Targeted Search and the Long Tail Effect

Huanxing Yang[*][†]
Department of Economics, Ohio State University

June 2011

Abstract

This paper develops a new search model to explain the long tail effect. Search targetibility or the quality of search is explicitly modeled. Consumers are searching for the right products within the right categories. Mainstream consumers are distinguished from long tail consumers in terms of the prevalence of consumer tastes in the population. We show that mainstream consumers enjoy higher utility and mainstream products are sold at lower prices. In the market equilibrium long tail consumers might be excluded. As search costs decrease or search targetibility increases, additional variety of goods catering to long tail consumers will be provided and the concentration of sales across different category of goods decreases. The effects of a decrease in search costs or an increase in search targetibility on consumer utility, prices, and profits depend on whether the type coverage increases. Decreases in search costs and increases in search targetibility have different qualitative effects.

Keywords: Search, targetibility, product variety, long tail
JEL: D83, L11, L86

1 Introduction

The widespread usage of the Internet has dramatically changed the variety and the distribution of products offered. On the one hand, the variety of goods available has been steadily increasing, with more and more niche products being offered. On the other hand, the distribution of sales has become flatter, with niche products gaining larger market shares. Anderson (2004, 2006, 2009) referred to this phenomenon as the "long tail." Specifically, in the book industry, from 2002 to 2007 the number of new titles grew almost 10% a year. Actually, the number of new titles in 2007 alone was more than those published throughout the 1970s.[1] Similar patterns are found in markets for music and DVDs. Rhapsody, an online music provider, has more downloads of the songs beyond its top 10,000 than those within its top 10,000. For video

[*]Arps 448, 1945 N. High St., Columbus, 43210. Email: yang.1041@osu.edu
[†]I would like to thank Heski Bar-Issac, Howard Marvel, and Jim Peck for helpful discussions.
[1]According to Frank Urbanowski, Director of MIT Press, the increased accessibility to backlist titles through the Internet lead to a 12% increase in the sale of these titles.

rental shops, "new release" movies usually account for a dominant share of revenue. However, for DVDStation, a company that allows consumers to search and reserve movies online, more than 50% of their rental revenue comes from titles that are not new releases.[2]

One explanation for the *long tail* effect is that the Internet decreases inventory costs. Due to space constraints, a brick-mortar store can only carry a limited variety of goods. These logistical constraints are absent for online stores, so they can carry a much larger variety of goods. With more variety of goods available online, consumers can have access to the products of their preferred tastes and sales will spread more to niche products. However, this supply-side story of product availability does not tell the whole story. Several recent studies, Brynjolfsson et al. (2003) on online bookstores, Brynjolfsson et al. (2007) on the clothing retailing industry, and Elberse and Oberholzer-Gee (2008) on the video industry, found that even after controlling for product availability, online sales still exhibit the long tail effect relative to offline sales.

This paper presents a new model of search to provide an explanation for the long tail effect. In particular, we explicitly model *search targetibility* or the *quality* of search, which enables us to distinguish decreases in search costs from increases in search targetibility, both caused by the widespread use of online search. We not only study how online search affects the variety of goods offered and the concentration of sales, but also study the effects on consumer utility, price dispersion, and the distribution of firms.

Specifically, consumers are of different types with distinctive tastes. A consumer of a particular type only demands a good of a corresponding type, which defines product categories. There is an exogenously given population of firms, and each firm can only choose to serve one type of consumers, or produce one category of goods. Within the right category, each consumer likes different firms' products to different degrees, or a consumer's valuation about a particular firm's product is a random draw from some distribution.[3] The distribution of consumer types is exogenously given. We call consumer types which have relatively large fractions of population *mainstream types*, and those having relatively small fractions of population *long tail types*. The timing is a follows. First, firms simultaneously choose product categories (which type of consumers to serve). Then, observing the type distribution of firms, firms simultaneously set their prices, and consumers conduct search and buy goods.

Consumers search sequentially. Before searching, each consumer has a targeted set of firms, which consists of all the firms of the right category (signal) and a certain fraction of firms of irrelevant categories (noise). One can think that this targeted set is generated by some online

[2] The facts in this paragraph can be found in Greco (2005), Brynjolfsson et.al (2006), and Bar-Issac et.al (2011).

[3] For a concrete exmaple, consider books. At the category level, some consumers only want to read detective stories (DS) and some consumers only want science fiction (SF). This distinction defines two consumer types: the type of DS and the type of SF. At the book (firm) level, a particular consumer of type, say, DS, likes different DS books to different degrees.

search engines. Within the targeted set, consumers search randomly. With this formulation, consumers are not only searching for the right category of goods, but also searching for the right products within the right category. The optimal amount of search depends on the probability of finding the right category. If the targeted set contains fewer firms of irrelevant categories (the targetibility of search increases or the noise decreases), for each type of consumer the probability of finding the right category increases. On the other hand, if there are more firms serving a particular consumer type, then the probability of finding the right category for that type is higher, as the signal to noise ratio in the targeted set is higher.

Given the set of consumer types covered, there is at most one equilibrium. In any equilibrium that covers more than two types, mainstream consumers enjoy higher utility and search more within the right category than long tail consumers do. Moreover, there are more firms serving mainstream types, and firms serving mainstream types charge lower prices and have higher sales per firm than those serving long tail types.[4] Intuitively, given that there are more mainstream type consumers, more firms will naturally serve mainstream types as they are potentially more profitable. Now the probability of finding the right category is higher for mainstream consumers, and they will search more within the right category, which intensifies competition among firms of the same category and leads to lower prices. As a result, mainstream consumers enjoy higher utilities. Given that firms serving mainstream consumers charge lower prices, the sales per firm for those firms are higher than those serving long tail consumers in order to restore the equal profit condition.[5]

Due to the coordination feature of exclusion, there are multiple equilibria with different sets of consumer types covered. To resolve the issue of multiple equilibria, we introduce a notion of stability by considering firms' joint deviation in choosing the types of consumers to serve. We show that, given parameter values, there is a unique stable equilibrium, which we call market equilibrium. The market equilibrium must be of monotonic configuration: if a consumer type is covered, then all the consumer types more mainstream than that type must be covered. Actually, the market equilibrium has the most types covered among all equilibria with monotonic configurations. In the market equilibrium some long tail types are potentially excluded. This is because to induce a particular type to search the probability of finding the right category for that type must be high enough. If the population of that type is too small, hence can only accommodate too few firms, then the probability of finding the right category for that type will be too small. In that case, consumers of that type will not bother searching and are excluded in the market equilibrium.

[4] This implies that for covered types, compared to the distribution of consumer types, the distribution of firms is skewed more toward long tail types.

[5] To prevent firms from deviating to serving another type, all the firms must earn the same profit in equilibrium.

When either the search costs decrease or the search targetibility increases, (weakly) more long tail types of consumers will be covered in the market equilibrium, leading to (weakly) more variety of goods offered and lower concentration of sales across types. The underlying reason is that both changes increase consumers' incentive to search. This provides an explanation for the long tail effect. When the consumer coverage does not increase, both changes will lead to lower prices, lower profits, and higher utility for each covered consumer type. This is due to the fact that increased search intensity within the right category intensifies competition among firms. When the consumer coverage does increase, the effects of both changes on profits and consumer utilities are ambiguous. This is because increased type coverage will decrease each consumer type's probability of finding the right category, as firms are spreading over more types. In some sense, more type coverage softens competition among firms. This effect tends to increase prices and profits and decrease utilities for consumers. As a result, the overall effect is ambiguous.

Decreases in search costs and increases in search targetibility have different qualitative effects. First, while a decrease in search costs always induces consumers to conduct more overall search, an increase in search targetibility might lead to less overall search. Second, an increase in search targetibility tends to reduce the difference between mainstream consumers and long tail consumers, as the probabilities of finding the right category become more equalized among consumers. It is not clear whether a decrease in search costs always has a similar effect. When consumers' match value is uniformly distributed, we show that decreases in search costs and increases in search targetibility have distinctive (sometimes opposite) effects on consumers' overall amount of search, the distribution of prices, and the distribution of firms across types.

Finally, in an extension we incorporate free entry of firms. While most of the results in the basic model hold qualitatively under free entry, some results depend on whether the measure of irrelevant firms in the targeted set (noise) increases with the measure of active firms in the market. In general, with free entry the effects of changes in search costs and search targetibility are dampened, as the total measure of active firms will endogenously adjust, which tends to partially offset the direct effects of the initial changes.

There is an extensive literature on consumers searching for prices among firms offering homogenous goods, e.g. the non-sequential search model of Varian (1980) and the sequential search model of Stahl (1989). This paper is more related to the literature on searching for variety of goods. Wolinsky (1986) is the first model that studies consumers searching for right products among heterogenous goods, followed by Bakos (1997) and Anderson and Renault (1999). In particular, Anderson and Renault show that the monopoly pricing result of the Diamond (1971) model and the marginal cost pricing result of the Bertrand competition are the two limiting cases of Wolinsky's model. In those models, consumers are ex ante identical and firms are symmetric; hence there is no issue of search targetibility. In our model, consumers are of different types and different firms might choose to serve different consumer types. This allows us to model search targetibility and address the long tail effect.

Bar-Isaac et al. (2010) provide a search model with endogenous product design to explain the long tail effect and the super star phenomenon.[6] In their model, firms are vertically differentiated or of different qualities. Firms choose prices and product design, which ranges from broad market designs that appeal to all consumers to some average extent to more niche designs that are very appealing to some consumers but very unattractive to other consumers. In equilibrium, higher quality firms choose the most broad design and lower quality firms choose the most niche design. As consumers' search costs decrease, more firms choose niche designs. In their model, consumers are ex ante homogenous and the increase in variety of goods offered is embodied in more firms choosing niche designs, while in our model consumers are ex ante heterogenous and the increase in variety of goods offered is reflected in more long tail types of consumers covered (or more categories of goods offered).

Hervas-Drane (2010) considers how online recommendation systems affect sales distribution in a search model. He shows that the presence of a general recommendation system tends to increase sales concentration, while a personalized recommendation systems tends to reduce it. In his model, search is either completely random (no recommendation) or not needed (with recommendation).[7][8]

The rest of the paper is organized as follows. Section 2 sets up the basic model. Section 3 analyzes consumers' search behavior and firms' pricing behavior. In Section 4 we characterize equilibria and establish the existence and uniqueness of stable equilibrium. Section 5 studies comparative statics about the market equilibrium and shows that decreases in search costs and increases in search targetibility lead to different effects. Section 6 incorporates free entry and Section 7 concludes.

2 Basic Model

There is a continuum of consumers with total measure m, and each consumer has a unit demand. On the producers' side, there is a continuum of risk-neutral firms with total measure 1. Each firm produces a single product and the marginal cost of production is normalized to 0. Consumers are of $N \geq 2$ types, labeled as t_1, t_2, ..., t_N. Consumers of different types have

[6] The super star phenomenon refers to the scenario that the most popular products gain market shares.

[7] Other differences are that, while in his model there is only a monopolist firm and the variety of goods offered is exogenously given, in our model we have competing firms and the variety of goods offered (the equilibrium coverage of consumer types) is endogenous.

[8] Somewhat related to targeted search, a recent paper by Bergemann and Bonatti (2011) studies the effects of targeted advertising on media markets.

distinctive tastes. The proportion of type t_n consumers is α_n, with $\sum_{n=1}^{N} \alpha_n = 1$. We assume that α_n is strictly decreasing in n, that is, $\alpha_1 > \alpha_2 ... > \alpha_N$. With this formulation, as n increases, type t_n consumers become gradually from popular types (mainstream consumers) to less popular types (long tail consumers). The distribution of consumer types is common knowledge. Each firm has to decide which type of consumers to serve by making its product cater to one particular type of consumers, and it can at most serve one type of consumers. A firm serving type t_n consumers is labeled as type T_n. We assume that consumers know their own types, but firms cannot observe consumers' type who visit them. The above assumptions imply that consumers know whether they are mainstream consumers or long tail consumers.

Consumers have to search for products. We assume that each consumer searches sequentially, with per search cost $s > 0$. If a consumer searches M times, he incurs total search cost of $M \times s$. If a consumer l of type t_i buys from firm k of type T_j, then his gross utility (net of search costs) is

$$u_{lk}(t_i, T_j) = \begin{cases} -p_k + \varepsilon_{lk} & \text{if } i = j \\ -p_k & \text{if } i \neq j \end{cases}, \tag{1}$$

where p_k is the price charged by firm k, and ε_{lk} is the match value between consumer l and firm k. The random variable ε_{lk} has a density function $f(\varepsilon)$, cumulative distribution function $F(\varepsilon)$, and support $[a, b]$, with $b > a > 0$. We assume the density function f is log concave, which is standard in the literature. Moreover, ε_{lk} is i.i.d across consumers and firms.

In the formulation of consumers' preferences (1), a type t_n consumer derives positive utility only if he buys from a T_n firm, and he derives 0 utility if he buys from a T_j firm with $j \neq n$. Moreover, there are variations of the match value between a t_n type consumer and a T_n firm, which is captured by the term ε_{lk}. The interpretation of the underlying preference is as follows. Different types of consumers demand goods of different categories and a particular type of consumers only derives positive utility from goods of a particular category. Among the firms that provide the right (or relevant) category of goods to a particular consumer, the degree to which the consumer likes the products varies across firms. To illustrate the idea, we use books (novels) as an example. At the category level, some consumers only want to read detective stories (DS) and some consumers only want science fiction (SF). This distinction defines two consumer types: the type of DS and the type of SF. At the book (firm) level, a particular consumer of type, say, DS, likes different DS books to different degrees. To summarize, the type of a consumer defines the category of goods that he wants, and there is no substitution among different categories.[9] Within the right category, consumers of the same type still have different tastes regarding different firms' products. As a result, consumers are

[9] The case with possible substitution among different categories is discussed in the conclusion.

not only searching for the right category of goods, but also searching for the right products within the right category.

Let β_n be the fraction of type T_n firms and p_n be the price (or a price distribution) charged by type T_n firms. The timing is as follows. In the first stage, firms simultaneously determine their types by choosing which types of consumers to serve. In the second stage, the type distribution of firms, $\{\beta_n\}$, becomes publicly known. Then firms simultaneously choose the prices $\{p_n\}$. Finally, rationally anticipating firms' prices, consumers conduct search and buy goods.

Consumers' search are not completely random. Denote ϕ_n as a type t_n consumer's probability of encountering a type T_n firm in each search. In particular, we assume that

$$\phi_n = \frac{\beta_n}{\beta_n + \sigma}, \tag{2}$$

where $\sigma > 0$ is a constant with the restriction that $\beta_n + \sigma \leq 1$ for all n. The underlying rationale of (2) is as follows. Before conducting search, a type t_n consumer has a *targeted set* of firms, and then searches randomly among the firms in the targeted set. In particular, the targeted set of firms include all the firms of the right category (all type T_n firms), and some firms of irrelevant categories with measure σ.[10] We can reasonably think that the targeted set of firms is generated by the technology of the internet. In the example of novels, with internet search engines a DS type consumer can type in the keyword Detective Stories, then he will be directed to all detective stories plus some novels of other categories. The term σ captures how refined the targeted set is, or the *targetibility* of search. A bigger σ means that the targeted set includes more firms of irrelevant categories, thus search has a lower targetibility. On the other hand, a smaller σ implies higher targetibility of search. Note that if $\sigma = 0$, then for all consumers the probability of finding the right category, ϕ_n, becomes 1. That is, search becomes completely targeted. In the other extreme, if $\beta_n + \sigma$ is always 1, then the targeted set includes all firms and search becomes completely random. As σ decreases, ϕ_n increases. From the formulation of (2), we see that as β_n increases, ϕ_n increases as well.[11] Given that the targeted set always contains the same measure of firms of irrelevant categories, an increase in the fraction of the firms of the right category would increase the chance of hitting a firm of

[10] It does not matter whether two consumers of the same type share the same set of irrelevant firms.

[11] An alternative way to model search targetibility is as follows:

$$\phi_n = \frac{\gamma\beta_n}{\gamma\beta_n + \sigma},$$

where $\gamma \in (0, 1]$. That is, the targeted set includes γ proportion of firms of the right category, and a measure σ of irrelevant firms. Keeping σ constant, search targetibility can be captured by the parameter γ. As γ increases, ϕ_n increases, or the targetibility of search increases. In this formulation, ϕ_n is increasing in β_n as well. The main results of this paper hold qualitatively under this alternative formulation.

the right category.[12]

We will focus on symmetric equilibria in the sense that firms of the same type will charge the same price. In other words, p_n is degenerate.

3 Search and Price

3.1 Consumers' search behavior

Suppose firms' type distribution is $\{\beta_n\}$ and consumers expect that the prices charged are $\{p_n^*\}$. Consider a type t_n consumer whose current utility is u_n if he stops searching. Now suppose he samples one more firm. With probability ϕ_n the new firm is a T_n type firm, and the consumer will prefer the new product if $\varepsilon - p_n^* \geq u_n$, with a utility gain $\varepsilon - (u_n + p_n^*)$. With probability $1 - \phi_n$, he encounters a firm not of type T_n and earns nothing. Therefore, the expected gain from one additional search is:

$$\frac{\beta_n}{\beta_n + \sigma} \int_{u_n + p_n^*}^{b} (\varepsilon - u_n - p_n^*) f(\varepsilon) d\varepsilon \equiv \phi_n g(u_n + p_n^*). \tag{3}$$

Searching one more firm is worthwhile if and only if the expected search gain is bigger than the search cost s. Or equivalently, a t_n type consumer will stop searching if and only if $u_n \geq \overline{u}_n$, where $\overline{u}_n$, the reservation utility for type t_n, is implicitly defined as

$$\phi_n g(\overline{u}_n + p_n^*) = s. \tag{4}$$

Define $\widehat{x}_n \equiv \overline{u}_n + p_n^*$. Now (4) can be rewritten compactly as $\phi_n g(\widehat{x}_n) = s$. We can interpret $\widehat{x}_n$ as type t_n consumers' *reservation match value* (in terms of ε). From (3), we can see that $g(x)$ is strictly decreasing in x. Thus, there is at most one $\widehat{x}_n$ (at most one $\overline{u}_n$ given p_n^*) satisfying (4).

Lemma 1 *(i) The reservation match value, $\widehat{x}_n$, is increasing in ϕ_n, increasing in β_n, and decreasing in σ; (ii) if β_n is close enough to* 0*, then type t_n consumers will not search.*

Proof. As ϕ_n increases, by (4) $g(\widehat{x}_n)$ must decrease. Given that $g(x)$ is strictly decreasing in x, $\widehat{x}_n$ must increase. Since ϕ_n is increasing in β_n and decreasing in σ, $\widehat{x}_n$ is increasing in β_n and decreasing in σ. This proves part (i). To show part (ii), note that to induce consumers of type t_n to search, $\overline{u}_n$ must be positive. Given that $p_n^* \geq 0$, $g(\widehat{x}_n)$ has an upper bound $E(\varepsilon)$, which is finite. If β_n is close enough to 0, hence ϕ_n is close to 0 (less than $s/E(\varepsilon)$), then there is no $\overline{u}_n \geq 0$ satisfying (4) and type t_n consumers will not search. ■

[12] Using an analogy, an increase in β_n means that the signal to noise ratio, β_n/σ, increases. Hence search is more likely to hit the relevant category.

A bigger reservation match value, $\widehat{x}$, means that consumers are more demanding in terms of stopping searching, hence the search intensity is higher, or search more on average within the right category. According to Lemma 1, the search intensity (within the right category) is increasing in the probability of finding the firms of the right category. This is because the expected gain from search is increasing in the probability of finding firms of the right category. If the probability of finding the right category of firms is low enough, consumers will not bother searching. As the targetibility of search increases (σ decreases), all types of consumers will have higher probabilities of finding the right category, and they will search more within the right category.

In general, a bigger β_n or ϕ_n does not imply less overall search (in expectation). Although a bigger ϕ_n implies that the consumer is more likely to encounter firms of the right category in each search, it will lead to a higher reservation match value $\widehat{x}_n$, which means that the consumer will search more within the right category. Thus whether a bigger ϕ_n will lead to more overall search is ambiguous. To see this more clearly, let ρ_n be a type t_n consumer's probability of buying after each round of search. The expected length of search for that type is simply $1/\rho_n$.[13] Therefore, a smaller ρ_n implies more overall search (in expectation). In particular, $\rho_n = \phi_n(1 - F(\widehat{x}_n))$. Now (4) can be written as

$$\rho_n \frac{g(\widehat{x}_n)}{1 - F(\widehat{x}_n)} = s. \tag{5}$$

Since $\widehat{x}_n$ is increasing in ϕ_n, by (5) an increase in ϕ_n leads to less overall search (an increase in ρ_n) if and only if $\frac{g(\widehat{x}_n)}{1-F(\widehat{x}_n)}$ is decreasing in $\widehat{x}_n$. However, the logconcavity of $f(\varepsilon)$ cannot pin down whether $\frac{g(\widehat{x}_n)}{1-F(\widehat{x}_n)}$ is decreasing in $\widehat{x}_n$. What we can show is that when $f(\varepsilon)$ is uniformly distributed, $\frac{g(\widehat{x}_n)}{1-F(\widehat{x}_n)}$ is decreasing in $\widehat{x}_n$, thus a higher ϕ_n implies less overall search. When $f(\varepsilon)$ has an exponential distribution, $\frac{g(\widehat{x}_n)}{1-F(\widehat{x}_n)}$ is constant, and the amount of overall search is independent of ϕ_n.

3.2 Firms' Pricing behavior

Each firm has two decisions to make: which type of consumers to serve by choosing type T_n, and what price to charge by choosing p_n. In this subsection we pin down type T_n firms' equilibrium price, p_n^*.

For that purpose, we first derive a type T_n firm's demand whose price is p_n, given that all other T_n firms charge p_n^* and type t_n consumers' reservation utility is $\overline{u}_n$. If a type t_n consumer visits a T_n firm, he buys from this firm if and only if $\varepsilon - p_n > \overline{u}_n$. So the probability of purchase from the firm in question is $1 - F(\overline{u}_n + p_n)$. Given that all other type T_n firms charge p_n^*,

[13] Specifially, the expected length of search can be expressed as $\sum_{t=1}^{\infty} t\rho_n(1-\rho_n)^{t-1}$.

if a type t_n consumer visits such a T_n firm, the probability of the consumer purchasing from that firm is $1 - F(\overline{u}_n + p_n^*) \equiv \rho_n$. Now consider the firm in question. In the first period, a number of $\frac{m\alpha_n}{\beta_n+\sigma}$ type t_n consumers visit the firm.[14] After the first period, a measure of $m\alpha_n(1-\rho_n\phi_n)$ type t_n consumers do not stop searching. As a result, in the second period a number of $\frac{m\alpha_n}{\beta_n+\sigma}(1-\rho_n\phi_n)$ type t_n consumers will visit the firm. By the same logic, in third period a number of $\frac{m\alpha_n}{\beta_n+\sigma}(1-\rho_n\phi_n)^2$ type t_n consumers will visit the firm, and so on. Summing up all the visits, we derive the following demand for a T_n firm which charges p_n:

$$\frac{m\alpha_n}{\rho_n\beta_n}[1 - F(\overline{u}_n + p_n)], \tag{6}$$

with profit

$$\Pi_n = \frac{m\alpha_n}{\rho_n\beta_n} p_n [1 - F(\overline{u}_n + p_n)]. \tag{7}$$

Note that ρ_n does not depend on p_n, the price charged by the firm in question. The profit maximizing price p_n^* is given by the first order condition:

$$p_n^* = \frac{1 - F(\overline{u}_n + p_n^*)}{f(\overline{u}_n + p_n^*)} = \frac{1 - F(\widehat{x}_n)}{f(\widehat{x}_n)}. \tag{8}$$

Lemma 2 *For each type n, (i) given β_n, the profit maximizing p_n^* and consumers' reservation utilities $\overline{u}_n$ are unique; (ii) p_n^* is decreasing in $\widehat{x}_n$ and $\overline{u}_n$ is increasing in $\widehat{x}_n$; (iii) p_n^* is decreasing and $\overline{u}_n$ is increasing in β_n, and p_n^* is increasing and $\overline{u}_n$ is decreasing in σ.*

Proof. Since $f(\varepsilon)$ is logconcave, $\frac{1-F(\varepsilon)}{f(\varepsilon)}$ is strictly decreasing in ε.[15] This implies that, given $\widehat{x}_n$, there is a unique p_n^* satisfying (8) and p_n^* is decreasing in $\widehat{x}_n$. Since $\overline{u}_n = \widehat{x}_n - p_n^*$, $\overline{u}_n$ is uniquely determined as well given $\widehat{x}_n$. By (4), $\widehat{x}_n$ is uniquely determined given β_n. Therefore, p_n^* and $\overline{u}_n$ are uniquely determined given β_n. Moreover, $\overline{u}_n$ is increasing in $\widehat{x}_n$. This proves part (i) and (ii). Part (iii) follows immediately from Lemma 1 and part (ii). ■

The results of Lemma 2 are intuitive. An increase in reservation match value means that consumers will search more within the right category, and with a log concave density function, each firm's demand becomes more elastic. As a response, firms' equilibrium price decreases. This tends to increase consumers' reservation utility. Since both an increase in the fraction of firms of the right category (β_n) and an increase in search targetibility (a reduction in σ) tend to increase consumers' reservation match value, both would lead to a decrease in equilibrium price and an increase in consumers' reservation utility.

[14] The total measure of type t_n consumers is $m\alpha_n$, and each consumer searches randomly in the targeted set of firms with meaure $\beta_n + \sigma$. Therefore, each type T_n firm gets a number of $\frac{m\alpha_n}{\beta_n+\sigma}$ type t_n consumers in the first period.

[15] See Bagnoli and Bergstrom (2005).

4 Market Equilibrium

A market equilibrium is characterized by firms' type distribution $\{\beta_n\}$, firms' optimal prices $\{p_n^*\}$, and consumers' reservation utilities $\{\overline{u}_n\}$ such that:

(i) Given $\{\beta_n\}$ and $\{p_n^*\}$, for each type t_n, type t_n consumers' optimal search behavior leads to $\overline{u}_n$;

(ii) Given consumer's optimal search behavior $\{\overline{u}_n\}$ and firms' type distribution $\{\beta_n\}$, the profit maximizing prices are $\{p_n^*\}$.

(iii) In the first stage, given firms' type distribution $\{\beta_n\}$, no firm of any type T_n has an incentive to deviate to becoming another type.

Since consumers' search behavior depends on firms' type distributions, potentially there could be multiple equilibria. Denote $I = \{n : \beta_n > 0\}$. That is, I is the set of consumer types that are served, which we call the *inclusion set.* In one extreme, I contains only a single element n. That is, all firms choose to be type T_n and only type t_n consumers are served. We call such equilibria *pure exclusive T_n equilibria.* In the other extreme, I contains all N elements. That is, all β_n's are strictly positive and all types of consumers are served. We call such equilibria as *all inclusive equilibria.* In between, I might contain at least two but not all elements. That is, more than two types of consumers are served but some type(s) of consumers are excluded. Denote an equilibrium associated with an inclusion set I as $\{\beta_n^I\}$, $\{p_n^{*I}\}$, and $\{\overline{u}_n^I\}$ for $n \in I$.

4.1 Characterizing equilibria

We start by characterizing equilibria, assuming they exist. Consider an inclusion set I. For $n \in I$, the expressions of firm's profits, (7), can be simplified as:

$$\Pi_n^{*I} = m\frac{\alpha_n}{\beta_n^I}p_n^{*I}. \tag{9}$$

By the analysis in the previous sections, equilibrium requirements (i) and (ii) can be explicitly written as (for $n \in I$):

$$\frac{\beta_n^I}{\beta_n^I + \sigma}g(\overline{u}_n^I + p_n^{*I}) = s;\ p_n^{*I} = \frac{1 - F(\overline{u}_n^I + p_n^{*I})}{f(\overline{u}_n^I + p_n^{*I})}. \tag{10}$$

Regarding equilibrium requirement (iii), there are two kinds of deviations to worry about. First, any included type T_n, $n \in I$, should have no incentive to deviate to an excluded type $T_{n'}$, $n' \notin I$. This kind of deviation is clearly not profitable. This is because a single firm's deviation to type n' will not induce type $t_{n'}$ consumers to search, thus deviation will lead to zero profit, while a positive profit is guaranteed if a firm remains as the current type T_n. Second, any included type T_n, $n \in I$, should have no incentive to deviate to another included type $T_{n'}$, $n' \in I$. To prevent this kind of deviations, all firms that serve any included types of consumers should get the same profit. That is, for any $n \in I$ and $n' \in I$, $n \neq n'$, in equilibrium the profit of a T_n type firm must equal to that of a $T_{n'}$ type firm:

$$\Pi_n^{*I} = \Pi_{n'}^{*I} \Leftrightarrow \frac{\alpha_n}{\beta_n^I} p_n^{*I} = \frac{\alpha_{n'}}{\beta_{n'}^I} p_{n'}^{*I} \tag{11}$$

Proposition 1 *(i) For any configuration of I, there is at most one equilibrium. (ii) In the equilibrium with more than two types served (I contains more than two elements), β_n^I is decreasing in n, p_n^{*I} is increasing in n, $\overline{u}_n^I$ is decreasing in n, and the consumer to firm ratio, $\frac{\alpha_n}{\beta_n^I}$, is decreasing in n. That is, there are more firms serving mainstream consumers, but there are less firms per consumer for mainstream consumers; firms serving mainstream consumers charge lower prices and have more sales than those serving long tail consumers; mainstream consumers enjoy higher utilities than long tail consumers.*

Proof. Note that given $\{\beta_n\}$, and hence $\{\phi_n\}$, by Lemma 1 $\{\widehat{x}_n\}$ are uniquely determined, and $\{p_n^*\}$ and $\{\overline{u}_n\}$ are uniquely determined following Lemma 2. Therefore, to show the uniqueness of equilibrium for any I, we only need to show the uniqueness of $\{\beta_n\}$ in equilibrium. First consider the case that I only contains a single element n (pure exclusive T_n equilibria). With this configuration, $\beta_n = 1$. It is obvious that the equilibrium is unique. Next consider the case that I contains more than two elements. Suppose, with inclusion set I, $\{\beta_n^I\}$ is an equilibrium distribution of firms' types, and $\{\beta_n^{I\prime}\} \neq \{\beta_n^I\}$ is another equilibrium distribution. Without loss of generality, suppose for some $i \in I$, $\beta_i^{I\prime} > \beta_i^I$. Given that $\sum_{n\in I} \beta_n^I = 1$ and $\sum_{n \in I} \beta_n^{I\prime} = 1$, there must be a $j \neq i$ and $j \in I$ such that $\beta_j^{I\prime} < \beta_j^I$. Since $\beta_i^{I\prime} > \beta_i^I$, by Lemma 1 we have $\widehat{x}_i^{I\prime} > \widehat{x}_i^I$, which by Lemma 2 implies that $p_i^{*I\prime} < p_i^{I*}$. Now the facts that $\beta_i^{I\prime} > \beta_i^I$ and $p_i^{*I\prime} < p_i^{*I}$ lead to $\Pi_i^{*I\prime} < \Pi_i^{I*}$. By similar logic, $\beta_j^{I\prime} < \beta_j^I$ implies that $\Pi_j^{*I\prime} > \Pi_j^{I*}$. Combining the above results with the equal profit condition (11) for $\{\beta_n^I\}$, we have $\Pi_i^{*I\prime} < \Pi_i^{*I} = \Pi_j^{*I} < \Pi_j^{*I\prime}$, which contradicts the equal profit condition (11) for $\{\beta_n^{I\prime}\}$. Therefore, if an equilibrium with inclusion set I exists, it must be unique. This proves part (i).

Now consider the equilibrium with an inclusion set I containing more than two elements. Let $n \in I$, $n' \in I$, and $n' > n$. We first show that $\beta_n^I > \beta_{n'}^I$. Suppose $\beta_n^I \leq \beta_{n'}^I$. Since $\alpha_n > \alpha_{n'}$, by the equal profit condition (11) we must have $p_n^{*I} < p_{n'}^{*I}$. By Lemma 1, $\beta_n^I \leq \beta_{n'}^I$ implies that $\widehat{x}_n^I \leq \widehat{x}_{n'}^I$. Since, by Lemma 2, p_n^{*I} is decreasing in $\widehat{x}_n^I$, it follows that $p_n^{*I} \geq p_{n'}^{*I}$. Thus we got the requisite contradiction, and $\beta_n^I > \beta_{n'}^I$ must hold.

Given that $\beta_n^I > \beta_{n'}^I$, by Lemma 1, $\widehat{x}_n^I > \widehat{x}_{n'}^I$. Since by Lemma 2 p_n^{*I} is decreasing in $\widehat{x}_n^I$, it follows that $p_n^{*I} < p_{n'}^{*I}$. Given that $p_n^{*I} < p_{n'}^{*I}$, by the equal profit condition (11) we must have $\frac{\alpha_n}{\beta_n^I} > \frac{\alpha_{n'}}{\beta_{n'}^I}$. Finally, since by Lemma 2, p_n^{*I} is decreasing in $\overline{u}_n^I$, $p_n^{*I} < p_{n'}^{*I}$ implies that $\overline{u}_n^I > \overline{u}_{n'}^I$. This proves part (ii). ∎

The intuition for part (ii) of Proposition 1 is as follows. Suppose there are more firms serving a long tail type of consumers than those serving a mainstream type. This world lead to two effects. On the one hand, the firms serving the mainstream type have more sales per firm than those serving the long tail type. On the other hand, the long tail type of consumers will search more than the mainstream type, and thus firms serving the mainstream type can charge a higher price than those serving the long tail type do. Combining these two effects, firms serving the mainstream type will earn a strictly higher profit than those serving the long tail type, which violates the equal profit condition and cannot be an equilibrium. Therefore, in equilibrium there must be more firms serving the mainstream type than those serving the long tail type. Given that there are more firms serving the mainstream type, mainstream consumers will search more (within the right category) than long tail consumers do, leading to a lower price charged by firms serving the mainstream type. Now to restore the equal profit condition, firms serving the mainstream type must have higher sales per firm than those serving the long tail type.

Proposition 1 shows that, among the types served, mainstream consumers always enjoy higher utility than long tail types do. In other words, mainstream consumers are better off simply by the fact that their tastes are shared by more people, and long tail consumers suffer simply by the fact their tastes are shared by fewer people. In particular, the benefit of mainstream consumers come from two sources: it is easier for them to find products of the right category, and those products are cheaper. Another interesting feature is regarding the distribution of firms. Among the covered types, although there are more firms serving mainstream types, compared to the distribution of consumer types the distribution of firms is skewed toward long tail types, as each long tail consumer brings a higher profit (price) than a mainstream consumer does.[16]

4.2 The existence of equilibrium and equilibrium selection

For an equilibrium with configuration I to exist, all the included types of consumers must have incentives to search. More formally, in an equilibrium with configuration I, for all $n \in I$, $\overline{u}_n^I \geq 0$. To ensure that some equilibrium exists, we make the following assumption: if a type t_n consumer encounters a T_n firm with probability 1 in each search, then he has an incentive to search. This assumption ensures that pure exclusive equilibria exist. More formally, Define

$$h(x) \equiv x - \frac{1 - F(x)}{f(x)}$$

[16] Following the discussion in subsection 3.1, in general it is not clear whether mainstream consumers (who have a higher ϕ_n) will conduct less overall search. If $f(\varepsilon)$ is uniformly distributed, then mainstream consumers conduct less overall search than long tail consumers. If the distribution of $f(\varepsilon)$ is exponential, then all covered types conduct the same amount of search.

Note that $h(x)$ is strictly increasing in x. Let $\widehat{x}$ be such that $g(\widehat{x}) = s$. We assume the following condition holds throughout the paper:

$$\widehat{x} \geq a \text{ and } h(\widehat{x}) \geq 0 \tag{12}$$

Note that condition (12) is satisfied if the search cost s is small enough.

Given that assumption (12) holds, there are definitely multiple equilibria. In particular, all pure exclusive equilibria exist. To see this, assumption (12) ensures that type t_n consumers have incentive to search in the T_n pure exclusive equilibrium.[17] Since there are no firms other than the T_n type, all the other types of consumers will not search. And this means that each individual firm has no incentive to deviate to other types. Thus any T_n pure exclusive equilibrium exists. A generalization of the above logic is that, once a particular type t_n of consumers is excluded, we do not need to worry about t_n type consumers' deviation to searching and firms' deviation to becoming T_n type, since to make such deviations profitable requires joint deviations of firms and consumers. This self-confirming feature of exclusion naturally leads to the multiplicity of equilibria.

The above discussion shows that an equilibrium with configuration I exists if and only if for all $n \in I$, $\overline{u}_n^I \geq 0$. Since, by Proposition 1, $\overline{u}_n^I$ is decreasing in n, the existence of equilibrium boils down to the condition that the least mainstream type covered has an incentive to search, or $\overline{u}_{\overline{n}}^I \geq 0$, where $\overline{n}$ is the largest element that belongs to I.

Lemma 3 *There exists a $\widehat{\beta} \in (0,1)$ such that an equilibrium with configuration I exists if and only if $\beta_{\overline{n}}^I \geq \widehat{\beta}$.*

Proof. Define $\widehat{\widehat{x}}$ such that $h(\widehat{\widehat{x}}) = 0$. Given assumption (12) and that fact that $h(x)$ is increasing in x, $\widehat{\widehat{x}}$ is uniquely defined. Moreover, by Lemma 2, $\overline{u} \geq 0$ if and only if $\widehat{x} \geq \widehat{\widehat{x}}$. By Lemma 1, $\widehat{x} \geq \widehat{\widehat{x}}$ is equivalent to $\phi \geq \widehat{\phi} \in (0,1)$, where $\widehat{\phi}$ is defined as $\widehat{\phi} g(\widehat{\widehat{x}}) = s$, which is uniquely defined by the monotonicity of $g(\cdot)$. Since ϕ_n is increasing in β_n, $\phi \geq \widehat{\phi}$ is equivalent to $\beta \geq \widehat{\beta}$, where $\widehat{\beta} \in (0,1)$ is defined as $\widehat{\phi} = \frac{\widehat{\beta}}{\widehat{\beta}+\sigma}$. Therefore, $\overline{u} \geq 0$ if and only if $\beta \geq \widehat{\beta}$. Now the condition ensures the existence of the equilibrium with configuration I, $\overline{u}_{\overline{n}}^I \geq 0$, is equivalent to $\beta_{\overline{n}}^I \geq \widehat{\beta}$. ■

Let z be the number of elements in I, or the number of consumer types served. Note that $z \in \{1, 2, ..., N\}$. For pure exclusive equilibria, $z = 1$, and for the all inclusive equilibrium, $z = N$. For z such that $1 < z < N$, there are more than one possible configurations of I that have the same z. Among the possible configurations, we are interested in one particular configuration.

[17] By the definition of $h(x)$, $\overline{u}_n^I(\widehat{x}) = h(\widehat{x})$ in T_n pure exclusive equilibrium.

Definition 1 *A configuration I is said to be monotonic if $n \in I$ implies that, for any n' such that $1 \leq n' < n$, $n' \in I$. A monotonic configuration I that contains z elements is called a z-monotonic configuration.*

In monotonic configurations, (relatively) mainstream types are covered and (relatively) long tail types are excluded. In a z-monotonic configuration all the first z mainstream types of consumers are served, while the last $N - z$ (long tail) types are excluded. Note that the T_1 pure exclusive equilibrium has the 1-monotonic configuration, and the all inclusive equilibrium has the N-monotonic configuration. Moreover, given z there is a unique z-monotonic configuration.

The next Lemma shows that, compared to non-monotonic configurations, firms have higher profits in equilibria with monotonic configurations.

Lemma 4 *For any z, $1 \leq z < N$, among all the configurations having the same z, firms' profits are highest in the equilibrium of z-monotonic configuration.*

To understand Lemma 4, notice that monotonic configurations always include the most popular (mainstream) types of consumers. This implies that firms can spread out relatively evenly across included types under monotonic configurations. Since some relatively less popular (long tail) types of consumers are included in non-monotonic configurations, firms' type distribution will skewed toward more popular types, as the segments of less popular types can accommodate fewer firms. With more firms congested among popular types, those firms have lower sales per firm, and popular consumer types will search more within the right category, which results in lower price charged. Both effects lead to lower profits.

To resolve the issue of multiple equilibria, we have to impose some equilibrium selection criterion. One natural criterion is to select the equilibrium with the highest profit for firms. The rationale is that firms will most likely to coordinate on the equilibrium with highest profits.[18] The result of Lemma 4 suggests that, with such a criterion, an equilibrium with monotonic configuration will always be selected. However, this is not true for the following reason. To maximize profits, firms have two tendencies. First, they try to cover as many consumers as possible, since doing that can increase sales per firm. This means that mainstream consumers are more likely to be covered. Second, fixing the total measure of consumers covered, firms tend to cover as many types as possible. By spreading over more types (segments), in each covered segment consumers will search less and firms can charge higher prices. In some sense, spreading over more segments increases product differentiation and softens competition. This tendency implies that mainstream consumers may not be necessarily covered. To see this, note that it is possible that a z-monotonic equilibrium does not exist, but an equilibrium with a

[18] Given parameter values, it is hard to characterize the equilibrium under which firms get the highest profit among all possible equilibria. This is because it depends on the distribution of types, $\{\alpha_n\}$.

non-monotonic configuration which has z elements exists. This is because, generally, including a more mainstream type would cause the distribution of firms skewed more toward mainstream types, leaving fewer firms covering the long tail types, which discourages long tail consumers from searching. Therefore, given parameter values, it is possible that the equilibrium with highest profits among all equilibria has a non-monotonic configuration.[19]

As mentioned before, the model has the flavor of coordination games due to the self-confirming feature of exclusion. To select a reasonable equilibrium, we have to resort to joint (or coordinated) deviations (among firms in the first stage game). Specifically, we introduce the concept of stability.

Definition 2 *An equilibrium with firm distribution $\{\beta_n\}$ is said to be stable if, for any n, any joint deviation to type T_n by any measure of firms which are currently not of type T_n is not profitable.*

For an equilibrium with more than two included types, to check whether it is stable we do not need to worry about the deviations to already included types, since such deviations will not be profitable. This is because deviating to an already included type will reduce sales per firm and the price of that type (due to more search of that consumer type), leading to lower profits than what deviating firms can get by remaining as the original types. Therefore, we only need to worry about joint deviation to the excluded types.

Later on we will show that equilibria with non-monotonic configurations are not stable. This implies that we can focus on equilibria with monotonic configurations. Before discussing the stability of equilibria, we first analyze equilibria of monotonic configurations. In the subsequent notation, a superscript z denotes a z-monotonic configuration.

Proposition 2 *(i) Firms' profits in the equilibrium of z-monotonic configuration, Π^{*z}, is increasing in z. (ii) For $n \leq z$, p_n^{*z} is increasing in z, and both β_n^z, $\overline{u}_n^z$ are decreasing in z; both β_z^z, $\overline{u}_z^z$ are decreasing in z. That is, for the mainstream types that are already covered, including more types lead to higher prices and fewer firms serving those types.*

Proof. Part (i). Consider a z-monotonic configuration, and a $(z+1)$-monotonic configuration, with $1 \leq z < N$. Since $\beta_{z+1}^{z+1} > 0$, $\sum_{n=1}^{z} \beta_z^{z+1} < 1$. Given that $\sum_{n=1}^{z} \beta_z^z = 1$, there must be some $k \leq z$ such that $\beta_k^{z+1} < \beta_k^z$. Now following Lemmas 1 and 2, we have $p_k^{*(z+1)} > p_k^{*z}$.

[19] Here we provide an example in which the equilibrium with highest profit has a non-monotonic configuration. Suppose there are three types, with $\alpha_1 = 0.34$, $\alpha_2 = \alpha_3 = 0.33$. The other parameter values are such that $\widehat{\beta} = 0.5 - \epsilon$, with ϵ being positive but very small. It is easy to see that a monotonic equilibrium including types 1 and 2 does not exist, as β_2 will be less than $\widehat{\beta}$. The only monotonic equilibrium is the one that only includes type 1. However, the equilibrium including only type 2 and 3 exists, since in such case $\beta_2 = \beta_3 = 0.5 > \widehat{\beta}$. It is not difficult to see that the equilibrium with only type 2 and 3 being covered yields a higher profit for firms than the equilibrium covering only type 1.

This implies that $\Pi_k^{*(z+1)} = m\frac{\alpha_k}{\beta_k^{z+1}}p_k^{*(z+1)} > m\frac{\alpha_k}{\beta_k^z}p_k^{*z} = \Pi_k^{*z}$. Therefore, by the equal profit condition, all firms have a higher profit in the equilibrium of $(z+1)$-monotonic configuration.

Part (ii). By the results in part (i), for any $n \leq z$, we have $\Pi_n^{*(z+1)} > \Pi_n^{*z}$. Following an argument similar to previous proofs, this condition implies that $\beta_n^{z+1} < \beta_n^z$, which further implies that $p_n^{*z+1} > p_n^{*z}$ and $\overline{u}_n^{z+1} < \overline{u}_n^z$. By Proposition 1, $\beta_z^{z+1} > \beta_{z+1}^{z+1}$. By Part (i), $\beta_z^z > \beta_z^{z+1}$. Thus $\beta_z^z > \beta_{z+1}^{z+1}$, which implies that $\overline{u}_z^z > \overline{u}_{z+1}^{z+1}$. ■

The results of Proposition 2 are intuitive. Including one more type means that fewer firms will be serving the previously included types, as some firms switch to serving the newly included type. This leads to two effects. First, sales per firm would increase. Second, for the previously included types of consumers, the probability of finding the right category of firms decreases. As a result, they will search less intensively within the right category and firms now can charge higher prices. Both effects tend to increase firms' profits. The second effect also makes the previously included types of consumers worse off.

Proposition 3 *Given parameter values, (i) any equilibrium with a non-monotonic configuration is not stable; (ii) there is a unique stable equilibrium, which is the monotonic equilibrium with the biggest z; such z^* is determined by $\beta_{z^*}^{z^*} \geq \widehat{\beta}$ and $\beta_{z^*+1}^{z^*+1} < \widehat{\beta}$.*

The underlying reason for non-monotonic equilibria being not stable is that mainstream types are more profitable for firms. Monotonic equilibria with the number of types covered less than z^* are not stable because covering more types tends to increase firms' profits. The monotonic equilibrium with the largest number of types covered is stable because no more types can be possibly covered. More specifically, the measure of each remaining long tail type is so small such that a measure of $\widehat{\beta}$ firms (jointly) deviating to becoming that type is not profitable.

In the subsequent analysis, we will focus on the unique stable equilibrium, or the monotonic equilibrium with the largest number of types being covered. To abuse terminology, we will simply call it the *market equilibrium*. Note that long tail consumers might be excluded in the market equilibrium. The underlying reason is that search is not perfectly targeted. If there are only a few firms serving a long tail type, the expected gain from searching is low as the probability of finding the firms of the right category is low. As a result, long tail consumers might simply not search. Anticipating this, if the measure of a long tail type of consumers is too low, firms might just exclude that type.

5 Equilibrium Properties and Comparative Statics

Applying the results of Proposition 1, we conclude that the following properties hold in the market equilibrium. Among the first z^* mainstream types covered, prices are strictly increasing and consumers' reservation utilities are strictly decreasing as types move toward less mainstream types.[20] Moreover, among covered types, although more firms are serving more mainstream types, the distribution of firms is skewed toward less mainstream types relative to the distribution of consumer types.

The market equilibrium depends on the distribution of consumer types, $\{\alpha_n\}$, search costs s, and the targetibility of search embodied in σ. In the rest of this section we will study comparative statics regarding the market equilibrium.

5.1 Consumer distribution

We first study how changes in consumer distribution, $\{\alpha_n\}$, affect the market equilibrium.

Proposition 4 *(i) The number of consumer types covered in the market equilibrium, z^*, has an upper bound $\frac{1}{\widehat{\beta}}$. (ii) Consider two distributions of consumer types, $\{\alpha_n\}$ and $\{\alpha'_n\}$. Suppose in the market equilibrium under $\{\alpha_n\}$ z^* types are covered. Moreover, $\alpha_n \leq \alpha'_n$ for $n < z^*$ and $\alpha_n \geq \alpha'_n$ for $n \geq z^*$. In the market equilibrium under $\{\alpha'_n\}$ the number of types covered is less than or equal to z^*.*

Proof. Part (i). By Lemma 3 a monotonic equilibrium with z^* exists if and only if $\beta_{z^*}^{z^*} > \widehat{\beta}$. Since in equilibrium β_n is decreasing in n, $\beta_{z^*}^{z^*} < \frac{1}{z^*}$. Therefore, $z^* < 1/\widehat{\beta}$.

Part (ii). We only need to show that $\beta'_{z^*} \leq \beta_{z^*}$. Suppose $\beta'_{z^*} > \beta_{z^*}$. Then by previous results, $p'_{z^*} < p_{z^*}$, and $\Pi^{*\prime}_{z^*} < \Pi^*_{z^*}$ since $\alpha'_{z^*} \leq \alpha_{z^*}$. Given that $\beta'_{z^*} > \beta_{z^*}$, there must be a $n < z^*$ such that $\beta'_n \leq \beta_n$. This implies that $p'^*_n \geq p^*_n$. Combining with the fact that $\alpha'_n \geq \alpha_n$, we have $\Pi^{*\prime}_n \geq \Pi^*_n$. By the equal profit condition in equilibrium, this contradicts $\Pi^{*\prime}_{z^*} < \Pi^*_{z^*}$. Therefore, we must have $\beta'_{z^*} \leq \beta_{z^*}$. ■

Proposition 4 implies that as the proportions of long tail consumers decrease, or the type distribution becomes more skewed toward mainstream types,[21] in equilibrium more long tail types of consumers will be excluded. Intuitively, as the proportions of long tail consumers decrease, the long tail types can potentially accommodate fewer firms. If the mass of the accommodated firms falls below the critical mass $\widehat{\beta}$, the long tail types are simply excluded.[22]

[20] Actaully, across all types consumer utility decreases as we move toward less mainstream types. This is because for long tail types that are excluded ($n > z^*$), their utility is zero.

[21] For a concrete example, consider the following family of distributions. For $2 \leq n \leq N$, $\alpha_n = k\alpha_{n-1}$, where $k \in (0, 1)$. That is, the fraction of types decreases exponentially. As k decreases, the distrubution becomes more skewed toward maintream types.

[22] When consumers' type distribution becomes more skewed toward mainstream types, its impact on firms' profits is ambiguous. On one hand, a decrease in the number of types covered tends to decrease profits. On the other hand, an increase in the population of the most mainstream types tends to increase sales and profits.

5.2 The Long tail effect

Now we study how changes in search costs and search targetibility affect the market equilibrium, fixing the distribution of consumer types $\{\alpha_n\}$. Define M_n, $n \leq z^*$, as the market share of the sales of type T_n products in the market equilibrium. It can be readily shown that $M_n = \frac{\alpha_n}{\sum_{i=1}^{z^*} \alpha_i}$. As the number of types covered, z^*, increases, all M_n, $n \leq z^*$, decreases. In other words, the *concentration of sales* across consumer types decreases as z^* increases.

Proposition 5 *In the market equilibrium, if either the search costs s decrease, or the targetibility of search increases (σ decreases), (i) the number of types of consumers covered, z^*, will (weakly) increase, and the concentration of sales will (weakly) decrease; (ii) if z^* remains the same, then for all the previously covered types $n \leq z^*$, both p_n^* and Π_n^* decrease and $\overline{u}_n$ increases; (iii) if z^* increases, then it is possible that firms' profits increase and for all $n \leq z^*$, p_n^* increases and $\overline{u}_n$ decreases.*

Part (i) of Proposition 5 provides an explanation for the long tail effect. As search costs decrease or the targetibility of search increases, more niche products (catering to long tail consumers) become available, some previously excluded long tail consumers start to participate in the market, and sales become less concentrated as they spread to newly provided niche products. The underlying reason is that both a decrease in search costs and an increase in search targetibility encourage consumers to search. As a result, the critical mass of firms that is required to serve a particular type in order to induce search, $\widehat{\beta}$, decreases, which potentially leads to more types being covered in the market equilibrium.

Part (ii) of Proposition 5 shows that if the coverage of consumer types does not increase when search costs decrease or the targetibility of search increases (this is the case if the initial market equilibrium is already all inclusive), it will lead to lower prices, lower profits, and higher consumer utilities for all the types already covered. This is because both changes encourage consumers to search more within the right category, which intensifies competition among firms.

However, when the coverage of consumer types does increase, there is an additional countervailing effect. More types covered would soften competition by increasing product differentiation, and this effect tends to increase firms' profits and lower consumer utilities. The overall effect is ambiguous. In part (iii) of Proposition 5, we construct an example in which the second effect dominates. This implies that a decrease in search costs or an increase in search targetibility may not always be a blessing for consumers, especially when the magnitude of changes is small. In particular, when a small change of magnitude causes more long tail types to be covered, while the newly covered long tail consumers are always better off, the previously covered mainstream consumers might be worse off, as some firms switch to cover some previously excluded long tail types, which reduces mainstream consumers' chance of finding their relevant categories of products. Nevertheless, when the change in magnitude is intermediate it is also possible that firms' profits and mainstream consumers' utilities both increase: the newly covered marginal types tend to increase the average sales for firms, and this may more than compensate for the profit loss resulting from lower prices among previously covered types.

Related to part (iii) of Proposition 5, the following two interesting and ironical phenomena could arise for previously covered consumers: a decrease in search costs could lead to less overall search, and an increase in search targetibility could lead to lower probabilities of finding the relevant categories of products. The underlying reason is that the distribution of firm types is endogenously determined, and the effect of the induced change in firm distribution could reverse the direct effect of a reduction in search costs or an increase in search targetibility.

To illustrate the first phenomenon, recall that $1/\rho_n$ indicates the (expected) amount of overall search, where ρ_n is a type t_n consumer's probability of buying after each round of search. Suppose the match value ε is uniformly distributed. Now equation (5) in subsection 3.1 can be written more explicitly as

$$\rho_n \frac{b - \widehat{x}_n}{2} = s. \tag{13}$$

Now suppose s decreases slightly to s' and the type coverage is increased to $z^* + 1$. By part (iii) of Proposition 5, for any $n \leq z^*$, $\beta_n' < \beta_n$. Thus $\phi_n' < \phi_n$. Following (13) and the fact that $s \simeq s'$, we have $\widehat{x}_n' < \widehat{x}_n$ and $\rho_n' > \rho_n$. That is, under s' type t_n consumers conduct less overall search.[23] The second phenomenon can be constructed in a similar fashion.[24]

5.3 The difference between search costs and search targetibility

Now we study different effects induced by a decrease in search costs and an increase in search targetibility. Roughly speaking, there are two major differences. First, other things equal, while a decrease in search costs always tends to induce consumers to search more, an increase in search targetibility might induce consumers to search less overall. On the one hand, an increase in search targetibility makes consumers search more within the right category ($\widehat{x}_n$ increases). On the other hand, consumers now have a high chance of hitting the right category. Therefore, the overall search could increase or decrease. Second, an increase in search targetibility tends to reduce the difference between mainstream consumers and long tail consumers, as the probabilities of finding the right category become more equalized among consumers. It is not clear whether a decrease in search costs always has a similar effect.

[23] Another way to understand the results is as follows. Define the *effective search costs* of type t_n consumers as s/ϕ_n, the search costs divided by that type's probability of finding the relevant category. When the type coverage z^* increases, for previously covered types ϕ_n will decrease as some firms switch to cover some previously excluded types. This effect tends to increase the effective search costs, which will lower consumer utilities and discourage consumers from searching.

[24] Specifically, an induced decrease in β_n is bigger relative to the initial decrease in σ such that ϕ_n decreases.

To derive clear analytical results, in this subsection we assume that the match value, ε, is uniformly distributed on $[a, b]$. With uniform distribution, the reservation match value and prices can be written explicitly as:

$$p_n^* = (b - \widehat{x}_n) = \sqrt{\frac{2s(b-a)}{\phi_n}}. \tag{14}$$

We say that the *concentration of firms* decreases if for any two covered types n and n', with $n' > n$, $\beta_n/\beta_{n'}$ decreases. That is, firms become more evenly distributed across types when the concentration of firms decreases. Note that sales per firm for type n is given by α_n/β_n. A decrease in the concentration of firms implies that sales per firm become less evenly distributed across types, with the sales per firm of firms serving mainstream types increasing and that of firms serving long tail types decreasing.[25] The next proposition shows that changes in search costs and changes in search targetibility have different effects.

Proposition 6 *Suppose the match value ε is uniformly distributed on $[a, b]$. (i) Suppose the equilibrium type coverage, z^*, does not change. When the search costs decrease, all covered consumer types will search more overall, the ratios of prices between any two covered types will not change, and the distribution or the concentration of firms will not change either. When the search targetibility increases, among covered types consumers will search less overall, the ratio of the price of any mainstream type to that of any relatively less mainstream type will increase, and the concentration of firms will increase and sales per firm will be more evenly distributed across types. (ii) Suppose the equilibrium type coverage, z^*, increases. When the search costs decrease, the ratio of the price of any mainstream type to that of any relatively less mainstream type will decrease, and the concentration of firms will decrease and sales per firm will be less evenly distributed across types. The effects of an increase in search targetibility on price ratios and concentration of firms are ambiguous.*

The predictions of Proposition 6 can be potentially tested, which might enable us to empirically distinguish reductions in search costs from increases in search targetibility. To understand the results, first consider the case that the type coverage does not increase. A reduction in search costs induces all covered types to search more. Thus all the prices decrease, but the ratios of prices across different types remain the same.[26] Thus the distribution and the concentration of firms will not change. On the other hand, an increase in search targetibility

[25] Recall that, for covered types, the distribution of firms is skewed more toward long tail types compared to the distribution of consumer types. A decrease in firms' concentration means that the distribution of firms becomes further away from the type distribution of consumers, or firms become more evenly distributed across types.

[26] The feature that price ratios do not change has to do with the uniform distribution of the match value. For general logconcave distributions, the price ratios will depend on the densities $f(\widehat{x}_n)$ and $f(\widehat{x}_{n'})$, which might change as both $\widehat{x}_n$ and $\widehat{x}_{n'}$ decrease.

increases all consumers' probability of finding the right category. Although consumers' reservation match value will increase correspondingly, its impact on the expected length of search is dominated by the effect of the initial increase in the probability of finding the right category, leading to less overall search. In quantitative terms, an increase in search targetibility has a bigger impact on long tail types. This is because a reduction in the noise σ would increase the signal-to-noise ratio more significantly in percentage terms when the initial signal-to-noise ratio is low. As a result, although all the prices decrease, the ratios of prices of mainstream types to those of long tail types increase as well (the price dispersion across types decreases). To restore the equal profit condition, some firms will switch from serving long tail types to serving mainstream types, leading to an increase in the concentration of firms, and sales per firm will tend to be more equalized across types. In the extreme case of full targetibility ($\sigma = 0$), in the market equilibrium all firms charge the same price, all consumers receive the same utility, and the distribution of firms exactly matches the distribution of consumer types.

When the type coverage does increase, a reduction in search costs causes fewer firms to serve the previously covered types. Other things equal, this tends to increase the difference in the probabilities of finding the right category across different types. This is because a reduction in the signal would reduce the signal-to-noise ratio more significantly in percentage terms when the initial signal-to-noise ratio is low. This implies that the ratios of prices of mainstream types to those of long tail types will decrease. To restore the equal profit condition, the ratios of firms serving mainstream types to those serving long tail types have to decrease, leading to a decrease in the concentration of firms, and sales per firm will tend to be less equalized across types. When the search targetibility increases, it has two effects. On the one hand, it tends to reduce the difference in the probabilities of finding the right category across different types, the effect we just mentioned in the last paragraph. On the other hand, an increase in type coverage tends to increase the difference in the probabilities of finding the right category across different types, an effect spelled out at the beginning of the paragraph. These two effects work against each other, and the resulting firms' concentration can either increase or decrease.

While it is hard to derive clean analytical results for general distributions of the match value, we believe that a similar pattern regarding the different effects of changes in search costs and changes in search targetibility holds more or less under more general distributions. This is because the following intuition is robust: an increase in search targetibility tends to reduce the difference between mainstream consumers and long tail consumers, as the difference in the probabilities of finding the right category decreases, while a decrease in search costs in general does not have a similar effect.

In the real world, the long tail effect is more realistically caused jointly by reductions in search costs and increases in search targetibility. However, it is reasonable to think that internet technology has more impact on increasing search targetibility than on reducing search costs. Conceivably, it is easier for online technology to achieve full search targetibility than to reduce the search costs all the way to zero.[27]

6 Free entry

In this section we study the effects of free entry. To incorporate free entry, we modify the first stage game. In particular, in the first stage firms simultaneously make the following decisions: whether to enter and which type of consumers to serve if entering. Entry entails a sunk cost k. Other aspects of the model are the same as the basic model. Denote the total measure of active firms in the market as γ. Note that the consumer-to-firm ratio is m/γ. Since in the basic model the total measure of firms is exogenously fixed as 1, both $\{\beta_n\}$ and σ can be interpreted as either fractions or measures. With the total measure of active firms endogenously determined under free entry, whether the noise or the measure of irrelevant firms, σ, will change with the measure of active firms γ is crucial. We will study the following two different cases in turns: σ is independent of γ, and σ increases with γ.

6.1 The noise is independent of the measure of active firms

In this subsection, we study the case that σ is independent γ. We treat both $\{\beta_n\}$ and σ as measures. The probability of finding the right category can be written as

$$\phi_n = \frac{\beta_n}{\beta_n + \sigma}. \tag{15}$$

Observe that (15) is equivalent to (2), the expression of ϕ_n in the basic model.[28] Given I, the set of types covered, $\gamma = \sum_{n \in I} \beta_n$.

With free entry, we need to add one more equilibrium requirement: for any type belonging to the inclusion set, $n \in I$, firms should earn zero profit. More specifically, the equilibrium

[27] Specifically, the Internet reduces search costs in the following way. Previously consumers need to go to brick-mortor stores physically to check whether the products are to their liking. With the Internet, they can search products at home by clicking the links. However, even on the Internet consumers still need to check the attributes of product to see whether they are of their tastes. This means that search is still costly. Regarding search targetibility, with internet search engines consumers can simply type in the category of products they want, and then the relevant links will automatically pop up. If the search engines are refined and powerful enough, all and only the relevant links will appear and full targetibility can be approximately achieved.

[28] If we treat both $\{\beta_n\}$ and σ as fractions, then $\phi_n = \frac{\beta_n \gamma}{\beta_n \gamma + \sigma}$, which is different from (2), the expression of ϕ_n in the basic model. Note that $\{\beta_n\}$ and σ in (2) of the basic model can be interpreted as both measures and fractions.

conditions can be written as:

$$\Pi_n^* = m\frac{\alpha_n}{\beta_n}p_n^* = k, \quad (16)$$

$$p_n^* = \frac{1-F(\widehat{x}_n)}{f(\widehat{x}_n)}, \quad (17)$$

$$s = \frac{\beta_n}{\beta_n+\sigma}g(\widehat{x}_n), \quad (18)$$

where the first condition (16) is the zero-profit condition. By previous results, p_n^* is strictly decreasing in β_n. Therefore, the equilibrium β_n is uniquely determined by the above three conditions, which does not depend on the distribution of firms across other types. This is why we drop the superscript of I for the equilibrium β_n and p_n^*. In some sense, with free entry the linkage among the included types is loosened. To see this, note that in the basic model if the measure of firms serving an included type changes, it necessarily changes the measure of firms serving another type, as the total measure of firms is 1. This linkage or congestion effect no longer exists under free entry. Instead, the measure of each type of firms is pinned down by the zero profit condition, and the measures of different included types can be determined independently.

Note that Proposition 1 is not affected by free entry, but the conditions that guarantee the existence of equilibrium need to be modified. For any included type, not only should consumers have incentive to search, but also firms should earn non-negative profits. Denote the equilibrium β_n as β_n^*, which solves (16)-(18). To ensure the existence of equilibrium, we assume the following condition holds:

$$m\frac{\alpha_1}{\widehat{\beta}}\frac{1-F(\widehat{\widehat{x}})}{f(\widehat{\widehat{x}})} > k, \quad (19)$$

which makes sure that the T_1 pure exclusive equilibrium exists. The following lemma characterizes the existence of equilibrium for any inclusion set I and shows that the unique stable equilibrium is the monotonic equilibrium with the largest number of types covered.

Lemma 5 *(i) Let $\overline{n}$ be the largest element of I. An equilibrium with inclusion set I exists if and only if*

$$\beta_n^* \geq \widehat{\beta} \Leftrightarrow m\frac{\alpha_{\overline{n}}}{\widehat{\beta}}\frac{1-F(\widehat{\widehat{x}})}{f(\widehat{\widehat{x}})} \geq k. \quad (20)$$

(ii) Given parameter values, there is a unique stable equilibrium, which is the monotonic equilibrium with the biggest z; such z^* is determined by

$$m\frac{\alpha_{z^*}}{\widehat{\beta}}\frac{1-F(\widehat{\widehat{x}})}{f(\widehat{\widehat{x}})} \geq k, \text{ but } m\frac{\alpha_{z^*+1}}{\widehat{\beta}}\frac{1-F(\widehat{\widehat{x}})}{f(\widehat{\widehat{x}})} < k.$$

The result of part (ii) of Lemma 5 is intuitive. With free entry, if some profitable consumer type ($\beta_n^* \geq \widehat{\beta}$) is excluded, then more firms can enter jointly to cover that type. Therefore, the unique stable equilibrium is the monotonic equilibrium with the biggest type coverage. Again, we call the unique stable equilibrium the market equilibrium.

Proposition 7 *(Free entry) In the market equilibrium, if either the search costs s decrease, or the targetibility of search increases (σ decreases), (i) the number of types of consumers covered, z^*, will (weakly) increase, and the concentration of sales will (weakly) decrease; (ii) for all the previously covered types $n \leq z^*$, both p_n^* and β_n^* decrease and $\overline{u}_n$ increases; (iii) the measure of active firms γ decreases if z^* remains the same, and it can either decrease or increase if z^* increases.*

Part (i) of Proposition 7 shows that decreases in search costs or increases in search targetibility again give rise to the long tail effect, as both encourage long tail consumers to search. Part (iii) shows that with free entry the measure of active firms can either decrease or increase, an effect absent from the basic model. Specifically, if the type coverage does not change, then the measure of active firms will decrease under free entry. This is because intensified search leads to lower prices and a lower gross profit, and the measure of firms serving each type has to decrease to restore the zero profit condition. Although for covered types prices decrease and utilities increase, they are partially offset by the induced decrease in the measure of firms of the right category. Another difference is that with free entry each previously covered type always benefits from a decrease in search costs or an increase in search targetibility, while in the basic model that is not the case. The main reason is that, with free entry, covering a previously excluded type has no direct effect on the measure of firms serving the already covered types, as the measure of firms serving each type is independently determined. However, with fixed measure of firms, covering a new type would reduce the measure of firms serving the already covered types, which reduces consumer utility by reducing those types' probabilities of finding the right category.

Regarding the results in Proposition 6 (the different effects of changes in search costs and those of increases in search targetibility), it is not difficult to see that they still hold qualitatively with free entry. This is because what drives the price ratios and concentration of firms is the equal (gross) profit condition, which also holds under free entry.

6.2 The noise increases with the measure of active firms

In this subsection we study the case that σ is increasing in γ. In general, the noise term σ could be written as $\sigma(\gamma)$, with $0 < \sigma'(\gamma) \leq 1$. Here we will only consider the special case in which the noise term is $\sigma\gamma$, or the noise is always of a constant proportion to the total measure of active firms. We focus on the special case for two reasons. First, it is easy to analyze. Second, the general case $\sigma'(\gamma) \in (0, 1)$ would yield similar qualitative results as the special case. Treating both $\{\beta_n\}$ and σ as fractions, for the special case the probability of finding firms of the right category can be written as:

$$\phi_n = \frac{\beta_n \gamma}{\beta_n \gamma + \sigma \gamma} = \frac{\beta_n}{\beta_n + \sigma}. \tag{21}$$

Observing (21), we see that the expression of ϕ_n is the same as that in the basic model, and it is independent of γ. Therefore, all the previous results hold (except those regarding firms' profits). Moreover, the equilibrium measure of active firms γ is (independent of other equilibrium features) determined by the free entry or zero profit condition.

It is worth noting that part (iii) of Proposition 5 holds with free entry and σ increasing with γ: when the type coverage increases, a decrease in search costs (or an increase in search targetibility) might make consumers of previously covered types worse off. This is in contrast to the case with free entry and σ independent of γ. To understand this result, observe that when σ increases with the measure of active firms, an increase in type coverage imposes a negative externality on already covered mainstream consumers, as the increased noise will reduce those consumers' probability of finding firms of the right category. On the other hand, when σ is independent of γ this externality is absent. In the basic model with γ fixed at 1, an increase in type coverage also imposes a negative externality on already covered mainstream consumers, but for a different reason: although the noise does not change, the probability of finding firms of the right category decreases since less firms remain serving the mainstream types as some firms switch to serving the newly covered types.

To study the effect of changes in search costs or search targetibility on γ, consider an increase in search targetibility. If the equilibrium type coverage z^* does not change, it is not difficult to see that the measure of active firms γ must decrease in equilibrium. This is because by part (ii) of Proposition 5, if γ remains the same then prices and firms' gross profits will decrease. Thus γ must decrease to restore the zero profit condition. If the equilibrium type coverage z^* increases, whether γ will increase or decrease is not clear. These effects are the same as those in part (iii) of Proposition 7.

7 Conclusion

This paper develops a new search model that incorporates search targetibility or quality of search. Consumers are searching for the right products within the right categories: different types of consumers demand different categories of goods, and the same type of consumers have different preference among the products of the right category. Mainstream consumers are distinguished from long tail consumers in terms of the prevalence of consumer tastes (types) in the population. We show that mainstream consumers search more within the right categories and enjoy higher utility, mainstream products are sold at lower prices, and among the covered types the distribution of firms is skewed more toward long types relative to the distribution of consumer types.

In the market equilibrium long tail consumers might be excluded. As search costs decrease or search targetibility increases, additional variety of goods catering to long tail consumers will be provided and the concentration of sales across different categories of goods decreases. This provides an explanation for the long tail effect. When the type coverage does not change, a decrease in search costs or an increase in search targetibility leads to lower profits, lower prices, and high consumer utilities for all covered types. However, when the type coverage increases, the effects of a decrease in search costs or an increase in search targetibility on prices, profits, and consumer utilities are ambiguous. Decreases in search costs and increases in search targetibility have different qualitative effects on consumers' overall search, the distribution of prices, and the distribution of firms across types

For simplicity, in the model we have assumed that each type of consumer only demands goods of the corresponding category. That is, there is no substitutability of goods across different types. In the real world, goods of different categories are more likely to be imperfect substitutes. For example, if a consumer who likes detective stories the most (a DS type) buys a science fiction (SF) book, his utility could still be potentially positive, though the utility is less than what he gets from buying a DS book. With the possibility of imperfect substitution across types, instead of being outrightly excluded, long tail types might participate in the market and buy goods that are not of their preferred category. Following the example, a DS type might buy some SF book if it is very hard to find DS books but SF books are in abundance in the market. We leave this line of extension for future research.

Appendix

Proof of Lemma 4.

Proof. First we show that it holds for $z = 1$. That is, in the T_1 pure exclusive equilibrium firms' profits are the highest among all pure exclusive equilibria. Consider the T_1 (with configuration I_1) and T_n (with configuration I_n, $n \geq 2$) pure exclusive equilibrium. From (10), we can clearly see that $p_1^{*I_1} = p_n^{*I_n}$. Now since $\alpha_1 > \alpha_n$, we have $\Pi_1^{*I_1} = m\alpha_1 p_1^{*I_1} > m\alpha_n p_n^{*I_n} = \Pi_n^{*I_n}$.

Next we show that it holds for z, $1 \leq z < N$. Consider the equilibrium of a configuration I which has z elements and is not monotonic. Let i be the smallest n such that $\beta_n = 0$. Since I is not z-monotonic, $i < z$. Let j be the largest n such that $n > i$ and $\beta_n > 0$. Now construct a new configuration I' from I as follows: move j out of I and replace it with i, without changing other elements. Essentially, under I and I' the same $z - 1$ types of consumers are served, and under I' a more mainstream type (i instead of j) is served. Note that if we repeat this process the new configuration will eventually become z-monotonic. Now, what we need to show is that firms get a higher profit in the equilibrium with configuration I' than that with configuration I.

Denote the equilibrium distribution of firm types under I and I' as $\{\beta_n\}$ and $\{\beta'_n\}$, respectively. In the next step we show that $\beta_j < \beta'_i$. Suppose the opposite, $\beta_j \geq \beta'_i$, is true. Now by Lemmas 1 and 2, we have $p^*_j \leq p^{*\prime}_i$. Given that $i < j$ so that $\alpha_i > \alpha_j$, it follows that $\Pi^*_j = m\frac{\alpha_j}{\beta_j}p^*_j < m\frac{\alpha_i}{\beta'_i}p^{*\prime}_i = \Pi^{*\prime}_i$. Since $\sum_{n=1}^{N}\beta_n = 1$ and $\sum_{n=1}^{N}\beta'_n = 1$, $\beta_j \geq \beta'_i$ implies that there must be some $k \in I$ and $k \neq j$ such that $\beta_k \leq \beta'_k$. Now following Lemmas 1 and 2, we have $p^*_k \geq p^{*\prime}_k$. Thus, $\Pi^*_k = m\frac{\alpha_k}{\beta_k}p^*_k \geq m\frac{\alpha_k}{\beta'_k}p^{*\prime}_k = \Pi^{*\prime}_k$. By the equal profit condition under both I and I', this leads to $\Pi^*_j = \Pi^*_k \geq \Pi^{*\prime}_k = \Pi^{*\prime}_i$, which contradicts the previous derived result $\Pi^*_j < \Pi^{*\prime}_i$. Therefore, we must have $\beta_j < \beta'_i$.

Now given that $\beta_j < \beta'_i$, since $\sum_{n=1}^{N}\beta_n = 1$ and $\sum_{n=1}^{N}\beta'_n = 1$, there must be some $k \in I$ and $k \neq j$ such that $\beta_k > \beta'_k$. By Lemmas 1 and 2, it follows that $p^*_k < p^{*\prime}_k$ and $\Pi^*_k < \Pi^{*\prime}_k$. Since in equilibrium all firms always get equal profit, this means that firms' equilibrium profit is higher under configuration I'. ■

Proof of Proposition 3.

Proof. Recall that, by previous analysis, we only need to worry about the deviations to some excluded types. Among all the possible deviations to a particular type that is excluded, the most profitable deviation is the one that just has a $\widehat{\beta}$ measure of firms deviating to becoming that type. This is because the profit of any type T_n firms is decreasing in β_n, while $\beta_n < \widehat{\beta}$ will lead to zero profit for T_n type firms, as type t_n consumers will not search by Lemma 3.

Part (i). Consider an equilibrium with a non-monotonic configuration I. Let $\overline{n}$ be the largest element in I, or $t_{\overline{n}}$ be the least mainstream type included. Firms' equilibrium profit is equal to type $T_{\overline{n}}$ firms' profit, which is $\Pi^{*I} = \Pi^{*I}_{\overline{n}} = m\frac{\alpha_{\overline{n}}}{\beta^I_{\overline{n}}}p^{*I}_{\overline{n}}$ Since I is not monotonic, there is some $i \notin I$ and $i < \overline{n}$, or t_i is some excluded mainstream type. Now consider the most profitable deviation to type T_i. That is, exactly a $\widehat{\beta}$ measure of firms deviating to becoming type T_i. Each deviating firm's profit is $\pi^d_i = m\frac{\alpha_i}{\widehat{\beta}}p^*_i$. Given that the original equilibrium exists, it must be the case that $\beta^I_{\overline{n}} \geq \widehat{\beta}$. By Lemma 2, $p^*_i \geq p^{*I}_{\overline{n}}$. Combining the above results with the fact that $\alpha_i > \alpha_{\overline{n}}$, we have $\pi^d_i > \Pi^{*I}_{\overline{n}} = \Pi^{*I}$. Therefore, there is a profitable (joint) deviation to an excluded type i, and the equilibrium is not stable.

Part (ii). Since, by Proposition 2, β^z_z is decreasing in z, following Lemma 3 we reach the conclusion that monotonic equilibria with more types covered are more difficult to exist. The number of types being covered in the monotonic equilibrium with the largest number of types being covered, z^*, is determined by $\beta^{z^*}_{z^*} \geq \widehat{\beta}$ and $\beta^{z^*+1}_{z^*+1} < \widehat{\beta}$.

By the result of part (i), only monotonic equilibria can be potentially stable. Given parameter values, all monotonic equilibria with z, $1 \leq z \leq z^*$, exist, and no monotonic equilibria with z, $z > z^*$, exists. We first show that any monotonic equilibrium with $1 \leq z < z^*$ is not stable. Consider the following deviation: a $\widehat{\beta}$ measure of firms deviating to becoming type T_{z^*}. Each deviating firm's profit is $\pi^d_{z^*} = m\frac{\alpha_{z^*}}{\widehat{\beta}}p^*_{z^*}$. Given that the monotonic equilibrium with z^* exists, it must be the case that $\beta^{z^*}_{z^*} \geq \widehat{\beta}$. By Lemma 2, $p^*_{z^*} \geq p^{*z^*}_{z^*}$. Therefore, $\pi^d_{z^*} \geq \Pi^{*z^*}_{z^*}$. Since firms' equilibrium profits are increasing in z, by Proposition 2, we have $\pi^d_{z^*} \geq \Pi^{*z^*}_{z^*} > \Pi^{*z}$. Thus the proposed deviation is a profitable one, and any monotonic equilibrium with $z < z^*$ is not stable.

Finally, we show that the monotonic equilibrium with z^* is stable. Consider the most profitable deviation to type $z^* + 1$: a $\widehat{\beta}$ measure of firms deviating to becoming type T_{z^*+1}. Each deviating firm's profit is $\pi^d_{z^*+1} = m\frac{\alpha_{z^*+1}}{\widehat{\beta}}p^*_{z^*+1}$. Suppose the deviation is profitable, $\pi^d_{z^*+1} > \Pi^{*z^*}$. Then the monotonic equilibrium with $z^* + 1$ would have existed. To see this, note that $\pi^d_{z^*+1} > \Pi^{*z^*}$ implies that in the $(z^* + 1)$-monotonic equilibrium, more firms will switch from other types to type $z^* + 1$ to restore the equal profit condition. This further implies that $\beta^{z^*+1}_{z^*+1} \geq \widehat{\beta}$ and the monotonic equilibrium with $z^* + 1$ exists, which contradicts the assumption that such an equilibrium does not exist. Therefore, it must be the case that $\pi^d_{z^*+1} < \Pi^{*z^*}$, or the deviation is not profitable. Given that α_n is decreasing, the most profitable deviations to type $z > (z^* + 1)$ are less profitable than that to type $z^* + 1$. Therefore, all the deviations are not profitable and the monotonic equilibrium with z^* is stable. ■

Proof of Proposition 5.

Proof. In the definition of $\widehat{\phi}$, which is given by $\widehat{\phi}g(\widehat{\widehat{x}}) = s$, $\widehat{\phi}(s)$ is increasing in s. This implies that $\widehat{\beta}(s)$ is increasing in s. Fixing s, $\widehat{\phi}$ is determined as well. Since $\widehat{\phi} = \frac{\widehat{\beta}}{\widehat{\beta}+\sigma}$, we can see that $\widehat{\beta}(\sigma)$ is increasing in σ. Therefore, both a decrease in σ and a decrease in s will lead to a decrease in $\widehat{\beta}$ and potentially more types of consumers covered in the market equilibrium. This proves part (i).

We will only present the proof of part (ii) and (iii) when σ decreases, as that of a decrease in s is similar. Suppose $\sigma' < \sigma$ but $z^{*\prime} = z^*$. We first show that for any $n \leq z^*$, $\widehat{x}'_n > \widehat{x}_n$. Suppose there is a $k \leq z^*$ such that $\widehat{x}'_k \leq \widehat{x}_k$. By Lemma 2, we have $p^{*\prime}_k \geq p^*_k$. And by Lemma 1, we have $\phi'_n \leq \phi_n$. Given that $\sigma' < \sigma$, it must be the case that $\beta'_k < \beta_k$. Therefore, $\Pi^{*\prime}_k > \Pi^*_k$. Since $z^{*\prime} = z^*$, $\beta'_k < \beta_k$ implies that there must be a $j \leq z^*$ and $j \neq k$ such that $\beta'_j > \beta_j$. Combining the above results with the fact that $\sigma' < \sigma$, we have $\phi'_j > \phi_j$ and $p^{*\prime}_j < p^*_j$ by Lemmas 1 and 2. Thus we have $\Pi^{*\prime}_j < \Pi^*_j$. By equal profit conditions, this contradicts $\Pi^{*\prime}_k > \Pi^*_k$. Therefore, we must have $\widehat{x}'_n > \widehat{x}_n$ for any $n \leq z^*$. Given this, by Lemmas 1 and 2, we immediately have $p^{*\prime}_n < p^*_n$ and $\overline{u}'_n > \overline{u}_n$. If there is some $n \leq z^*$ such that $\Pi^{*\prime}_n > \Pi^*_n$, applying similar logic as before we can derive some contradiction. Therefore, we must have $\Pi_n^{*\prime} < \Pi_n^*$.

Part (iii). We only need to provide an example. Suppose under initial σ, $\beta_{z^*+1}^{z^*+1} = \widehat{\beta} - \varepsilon$, and $\sigma' = \sigma - \eta$, with both ε and η being positive but very small. Moreover, under σ', $\beta_{z^*+1}^{z^*+1\prime}$ is slightly bigger than $\widehat{\beta}$ so that $z^{*\prime} = z^* + 1$. That is, type $z^* + 1$ is covered under σ'. Given that $\beta_{z^*+1}^{z^*+1\prime} > 0$, there must be a type $k \leq z^*$ such that $\beta_k' < \beta_k$. Since σ' is very close to σ, now by Lemmas 1 and 2, we have $\Pi_k^{*\prime} > \Pi_k^*$. Therefore, firms' profits increase. This further implies that $\beta_n' < \beta_n$, $p_{\overline{n}}^{*\prime} > p_{\overline{n}}^*$, and $\overline{u}_n' < \overline{u}_n$ for all $n \leq z^*$. ■

Proof of Proposition 6.

Proof. Let n and n' be two arbitrarily covered types in the market equilibrium, with $n' > n$. By the equal profit condition and (14), we get

$$(\frac{\beta_n}{\beta_{n'}})^2 \frac{1 + \sigma/\beta_{n'}}{1 + \sigma/\beta_n} = (\frac{\alpha_n}{\alpha_{n'}})^2. \tag{22}$$

Part (i). Inspecting (22), we see that it does not depend on search costs s. Given that z^* does not change, a decrease in s will not affect the ratio $\beta_n/\beta_{n'}$. Therefore, a decrease in s will not affect $\{\beta_n\}$ or the concentration of firms. This implies that $\{\phi_n\}$ will not change either. By (14), $p_n^*/p_{n'}^* = \sqrt{\phi_{n'}/\phi_n}$. Thus the price ratio $p_n^*/p_{n'}^*$ will not change either. Recall that the expected length of search is $1/\rho_n$, and $\rho_n = \phi_n(1 - F(\widehat{x}_n))$. Since a decrease in s will not affect ϕ_n but will cause $\widehat{x}_n$ to increase, ρ_n will decrease. Thus a decrease in s will induce more overall search for any covered type.

Suppose $\sigma' < \sigma$. We want to show $\beta_n'/\beta_{n'}' > \beta_n/\beta_{n'}$. Suppose the opposite is true, $\beta_n'/\beta_{n'}' \leq \beta_n/\beta_{n'}$. Given that z^* does not change, and $\beta_n > \beta_{n'}$ and $\beta_n' > \beta_{n'}'$, it implies that

$$\frac{1 + \sigma'/\beta_{n'}'}{1 + \sigma'/\beta_n'} < \frac{1 + \sigma/\beta_{n'}'}{1 + \sigma/\beta_n'} < \frac{1 + \sigma/\beta_{n'}}{1 + \sigma/\beta_n}. \tag{23}$$

Now the left hand side of (22) under σ' is strictly less than that under σ. This contradicts (22), by which they should equal to each other. Therefore, we must have $\beta_n'/\beta_{n'}' > \beta_n/\beta_{n'}$, or the concentration of firms increases. The change in price ratio can be expressed as

$$p_n^{*\prime}/p_{n'}^{*\prime} - p_n^*/p_{n'}^* = \sqrt{\frac{1 + \sigma'/\beta_n'}{1 + \sigma'/\beta_{n'}'}} - \sqrt{\frac{1 + \sigma/\beta_n}{1 + \sigma/\beta_{n'}}} > 0,$$

where the inequality follows (23).

Regarding the expected length of search, by (13) we have

$$\rho_n \frac{b - \widehat{x}_n}{2} = s = \rho_n' \frac{b - \widehat{x}_n'}{2}. \tag{24}$$

By the proof of Proposition **5,** $\widehat{x}'_n > \widehat{x}_n$. Now by (24), $\rho'_n > \rho_n$. Thus when σ increases, any covered type searches less overall.

Part (ii). Suppose $s' < s$ and $z^{*'} > z^*$. By previous results, $\beta'_n < \beta_n$ and $\beta'_{n'} < \beta_{n'}$. We want to show $\beta'_n/\beta'_{n'} < \beta_n/\beta_{n'}$. Suppose the opposite is true, $\beta'_n/\beta'_{n'} \geq \beta_n/\beta_{n'}$. This implies that

$$\frac{1+\sigma/\beta'_{n'}}{1+\sigma/\beta'_n} > \frac{1+\sigma/\beta_{n'}}{1+\sigma/\beta_n}.$$

Now the left hand side of (22) under s' is strictly greater than that under s. This contradicts (22), by which they should equal to each other. Therefore, we must have $\beta'_n/\beta'_{n'} < \beta_n/\beta_{n'}$, or the concentration of firms decreases. This further implies that the change in price ratio

$$p^{*\prime}_n/p^{*\prime}_{n'} - p^*_n/p^*_{n'} = \sqrt{\frac{1+\sigma/\beta'_n}{1+\sigma/\beta'_{n'}}} - \sqrt{\frac{1+\sigma/\beta_n}{1+\sigma/\beta_{n'}}} < 0.$$

■

Proof of Lemma 5.

Proof. (i). Recall from Lemma 3 that $\widehat{\widehat{x}}$ is the minimum reservation match value to induce consumers to search, and $\widehat{\beta}$ is the corresponding minimum measure of firms of the right category. For any $n \in I$, if

$$m\frac{\alpha_n}{\widehat{\beta}}\frac{1-F(\widehat{\widehat{x}})}{f(\widehat{\widehat{x}})} \geq k,$$

then there is a $\beta^*_n > \widehat{\beta}$ such that $m\frac{\alpha_n}{\beta^*_n}p^*_n = k$. This is because by previous results p^*_n is increasing in β_n, thus the gross profit Π^*_n is decreasing in β_n. Note that the above condition is the most stringent for the largest n, $\overline{n}$. Therefore, condition (20) is sufficient to ensure the equilibrium with inclusion set I exists.

(ii). First we show that any equilibrium with a non-monotonic configuration I is not stable. Let $\overline{n}$ be the largest element in I. By (i), $m\frac{\alpha_{\overline{n}}}{\widehat{\beta}}\frac{1-F(\widehat{\widehat{x}})}{f(\widehat{\widehat{x}})} \geq k$. Since I is non-monotonic, there is an $i \notin I$ and $i < \overline{n}$. The fact that $\alpha_i > \alpha_{\overline{n}}$ implies that $m\frac{\alpha_i}{\widehat{\beta}}\frac{1-F(\widehat{\widehat{x}})}{f(\widehat{\widehat{x}})} \geq k$. Now if exactly $\widehat{\beta}$ measure of new firms choose to be type T_i firms, type t_i consumers will search and those firms can earn a non-negative profit. Thus the equilibrium with configuration I is not stable.

By similar logic, any equilibrium of monotonic configuration with $z < z^*$ is stable. This is because new firms can profitably enter to serve type t_{z^*} consumers. The z^*-monotonic equilibrium is stable because no more firms can profitably enter and serve types less mainstream than type t_{z^*}. ■

Proof of Proposition 7.

Proof. (i). Recall that in Lemma 3, $\widehat{\beta}$ is defined by $\frac{\widehat{\beta}}{\widehat{\beta}+\sigma}g(\widehat{\widehat{x}}) = s$. From the expression we can easily see that both a decrease in σ and a decrease in s will lead to a decrease in $\widehat{\beta}$, and potentially more types of consumers will be covered in the market equilibrium (following Lemma **5**).

(ii). Suppose $\sigma' < \sigma$ (the proof regarding a decrease in s is similar). We want show that $\beta_n^{*\prime} < \beta_n^*$. Suppose $\beta_n^{*\prime} \geq \beta_n^*$. Combining the above statement with the fact that $\sigma' < \sigma$, by (18) we reach the conclusion that $\widehat{x}_n' > \widehat{x}_n$, which by (17) implies that $p_n^{*\prime} < p_n^*$. Now from (16) we have $\Pi_n^{*\prime} < \Pi_n^*$, a contradiction of the fact that both should equal to k. Therefore, we must have $\beta_n^{*\prime} < \beta_n^*$. Now by (16) $p_n^{*\prime} < p_n^*$, which further implies that $\widehat{x}_n' > \widehat{x}_n$ and $\overline{u}_n' > \overline{u}_n$ by (17).

(iii). If z^* remains the same, γ would decrease. This is because by (ii) $\beta_n^{*\prime} < \beta_n^*$ for all $n \leq z^*$. If z^* increases, then there are additional firms entering into serving more long tail types, and the change in the total measure of active firms γ is ambiguous. ■

References

[1] Anderson, C. "The Long Tail," *Wired*, Issue 12.10 Oct., 2004.

[2] Anderson, C. *The Long Tail: Why the Future of Business Is Selling Less of More.* 2006, New York: Hyperion.

[3] Anderson, C. *The Longer Long Tail: How Endless Choice is Creating Unlimited Demand.* 2009, London: Random House Business Books.

[4] Anderson, S. and R. Renault. "Pricing, Product Diversity and Search Costs: A Bertrand-Chamberlin-Diamond Model," *Rand Journal of Economics*, 1999, 30, pp. 719-735.

[5] Bagnoli, M. and T. Bergstrom. "Log-concave Probability and its applications," *Economic Theory*, 2005, 26, pp. 445-469.

[6] Bakos, Y. "Buyer Search Costs: Implications for Electronic Marketplaces," *Management Science*, 1997, 43(12), 1676-1692.

[7] Bar-Isaac, H., G. Caruana, and V. Cunat. "Search, Design, and Market Structure," *American economic Review*, 2011, forthcoming.

[8] Bergemann, D. and A. Bonatti. "Targeting in Advertising Markets: Implications for Offline vs. Online Media," 2011, working paper.

[9] Brynjolfsson, E., Y.J. Hu, and D. Simester. "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Products Sales," 2007, working paper.

[10] Brynjolfsson, E., Y.J. Hu, and M. D. Smith. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science*, 2003, 49, pp. 1580-1596.

[11] Brynjolfsson, E., Y.J. Hu, and M. Smith. "From Niches to Riches: Anatomy of the Long Tail," *Sloan Management Review*, 2006, 47, pp. 67-71.

[12] Diamond, P. "A Model of Price Adjustment," *Journal of Economic Theory*, 1971, 3, pp. 156-168.

[13] Elberse, A. and F. Oberholzer-Gee. "Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales," 2008, working paper.

[14] Greco, A. *The Book Publishing Industry*, 2nd Edition, 2005, Lawrence Erlbaum Associates: New Jersey.

[15] Hervas-Drane, A. "Search, Information Provision, and Sales Distribution," 2010, working paper.

[16] Stahl, D.O. "Oligopolistic Pricing with Sequential Consumer Search," *American Economic Review*, 1989, 79, pp. 700–712.

[17] Varian, H.R. "A Model of Sales," *American Economic Review*, 1980, 70, pp. 651–659.

[18] Wolinsky, A. "True Monopolistic Competition as a Result of Imperfect Information," *Quarterly Journal of Economics*, 1986, 101, pp. 493-511.