \frac{s}{\alpha}\right)\varepsilon$ (i.e., those who buy from this firm pay ε less). The benefit is that consumers with q slightly higher that $p_H - v + \frac{s}{\alpha}$ that continue to search if they sample a low-quality product at the price v will purchase the low-quality good if its price is $v - \varepsilon$. More precisely, a consumer will buy the low-quality product at price $v - \varepsilon$ instead of continuing to search for a high-quality product if $\varepsilon \geq \alpha(v + q - p_H) - s$, i.e., if her type is $q \leq p_H - v + \frac{s+\varepsilon}{\alpha}$. Therefore, the (first-order) benefit of reducing the price by ε is $m\frac{\varepsilon}{\alpha}v$. A low-quality firm has no incentive to deviate from $p_L = v$ if the loss exceeds the benefit, i.e., if

$$p_H - v + \frac{s}{\alpha} \geq \frac{v}{\alpha} \Leftrightarrow p_H \geq v + \frac{v - s}{\alpha} \ . \tag{19}$$

Now let us consider high-quality firms. In equilibrium, a high-quality firm sells to consumers with an intermediate q who buy whatever product they sample at their first search and to high-q consumers who search for a high-quality product. Thus the demand they face is $m\left[\frac{s}{\alpha} + \frac{1}{\alpha}(1 - (p_H - v + \frac{s}{\alpha}))\right]$.

[26]Note that the measure of consumers is m and the measure of firms is 1. So each firm has m first-time visitors.

Suppose a high-quality firm unilaterally reduces its price by a small ε. Its (first-order) loss is the lower price (by ε) paid by existing customers. The benefit is that it acquires additional new customers with relatively low q — more precisely, those consumers with $q > p_H - v - \varepsilon$ who sample this firm first — yielding the (first-order) benefit $m\varepsilon p_H$. In an equilibrium with an interior solution of p_H, the loss should be equal to the benefit, which determines p_H as

$$p_H = \frac{1 + s + v - s/\alpha}{1 + \alpha} . \tag{20}$$

To sustain the proposed equilibrium, we need to verify the conditions $p_H - v + \frac{s}{\alpha} \le 1$, $p_H \ge v$ and (19). They are equivalent to

$$\max\left\{v, v + \frac{v - s}{\alpha}\right\} \le p_H \le 1 + v - \frac{s}{\alpha} .$$

Given the assumption $s < v$, one can check that the equilibrium price in (20) satisfies these constraints if (9) holds. ■

Proof of Proposition 9. Given the consumer search behavior described in the main text, we need to ensure that firms have no incentive to change their category choices or prices. We first consider prices and then category choices.

The low-quality firms charge the monopoly price $p_L = v$ and cannot benefit from changing it according to the standard Diamond paradox argument. Regarding the high-quality firms, it is unprofitable for them to reduce p_H below $v + q_1$. But what about if a firm unilaterally raises its price to $p_H + \varepsilon$? In equilibrium the demand for each high-quality firm is

$$\frac{1}{\alpha}\left[\frac{m}{2}(q_2 - q_1) + m(1 - q_2)\right] .$$

This is because half of the consumers with $q \in [q_1, q_2]$ eventually buy from a high-quality firm, and all consumers with $q > q_2$ buy from a high-quality firm. And the measure of all high-quality firm is α. So the (first-order) benefit of raising the price slightly is ε times this equilibrium demand. The (first-order) loss caused by this small price increase is derived from those consumers who sample this firm first, have $q \in [q_1, q_1 + \varepsilon]$ and who like this firm's product type but will refrain from buying due to the higher price. So the lost demand is $\frac{m\varepsilon}{2\alpha}$, which leads to a loss of $\frac{m\varepsilon}{2\alpha} p_H$. Thus, in equilibrium

p_H should satisfy

$$\frac{m\varepsilon}{2\alpha}p_H \geq \frac{\varepsilon}{\alpha}\left[\frac{m}{2}(q_2 - q_1) + m(1 - q_2)\right] \ .$$

By using $q_1 = p_H - v$ and $q_2 = q_1 + 2s$ in (10) and (11), this condition simplifies to

$$p_H \geq \frac{2}{3}(1 + v - s) \ .$$

To ensure that $q_1 \geq 0$ and $q_2 \leq 1$, we need $p_H \geq v$ and $p_H \leq 1 + v - 2s$. Therefore, to sustain the pricing equilibrium, we need

$$\max\{v, \frac{2}{3}(1 + v - s)\} \leq p_H \leq 1 + v - 2s \ . \tag{21}$$

Now consider the firms' category choice. In the proposed equilibrium, a low-quality firm's profit is

$$\pi_L = p_L \frac{mq_1}{1-\alpha} = \frac{m}{1-\alpha}v(p_H - v) \ .$$

A high-quality firm's profit is

$$\pi_H = p_H \frac{1}{\alpha}\left[\frac{m}{2}(q_2 - q_1) + m(1 - q_2)\right] = \frac{m}{\alpha}p_H(1 - p_H + v - s) \ .$$

Suppose that a high-quality A firm deviates and lists in category A. Given all AL firms are charging $p_L = v$, it can act as a monopoly. If it charges a price $p \in [v, v + q_1]$, then its deviation profit is $\frac{mp}{1-\alpha}[q_1 - (p - v)]$. So the optimal deviation price is

$$\hat{p}_L = \max\{v, \frac{v + q_1}{2}\} \ .$$

In particular, if $p_H \leq 2v$ then the optimal deviation is $\hat{p}_L = v$. In this case a high-quality firm will not deviate if

$$\pi_H \geq \frac{m}{1-\alpha}v(p_H - v) = \pi_L \ .$$

Suppose now that a low-quality A firm deviates and lists in category AB and charges a price $p \leq v$. If a consumer with $q \geq q_1$ encounters this firm, what will she do? Given the assumption of a sufficiently high γ, only those consumers who like product A may buy. For those consumers with $q \in [q_1, q_2]$, they will buy this low-quality product with $p \leq v$ if they like product A. This yields demand $\frac{m(q_2 - q_1)}{2\alpha} = \frac{ms}{\alpha}$. For those consumers with $q \geq q_2$, they will buy this low-quality product if

$$v + q - p_H - 2s \leq v - p \quad \Leftrightarrow \quad q \leq q_2 + v - p \ .$$

The number of consumers with $q \geq q_2$ who come to visit this AL firm is $\frac{m(1-q_2)}{\alpha}(1 + \frac{1}{2} + (\frac{1}{2})^2 + \cdots) = \frac{2m(1-q_2)}{\alpha}$. So the demand from this source is

$$\frac{2m(1-q_2)}{\alpha}\frac{v-p}{1-q_2} = \frac{2m(v-p)}{\alpha} \ .$$

So this AL firm's deviation profit, when $p \leq v$, is

$$p\left[\frac{ms}{\alpha} + \frac{2m(v-p)}{\alpha}\right] \ .$$

So the optimal deviation price is $\min\{v, \frac{v}{2} + \frac{s}{4}\}$. Given our assumption that $v \geq s$ the optimal price is $\frac{v}{2} + \frac{s}{4}$, and the optimal deviation profit is $\frac{m}{2\alpha}(v + \frac{s}{2})^2$. Therefore, a low-quality firm will not deviate if

$$\pi_L \geq \frac{m}{2\alpha}(v + \frac{s}{2})^2 \ .$$

In sum, under the assumption of $p_H \leq 2v$, firms have no incentive to change their category choices if

$$\pi_H \geq \pi_L \geq \frac{m}{2\alpha}(v + \frac{s}{2})^2$$

or more explicitly if

$$p_H(1 - p_H + v - s) \geq \frac{\alpha}{1-\alpha}v(p_H - v) \geq \frac{1}{2}(v + \frac{s}{2})^2 \ . \tag{22}$$

Therefore, the proposed quality revealing equilibrium exists if $p_H \leq 2v$, and both conditions (21) and (22) hold. ■

References

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

——— (2006): "Advertising Content," *American Economic Review*, 96(1), 93–113.

Armstrong, M., J. Vickers, and J. Zhou (2009): "Prominence and Consumer Search," *Rand Journal of Economics*, 40(2), 209–233.

Armstrong, M., and J. Zhou (2011): "Paying for Prominence," *Economic Journal*, 121(556), F368–F395.

Athey, S., and G. Ellison (2011): "Position Auctions with Consumer Search," *Quarterly Journal of Economics*, 126(3), 1213–1270.

Bar-Isaac, H., G. Caruana, and V. Cunat (2012): "Search, Design, and Market Structure," *American Economic Review*, 102(2), 1140–60.

Brown, J., and A. Goolsbee (2002): "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry," *Journal of Political Economy*, 110(3), 481–507.

Burdett, K., and K. Judd (1983): "Equilibrium Price Dispersion," *Econometrica*, 51(4), 955–969.

Chen, Y., and C. He (2011): "Paid-Placement: Advertising and Search on the Internet," *Economic Journal*, 121(556), F309–F328.

de Corniere, A. (2013): "Search Advertising," mimeo, Oxford.

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

Goldmanis, M., A. Hortascu, C. Syverson, and O. Emre (2010): "E-Commerce and the Market Structure of Retail Industries," *Economic Journal*, 120(545), 651–682.

Janssen, M., and M. Non (2008): "Advertising and Consumer Search in a Duopoly Model," *International Journal of Industrial Organization*, 26(1), 354–371.

Mayzlin, D., and J. Shin (2011): "Uninformative Advertising as an Invitation to Search," *Marketing Science*, 30(4), 666–685.

Robert, J., and D. Stahl (1993): "Informative Price Advertising in a Sequential Search Model," *Econometrica*, 61(3), 657–686.

Stahl, D. (1989): "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review*, 79(4), 700–712.

Weitzman, M. (1979): "Optimal Search for the Best Alternative," *Econometrica*, 47(3), 641–654.

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