The Market for Keywords*

Kfir Eliaz† and Ran Spiegler‡

January 14, 2014

Abstract

Can a competitive market implement an ideal search engine? To address this question, we construct a two-sided market model in which consumers with limited, idiosyncratic vocabulary use keywords to search for their desired products. Firms get access to a keyword if they pay its competitive price-per-click. An underlying "broad match" function determines the probability with which a firm will enter the consumer's search pool as a function of the keyword it "buys" and the consumer's queried keyword. The main question we analyze is whether there exists a broad match function that gives rise to an efficient competitive equilibrium outcome. We provide necessary and sufficient conditions, in terms of the underlying search cost and the joint distribution over consumers' tastes and vocabulary, and characterize equilibrium keyword prices under such equilibria. The Battachayyara coefficient, a measure of closeness of probability distributions, turns out to play a key role in the analysis.

KEYWORDS: keywords

1 Introduction

The effect of search frictions on the functioning of markets has been one of the important themes of microeconomic theory in the last few decades. What economists typically mean by a search friction is the resources that an agent needs to devote in order to find what he wants. Yet another important search friction, which the literature has not addressed, is people's limited ability to *describe* their wants. Typos, misspellings ("Eliza", "Spiegel" and "Meerkat" are important real-life instances), or simply forgetting the name of an object, are perhaps the simplest manifestations of

*This paper benefitted from an ISF grant, no. 1153/13. We thank Yair Antler, Alex Frug, seminar participants at ESSET 2013, and especially Sergiu Hart for helpful comments.

†Tel-Aviv University and University of Michigan. Email: kfire@post.tau.ac.il.

‡Tel Aviv University and University College London. URL: http://www.tau.ac.il/~rani. E-mail: rani@post.tau.ac.il.

this friction. A more interesting example is being able to articulate general features of a product but not its specific characteristics - the genre of a novel, the category of an electric appliance, etc. Synonyms (or, more broadly, similar concepts, e.g. "aparthotel" vs. "serviced apartment") are yet another example. Finally, people often have only a vague idea of their wants and cannot find precise words to describe them succinctly and intelligibly.

As usual, when there is a market friction, we see the emergence of institutions that try to address it. One function of a sales person at a department store is to listen to customers' half-intelligible descriptions of what they look for and direct them to the relevant department. Reference books and Yellow Pages use a classification system and an index to simplify the search process. More modernly, online search engines attempt to bridge the gap between consumers' description of their wants and their actual wants. Consumers consult a search engine by entering a query, which is typically a collection of keywords, and the search engine responds with a list of objects (websites). A search engine is more than an address book; an effective search engine identifies typos, recognizes synonyms, and provides relevant specimens of the category described in the consumer's query. Such devices shape the search pool that consumers get when they submit a query, and ideally the search pool is optimally suited to their coarse, vague and imprecise description of what they want. We refer to the totality of such devices as "*broad matching*".

Traditionally, the broad matching function has been centralized; the search engine (like classified guidebooks) follows some algorithm for optimizing the search process, given consumers' limited ability to describe their wants. The theoretical question we pose in this paper is: *Can broad matching be decentralized and implemented by a competitive market*? In other words, can a competitive market for keywords (or queries, more generally) mimic the optimal algorithm of an ideal search engine, such that the incentives of firms will determine the search pools that consumers get in response to their queries? Of course, real-life online search engines partly decentralize the allocation of websites to queries, by auctioning so-called "sponsored links". However, sponsored links currently coexist with so-called "organic search", namely the centralized non-market search algorithm. Our question can thus be restated: if organic search is abolished such that all we are left with is the sponsored links, will the optimal quality of search be affected?

To address this question, we construct a simple model of a two-sided market in which search is conducted via keyword queries. There is a finite set of products X and a finite set of words W, where $|W| \geq |X|$. Each consumer type is defined by a pair

(x, w), where x is the only product he wants and w is the only word he can articulate. The population of consumers is characterized by a distribution μ over consumer types. On the other side of the market, each product x is offered by a measure one of "x firms". If a consumer who wants x is matched with a y firm, each party gets a payoff of 1 (0) if $x = y$ ($x \neq y$). Consumers can find their desired product only by conducting rational sequential search from a pool of firms associated with the word they can articulate, with a search cost of s per random draw. An *ideal search engine* is a collection of such search pools that maximizes total consumer welfare.

In contrast, a market-based search engine allocates words to firm types through a system of "prices per click" for each word. If a firm "buys" a word v, it gets access to the search pool associated with the queried word w with probability $b(w|v)$. The firm can thus calculate the effective "conversion rate" associated with v, namely the probability that when a consumer samples ("clicks") the firm through this process, he will transact with it. The function b captures the "broad match" technology of the market for keywords; it is an exogenous aspect of the search engine, which is designed by a central planner. It can be viewed as a "network of platforms", where each word is a platform, and $b(w|v)$ is the probability that firms with access to the platform v are brought into contact with consumers with access to the platform w. The important difference between this "broad match" method and the one followed by centralized "organic search" is that the search engine does not directly identify the firm types; its inputs are the words themselves, not the objects that are associated with them. The search engine relies on the decentralized competitive market incentives that induce an appropriate allocation of words to firms.

To see why broad match is important for a market-based search engine, suppose that $s = 0$, and imagine that the search engine adopts a "narrow match" method, i.e. $b(w|v) = 1$ (0) if $w = v$ ($w \neq v$). In this case, the search engine indeed functions as a mere address book. For any word w, competitive forces will favor firms that offer $\arg\max_x \mu(x, w)$, namely the products with the mass appeal within the group of consumers who query w. The reason is that these firms will enjoy the highest "conversion rate" and therefore have the highest willingness to pay for the word. The keyword's equilibrium price will be driven up to this highest conversion rate, crowding out firms that could potentially serve the "long tail" of the consumer population associated with w.

Going to the other extreme by adopting an indiscriminate broad match method (i.e. $b(w|v) = 1$ for all w, v) would only make things worse. Such a method reduces the two-sided market into a single platform consisting of one "mega-word", thus obliterating any correlation between consumers' queries and their wants. Competitive pressures will now favor firms that offer $\arg\max_x \alpha(x) = \Sigma_w \mu(x, w)$, the most popular product throughout the entire consumer population. The objective is thus to design an intermediate broad match function that will address the underlying search frictions without exacerbating the "long tail" problem.

For any broad match function b, we define a competitive equilibrium as an allocation of words to firms and a search strategy for consumers, such that (i) consumer behavior is optimal given the search pools induced by the allocation, and (ii) each word is allocated to the firms with the highest willingness to pay - namely, the highest conversion rate from the word. The equilibrium price-per-click of the word is equal to that highest conversion rate. We impose the following strictness requirements that avoid reliance on ties. First, consumers search only when this is strictly beneficial to them. Second, the conversion rate from each word is maximized by exactly one type of firms, such that the allocation of words to firm types is given by a well-defined function $f : W \to X$. The primary motivation for the latter requirement is to get equilibria that are robust to small perturbations in the consumer type distribution.

Our main results provide necessary and sufficient conditions for the implementability of the ideal search engine by competitive equilibrium in a market-based search engine. Specifically, we ask whether there exist a broad match function b and a competitive equilibrium in the market for keywords induced by b that generates the total consumer welfare induced by the ideal search engine. The answer turns out to depend on whether search costs are strictly positive.

When $s = 0$, an optimal search pool associated with w should give positive representation, however small, to any firm type x for which $\mu(x, w) > 0$. We show that in this case, market implementability of an ideal search engine is possible if and only if the consumer type distribution has the property that each product x has a distinct conditional distribution of queries, denoted β_x. When the condition for implementability is met, we construct a particular broad match function that has certain optimality properties, and induces equilibrium keyword prices with a striking structure:

$$p^*(w) = \frac{1}{\sum_y BC(\beta_{f(w)}, \beta_y)}$$

where $BC(\beta_x, \beta_y) = \sum_w \sqrt{\beta_x(w)\beta_y(w)}$ is the *Battachayyara coefficient* of the conditional query distributions β_x and β_y. The Battachayyara coefficient is a standard measure of closeness between two probability distributions, with applications to pattern recognition (see Theodoridis and Koutroumbas (2008)). In our context, a higher

Battachayyara coefficient for β_x and β_y means that statistically, consumers who want x and y submit more similar queries. Thus, p^* captures in a succinct way the intuition that queries that describe more accurately what the consumer wants will have higher market prices.

When $s > 0$, the composition of search pools under the ideal search engine reflects an attempt to minimize search costs. In particular, it may be optimal to neglect very small taste niches, because of the negative search externality that addressing them would inflict on the other consumer types. We first provide a simple characterization of the ideal search engine. We then show that it is implementable by a competitive equilibrium in a market-based search engine if and only if the following inequality holds for every pair of products x, y :

$$\sqrt{\frac{\alpha(x)}{\alpha(y)}} \cdot BC(\beta_x, \beta_y) < 1$$

That is, as long as the conditional query distributions that characterize x and y are sufficiently different relative to their relative popularity, the ideal search engine can be mimicked by the market-based search engine. The induced equilibrium prices are a variant on those obtained for $s = 0$. When all products are equally popular, the characterizations of the $s > 0$ and $s = 0$ cases coincide. Thus, a chief lesson from our exercise is that conditional query distributions are crucial for understanding both the limits to market implementability of ideal search engines and the equilibrium prices that emerge in "markets for keywords".

Related literature

Our paper is related to the literature on two-sided markets (see Spiegler (2000), Rochet and Tirole (2003), Caillaud and Jullien (2001,2003) and Armstrong (2006)), which analyzes interaction between platforms of different kinds with two or more sides of a market. Some works within this tradition (e.g. Hagiu and Jullien (2011)) explicitly address search platforms. Like much of this literature, we assume single-homing on the consumers' side and multi-homing on the firms' side. The key novelty in the present paper in relation to this literature is the introduction of "broad matching", or "platform networks". All the papers we are aware of implicitly assume "narrow matching"; multiple platforms are considered only in the context of competition among platforms, and interaction between a consumer and a firm invariably requires that they are both attached to the same platform.

The "platform network" aspect of our model relates it to the literature on buyer-seller networks. In this literature (e.g., see the pioneering work of Kranton and Minehart (2001)), buyers and sellers can only trade with partners that are linked to them through the network. The question that is typically studied is what are the efficient networks and what type of networks would form if links were strategic decisions in a non-cooperative game.

Another closely related strand involved models of keyword pricing. This literature (e.g. Edelman, Ostrovsky and Schwarz (2007)) mostly focuses on the mechanism-design problem of auctioning multiple "sponsored links". Typically, the links are assumed to have different values to the bidders, and which implicitly capture some underlying consumer search process. Few papers in this tradition (see Athey and Ellison (2011)) explicitly incorporate consumer search. Chen and He* (2011) and Eliaz and Spiegler (2011) also model explicitly the interaction between keyword and product prices. Again, this literature invariably assumes "narrow matching", usually in a trivial manner by focusing on a single keyword. Another important difference is that we assume a competitive environment with many firms of each type, whereas most of the literature on search engine pricing assumes small numbers of firms, so that the auction-theoretic dimension is not redundant (Eliaz and Spiegler (2011) is an exception in this regard).

Of course, our model attenuates important dimensions in both strands (such as product pricing in the literature on two-sided markets, or asymmetric information in the literature on keyword auctions), in order to focus on the novel element, namely "broad matching". We wish to emphasize that we do not approach the problem from a mechanism-design perspective. Rather, we are interested in the question of whether a competitive market for keywords can mimic an "ideal search engine" in a two-sided market, in analogy to the textbook question of whether a competitive market for goods can implement an efficient allocation in an exchange economy. We leave mechanism-design questions to future research.

Finally, in the last decade there has been much writing, both academic and popular, about the "long tail" phenomenon (see Brynjolfsson et al. (2006) or Anderson (2007)), namely the fact that tastes for many kinds of products are highly differentiated, such that a large segment of the consumer population belongs to a large number of small taste niches, and the observation that online commerce facilitates the flourishing of firms that serve the "long tail" because it greatly shrinks the frictions that characterize brick-and-mortar commerce (e.g. storage costs). The key friction that remains (and possibly gets magnified) in such environments seems to be consumers' limited awareness of the existence of products that cater to their particular tastes, and consumers' limited ability to describe their tastes in order to locate such products on the internet. The magnitude of the "long tail" phenomenon means that the welfare implications of well-designed broad matching are large.

2 Preliminaries: Search Pools

Let X be a finite set of products. Denote $|X| = n \geq 2$. There is a population of consumers, and each consumer wants a single product $x \in X$. We refer to such consumers as x consumers. There is also a population of firms, and each firm sells a single product $x \in X$. We refer to such firms as x firms.

A search pool is a pair (d, m), where $d : X \to [0, \infty)$ is a function that describes the measure of x consumers in the pool, and $m : X \to [0, \infty)$ is a function that describes the measure of x firms in the pool. Thus, the fraction of x firms in the pool is

$$\lambda(x) = \frac{m(x)}{\sum_y m(y)}$$

An x consumer in the pool repeatedly draws at random firms from the pool (with replacement), at a constant search cost per draw $s \in [0, 1)$. We refer to such a draw as a "click". Once the consumer draws an x firm, he transacts with the firm and stops the search process.

Thus, we can define the expected number of transactions that an x firm obtains in the search pool as follows:

$$\pi(x) = \frac{d(x)}{m(x)}$$

whenever $m(x) > 0$. When $m(x) = d(x) = 0$, we write $\pi(x) = 0$. Our concept of market equilibrium in the next section will effectively rule out the case of $d(x) > 0$ and $m(x) = 0$.

The expected number of clicks that any firm obtains in the search pool is

$$c = \frac{\sum_x \frac{d(x)}{\lambda(x)}}{\sum_y m(y)} = \sum_x \pi(x)$$

The reason is as follows. Since the stopping probability of an x consumer is $\lambda(x)$, the expected number of clicks from such a consumer is $1/\lambda(x)$. The total number of clicks by consumers is thus $\sum_x d(x)/\lambda(x)$. These clicks are uniformly distributed across all firms in the pool, hence each firm gets a fraction $1/\sum_y m(y)$ of the total number of clicks.

3 The Market for Keywords

Let W be a finite set of words, where $|W| \geq n$ (we use the terms "word" and "keyword" interchangeably). A consumer type is defined by the pair (x, w), where x is the (only) product he wants and w is the (only) word he can use to express his wants. When discussing a consumer type (x, w), we sometimes refer to w as his vocabulary. Let $\mu \in \Delta(X \times W)$ be the distribution of consumer types in the population. Denote its support by T_μ. We denote

$$\begin{aligned} \alpha(x) &= \sum_{w \in W} \mu(x, w) \\ \beta_x(w) &= \frac{\mu(x, w)}{\alpha(x)} \end{aligned}$$

We assume $\alpha(x) > 0$ for every $x \in X$. In addition, for every $w \in W$, there exists $x \in X$ such that $\beta_x(w) > 0$. That is, every product is wanted by some consumers, and every word constitutes the vocabulary of some consumers. We denote the conditional probability distribution $(\beta_x(w))_{w \in W}$ by β_x.

For any pair of products $x, y \in X$ let $BC(\beta_x, \beta_y)$ be the *Bhattacharyya coefficient* between the conditional distributions β_x and β_y :

$$BC(\beta_x, \beta_y) \equiv \sum_{w \in W} \sqrt{\beta_x(w)\beta_y(w)}$$

Technically, this is simply the direction cosine between two unit vectors in $\mathbb{R}^{|W|}$, $(\sqrt{\beta_x(w)})_{w \in W}$ and $(\sqrt{\beta_y(w)})_{w \in W}$. In particular, $BC(\beta_x, \beta_y) \leq 1$, where the inequality is strict whenever $\beta_x \neq \beta_y$. The average Bhattacharyya coefficient between β_y and every β_x is denoted $\overline{BC}(y)$ and defined as follows:

$$\overline{BC}(y) = \frac{1}{n} \sum_{x \in X} BC(\beta_x, \beta_y)$$

On the other side of the market, there is a measure one of x firms, for every $x \in X$. Let $b : W \times W \to [0, 1]$ be a directed random graph over words, to which we refer as the "*broad match function*". The following notation will be convenient: $b(w|v)$ is the probability of a link from v to w. (To avoid misunderstandings, we do *not* require $\Sigma_w b(w|v) = 1$.)

We view the consumer type distribution μ, the broad match function b and the search cost s as the primitives of the market. When $b(w|w) = 1$ and $b(w|v) = 0$ for all

$w \neq v$, we refer to b as the *narrow match* function. When $b(w|v) = 1$ for all w, v, we refer to b as the *fully broad match* function.

Let $f : W \rightarrow X$ be an allocation of words to products (or firm types), and denote $k_f(x) = |f^{-1}(x)|$. Let $a : X \times W \rightarrow \{0, 1\}$ be a function that indicates the decision of each consumer type whether to engage in active search. The pair (f, a) induces a collection of search pools $(d_a(w), m_f(w))_{w \in W}$, defined by

$$\begin{aligned} d_a(x, w) &= \mu(x, w)a(x, w) \\ m_f(x, w) &= \sum_{v \in f^{-1}(x)} b(w|v) \end{aligned}$$

Using the notation from Section 2, $\lambda_f(x, w)$ denotes the induced fraction of x firms in the pool w; $\pi_{f,a}(x, w)$ denotes the expected number of transactions that a single x firm obtains in the pool w; and $c_{f,a}(w)$ is the expected number of clicks that a single firm obtains in the pool w.

The interpretation is as follows. A search pool is defined by a word w. The population of consumers in the pool consists of those consumers who can express w and chose to search. The population of firms is defined by the allocation of words to products and the broad match function. Conditional on "buying" the word v, a firm enters the pool w with probability $b(w|v)$. Thus, the population of firm types in the pool w consists of all firm types that were allocated some word, weighted by the broad match function.

We define the conversion rate of firm type x from the word v, induced by (f, a), as follows:

$$CR_{f,a}(x, v) = \frac{\sum_w b(w|v)\pi_{f,a}(x, w)}{\sum_w b(w|v)c_{f,a}(w)}$$

To understand this expression note that when an x-firm is allocated the word v, it potentially enters multiple pools w, mediated by the broad match function. For each such pool, we can calculate the number of clicks and the number of transactions the firm can expect. The conversion rate is ratio between the total number of transactions and the total number of clicks the firm, aggregated over all the search pools.

Note that $CR_{f,a}(x, v)$ is sensitive to the representation of other firm types in the search pools x firms get access to by "buying" the word v. To see why, let $X = \{x, y\}$, and consider a search pool consisting predominantly of y firms, where both consumer taste types are present. If an x firm enters the pool, it encounters consumers at any round of search: first clicks, second clicks, etc. However, since y consumers can easily find their desired product, they will typically end their search after the first round. As a result, the majority of consumers in advanced search rounds like x. Since the

x firm's overall conversion rate is determined by a weighted average across all rounds of search, its conversion rate will be relatively high. Thus, when a firm type becomes more scarce in the search pools it gets access to, it enjoys a higher conversion rate. This observation will be important for our analysis.

We are now ready for the main definition in this paper.

Definition 1 *The pair (f, a) is a* ***market equilibrium*** *if the following conditions hold:*
(i) For every (x, w) with $\mu(x, w) > 0$, $a(x, w) = 1$ if and only if $\lambda_f(x, w) > s$.
(ii) For every (x, w) with $x \neq f(w)$, $CR_{f,a}(f(w), w) > CR_{f,a}(x, w)$.

Condition (i) captures the rationality of consumer search: each consumer engages in active search if and only if the probability he finds his desired product in the pool he has access to exceeds the search cost. Condition (ii) is a "market clearing" property: each word w is allocated to the firm types that value it the most, namely those with the highest conversion rate from the word. The reason we impose a strict inequality is that we want the equilibrium allocation to be stable w.r.t small fluctuations in μ.

Definition 2 *Given a market equilibrium (f, a), the* ***equilibrium price-per-click*** *of each word $w \in W$ is $p_{f,a}(w) = CR_{f,a}(f(w))$.*

This captures the idea that competitive forces push the price-per-click of each word up until it hits the maximal willingness to pay for it. Our definition of market equilibrium views the search engine as a quasi-Walrasian auctioneer who suggests an allocation f and a system of prices $(p(w))_{w \in W}$, with the requirement that the market for keywords clears. The broad match function b can be viewed as being analogous to initial endowments in an exchange economy.

From now on, we will refer to $p_{f,a}(w)$ as simply the price of w under (f, a). Note that since the denominator in the definition of $CR_{f,a}(x, w)$ is only a function of w, $CR_{f,a}(x, w) > CR_{f,a}(y, w)$ if and only if $\sum_w b(w|v)\pi_{f,a}(x, w) > \sum_w b(w|v)\pi_{f,a}(y, w)$, i.e. if v generates a greater expected number of transactions for x than for y. Nevertheless, we prefer to express condition (ii) in Definition 1 in terms of conversion rates, because this is the relevant quantity for the price-per-click of words.

Our definition of equilibrium prices may be motivated as follows. Suppose consumers search online and each keyword is associated with a single line (sponsored link) on the consumer's computer screen. If a consumer enters a keyword and the firm that comes up does not sell the product he likes, the consumer "refreshes" the screen and the search engine draws at random another firm from the search pool given by the broad match function. When firms bid for a word v, they essentially bid for the right to appear as a search outcome in the search pools to which v gives access (via the broad match function). Thus, as long as there are at least two firms of each type, there would be a tie (among firms of the same type) for winning the keyword, such that the price would equal the highest conversion rate.

In the sequel, we will be interested in welfare properties of market equilibria. For a given (f, a) - not necessarily an equilibrium - define the social welfare as follows:

$$U(f,a) = \sum_{w} \sum_{x} \mu(x,w) u_{f,a}(x,w)$$

where

$$u_{f,a}(x,w) = \begin{cases} 0 & if \quad a(x,w) = 0 \\ -\infty & if \quad a(x,w) = 1 \text{ and } m_f(x,w) = 0 \\ 1 - \frac{s}{\lambda_f(x,w)} & if \quad a(x,w) = 1 \text{ and } m_f(x,w) > 0 \end{cases}$$

is the net utility of consumer type (x, w) under (f, a). We will say that a market equilibrium is efficient if it maximizes social welfare.

The search engine in our model is presented as a benevolent designer interested in maximizing welfare. In this respect we are not taking a standard approach in the literature on search engine pricing, which derives the pricing method that maximizes the search engine's profits. However, these two approaches are not in conflict in our framework. Because equilibrium prices are equal to firms' conversion rates, the search engine extracts the entire surplus of firms. Hence, a profit maximizing search engine would also want to maximize social welfare.

Our framework may be interpreted as capturing an economy consisting of a *network of platforms*. A "platform" is essentially an exchange where buyers meet sellers who are linked to that exchange through the network. In our model each keyword constitutes an exchange which takes the form of a search pool. Each consumer in the network economy is characterized by two primitives: the product he wants and the platform he has access to. Each consumer can only meet the sellers that are linked to his platform. Given a broad match function, a market equilibrium induces the distribution of sellers in each platform.

4 Examples

In this section we provide two examples that illustrate search frictions that are captured by various consumer type distributions, and their welfare implications in market equilibrium under various broad match functions. Throughout this section, we assume $s = 0$. This means that social welfare is measured by the fraction of consumers who find the product they want.

4.1 Specificity of Consumers' Vocabulary

This example captures the idea that consumers are sometimes unable to describe the exact product they like and can only communicate the product category. Let $X = \{x, y\}$, $W = \{w_x, w_y, \phi\}$. The first two words describe the specific products x and y, whereas the third word is a generic word that describes the product category. The consumer type distribution is

$$\begin{aligned}
\mu(x, w_x) &= \alpha\beta \\
\mu(y, w_y) &= (1-\alpha)\beta \\
\mu(x, \phi) &= \alpha(1-\beta) \\
\mu(y, \phi) &= (1-\alpha)(1-\beta)
\end{aligned}$$

where $\alpha > \frac{1}{2}$. The interpretation is that a fraction α $(1-\alpha)$ of consumers like x (y); and independently, a fraction β can express exactly the name of their favorite product, while the others can only utter the general category. We examine three alternative broad match functions, and in each case look for market equilibria that maximize social welfare.

Narrow match
Here the population of firm types in the search pool w consists only of $f(w)$. Therefore, the equilibrium that maximizes social welfare is the following. The allocation of firm types to words is $f(w_x) = f(\phi) = x$, $f(w_y) = y$. All consumer types except (y, ϕ) choose to engage in active search. All words have an equilibrium price of 1. Social welfare is $\alpha + \beta(1-\alpha)$. That is, consumers who want the popular product are served, independently of their vocabulary, whereas consumers who want the less popular product are served only if they can express its exact name.

Fully broad match

Here all firm types that were allocated some word are present in all search pools. The equilibrium that maximizes social welfare is the following. All words are allocated to the product x, and only consumer types who want x engage in active search. All words have an equilibrium price of 1, and social welfare is α. That is, only consumers who want the popular product are served.

Optimal broad match

We now construct a broad match function that implements the social optimum in market equilibrium. Assume $b(w|w) = 1$ for all w. In addition, for both $z = x, y$, let $b(w_z|w) = 0$ if $w \neq w_z$. Thus, when a consumer submits a keyword that describes a specific product, his search pool will purely consist of the firm types that were allocated that very same keyword. Finally, set $b(\phi|w_x) = 0$ - that is, a firm bidding for w_x will not enter the search pool of the generic word. The only parameter left to be determined is thus $b(\phi|w_y)$, namely the probability that a firm type enters the search pool of the generic word ϕ conditional on being allocated the specific word w_y.

We guess an equilibrium: $f(w_x) = f(\phi) = x$, $f(w_y) = y$; and $a(z, w) = 1$ for every $(z, w) \in T_\mu$. If $b(\phi|w_y) > 0$, this ensures that $\lambda_f(z, w) > 0$ for every $(z, w) \in T_\mu$. Since $s = 0$, this would both maximize social welfare and satisfy condition (i) in Definition 1. Let us turn to condition (ii). First, consider the allocation of ϕ. Since $b(w_x|\phi) = b(w_y|\phi) = 0$, the value of ϕ for a firm stems only from the consumers whose vocabulary is ϕ. Therefore, the conversion rate for x (y) from this word is α $(1 - \alpha)$. Since $\alpha > \frac{1}{2}$, allocating ϕ to x firms is consistent with condition (ii), and $p_{f,a}(\phi) = \alpha$. Likewise, since $b(w|w_x) = 0$ for all $w \neq w_x$, the value of w_x for a firm stems only from the consumers whose vocabulary is w_x. Since *all* of these consumers want x, allocating w_x to x firms is consistent with condition (ii), and $p_{f,a}(w_x) = 1$.

Checking condition (ii) for w_y is more complicated. In order to satisfy condition (i), we had to set $b(\phi|w_y) > 0$. However, this means that the value of w_y for a firm also stems from consumers who submit the generic word. The following inequality must hold in order for w_y to be allocated to y firms:

$$b(w_y|w_y)\cdot\pi_{f,a}(y, w_y) + b(\phi|w_y)\cdot\pi_{f,a}(y, \phi) > b(w_y|w_y)\cdot\pi_{f,a}(x, w_y) + b(\phi|w_y)\cdot\pi_{f,a}(x, \phi) \quad (1)$$

Let us apply the definition of $\pi_{f,a}$ and our guess of f to this inequality. The payoff of firm y from word w_y, $\pi_{f,a}(y, w_y)$, is equal to $\mu(y, w_y)$, the measure of consumers of type (y, w_y), divided by the measure of y-firms in the search pool associated with w_y. The latter is equal to the sum of probabilities that a y-firm will get access to the search pool of w_y given each of the words that it buys. Since according to f firm y buys only w_y and $b(w_y|w_y) = 1$, there is a measure one of y firms in the w_y pool. It follows that $\pi_{f,a}(y, w_y) = \beta(1-\alpha)$. In a similar fashion we can compute the payoff of y firms from the word ϕ and the payoff of x firms from the words ϕ and w_y. Applying the specification of b described above yields:

$$\begin{aligned}
\pi_{f,a}(y,\phi) &= \frac{\mu(y,\phi)}{b(\phi|w_y)} = \frac{(1-\alpha)(1-\beta)}{b(\phi|w_y)} \\
\pi_{f,a}(x,w_y) &= \frac{\mu(x,w_y)}{b(w_y|w_x)+b(w_y|\phi)} = 0 \\
\pi_{f,a}(x,\phi) &= \frac{\mu(x,\phi)}{b(\phi|w_x)+b(\phi|\phi)} = \alpha(1-\beta)
\end{aligned}$$

It follows that inequality (1) reduces to

$$b(\phi|w_y) < \frac{(1-\alpha)}{\alpha(1-\beta)}$$

Thus, any $b(\phi|w_y) \in (0, \frac{1-\alpha}{\alpha(1-\beta)})$ will successfully complete an optimal broad match function.

To see why $b(\phi|w_y)$ cannot be too large, suppose that β is close to zero and that α is close to 1 - that is, the vast majority of consumers want x but do not know its name. When y firms are allocated w_y, they enter two search pools, associated with w_y and ϕ. Their conversion rate from the former pool is 1, but their conversion rate from the latter pool tends to be low because the vast majority of consumers who submit ϕ want x. Moreover, there are a lot more consumers in the ϕ pool than in the w_y pool, hence the effective conversion rate is dominated by the former. This gives x firms an incentive to outbid y firms for w_y. In order to mitigate this incentive, we need to raise the effective conversion rate from w_y for y firms, by making them extremely scarce in the ϕ pool. This is achieved by a sufficiently low $b(\phi|w_y)$.

4.2 Misinformation

Perhaps the simplest example of a gap between the desired product and the keyword used to describe it, is when the consumer is misinformed about the product's name. In this case, broad match can be viewed as a partial substitute for correcting misinformation. The following example demonstrates the incentive issues that arise when this function is decentralized in a competitive market for keywords. Let $X = W = \{x, y\}$.

Assume

$$\begin{aligned}
\mu(x,x) &= \alpha(1-\varepsilon) \\
\mu(x,y) &= \alpha\varepsilon \\
\mu(y,y) &= (1-\alpha)(1-\varepsilon) \\
\mu(y,x) &= (1-\alpha)\varepsilon
\end{aligned}$$

where $\alpha > \frac{1}{2}$ and $\varepsilon < \frac{1}{2}$. The story is that the names of x and y are very similar and thus easily confused with one another. Thus, α is the fraction of consumers who want product x, and ε is the (independent) probability that consumers are misinformed.

Note that the "rational expectations" aspect of condition (i) in the definition of market equilibrium means that the consumer's mistake cannot be interpreted as an accidental typo: the consumer type (x,y) genuinely believes (with probability one) that the name of product x is y. Therefore, he does not update his beliefs and learns the correct name after taking a number of unsuccessful draws from his search pool.

Narrow match

The equilibrium that maximizes social welfare is the following: for every $z \in X$, $f(z) = z$ and $a(z,z) = 1$, $a(z,-z) = 0$. That is, only consumers who know the correct name of their desired product engage in active search, while the others give up on search. Both words have an equilibrium price of 1. Social welfare is $1-\varepsilon$. That is, only well-informed consumers are served.

Fully broad match

This means abandoning all sensitivity to the spelling of the queried keyword. In this case, the equilibrium that maximizes social welfare is the following. All words are allocated to the product x, and only consumer types who want x engage in active search. All words have an equilibrium price of 1, and social welfare is α. That is, only consumers who want the popular product x are served.

Optimal broad match

The above extreme broad match functions illustrate the basic tension that our model captures. On one hand, narrow match means that misinformed consumers are not served. On the other hand, trying to resolve this market failure by fully broad match causes another market failure, whereby the product with mass appeal takes over the entire market and crowds out the "niche" product.

Let us construct a broad match function that induces an efficient market equilibrium, where $f(z) = z$ for both $z = x, y$, and all consumer types engage in active search. Since $s = 0$, condition (i) in the definition of market equilibrium is satisfied. Let us turn to condition (ii). The following inequalities need to hold:

$$\begin{aligned} b(x|x) \cdot \pi_{f,a}(x,x) + b(y|x) \cdot \pi_{f,a}(x,y) &> b(x|x) \cdot \pi_{f,a}(y,x) + b(y|x) \cdot \pi_{f,a}(y,y) \\ b(y|y) \cdot \pi_{f,a}(y,y) + b(x|y) \cdot \pi_{f,a}(y,x) &> b(y|y) \cdot \pi_{f,a}(x,y) + b(x|y) \cdot \pi_{f,a}(x,x) \end{aligned}$$

Plugging the definition of π, we obtain

$$\begin{aligned} \alpha(1-\varepsilon) + \alpha\varepsilon &> \frac{b(x|x)(1-\alpha)\varepsilon}{b(x|y)} + \frac{b(y|x)(1-\alpha)(1-\varepsilon)}{b(y|y)} \\ (1-\alpha)(1-\varepsilon) + (1-\alpha)\varepsilon &> \frac{b(y|y)\alpha\varepsilon}{b(y|x)} + \frac{b(x|y)\alpha(1-\varepsilon)}{b(x|x)} \end{aligned}$$

Thus, we need to set two parameters:

$$\begin{aligned} r_1 &= \frac{b(x|x)(1-\alpha)}{b(x|y)\alpha} \\ r_2 &= \frac{b(y|x)(1-\alpha)}{b(y|y)\alpha} \end{aligned}$$

that satisfy the inequalities

$$\begin{aligned} 1 &> r_1\varepsilon + r_2(1-\varepsilon) \\ 1 &> \frac{1}{r_1}(1-\varepsilon) + \frac{1}{r_2}\varepsilon \end{aligned}$$

Now, if we set

$$r_1 = \frac{1}{r_2} = \sqrt{\frac{1-\varepsilon}{\varepsilon}}$$

both inequalities reduce to $\varepsilon(1-\varepsilon) < \frac{1}{4}$, which necessarily holds. This construction will be generalized in the next section.

5 Analysis

Let us first examine the limitations of the market for keywords under the two extreme broad match functions: narrow match and fully broad match.

Proposition 1 (Optimal equilibrium under narrow match) *Let* $b(w|w) = 1$ *and* $b(w|v) = 0$ *for all* $w \neq v$*. Then, for generic* μ*, the maximal social welfare* U *that can be sustained in market equilibrium is*

$$(1-s)\sum_{w\in W}\left(\max_{x\in X}\mu(x,w)\right)$$

Proof. First, we construct a market equilibrium (f,a) that implements this level of social welfare. Let $f(w) = \arg\max_{x\in X}\mu(x,w)$. For generic μ, this is a well-defined function. Let $a(x,w) = 1$ if and only if $x = f(w)$. It is easy to see that both conditions of the definition of market equilibrium are satisfied. Moreover, consumers who engage in active search find their desired product in their first round of search, hence their payoff is $1-s$.

Now suppose there is another equilibrium (f,a). For each word w, let $X(w)$ be the set of products for which $a(x,w) = 1$. Then, the firm type with the highest conversion rate from w is $\arg\max_{x\in X(w)}\mu(x,w)$, hence this is also $f(w)$, which is well-defined for generic μ. But this means that condition (i) in the definition of market equilibrium is satisfied only if $X(w)$ is a singleton, denoted $x(w)$, such that social welfare is $(1-s)\Sigma_w\mu(x(w),w)$, which cannot exceed the above level by definition. ■

Thus, under narrow match, it is impossible to do better than serving the most popular product at each word, and this leaves out consumers who want the less popular products for a given word. In the two examples of the previous section, this meant that consumers who prefer the niche product but only know the name of the product category, or consumers who are misinformed about the name of their desired product, are not served. Note that when $X \subseteq W$ and $\beta_x(x) = 1$ for all x - i.e. when consumers always know the name of the product they want - narrow match is optimal in the sense that it induces an efficient market equilibrium.

Proposition 2 (Optimal equilibrium under fully broad match) *Let* $b(w|v) = 1$ *for all* w, v*. Then, for generic* μ*, the maximal social welfare* U *that can be sustained in market equilibrium is* $(1-s)\max_{x\in X}\alpha(x)$*.*

Proof. Under fully broad match, the model becomes equivalent to a specification where W consists of a single word w, and $\mu(x,w) = \alpha(x)$, and the broad match function is a narrow match function. As we saw in the previous result, the maximal social welfare that can be sustained in market equilibrium is $(1-s)\max_{x\in X}\alpha(x)$. ■

Interestingly, the maximal social welfare that is implementable in market equilibrium is weakly lower under fully broad match than under narrow match. Thus, if we think of broad match as a way of addressing the neglect of "long tails" induced by narrow match, going all the way to fully broad match throws the baby with the bathwater, because it leads to a wholesale neglect of the long tail. Whereas narrow match selects the set of word-specific popular products, fully broad match leads to the crowding out of all but the most popular product among the entire population of consumers.

In what follows, we examine the conditions for the existence of a broad match function that lies between the above extremes, that will allow for efficient market equilibrium. It turns out that we need to distinguish between the case of $s = 0$ and the case of $s > 0$.

5.1 Efficient Equilibrium under $s = 0$

The maximal social welfare when $s = 0$ is 1, namely every consumer type ends up getting the product he wants (the expected duration of his search does not enter his utility because $s = 0$). That is, (f, a) must satisfy $\lambda_f(x, w) > 0$ and $a(x, w) = 1$ whenever $\mu(x, w) > 0$.

Proposition 3 *There exists a broad match function b^* that induces an efficient market equilibrium if and only if $\beta_x \neq \beta_y$ for every distinct x, y. In particular, we can choose b^* to be*

$$b^*(w, v) = \frac{\alpha(f(v))\sqrt{\beta_{f(v)}(w)}}{k_f(f(v))} \tag{2}$$

where f is an arbitrary onto function.

Proof. First, observe that any (f, a) that maximizes social welfare automatically satisfies condition (i) in Definition 1. In particular, note that f is onto, since $\alpha(x) > 0$ for every x. The question is whether there exist such (f, a) that will also satisfy condition (ii).

Recall that if an individual x firm "buys" the word v, its number of transactions is

$$\sum_w \frac{b(w|v)\mu(x, w)a(x, w)}{\sum_{v' \in f^{-1}(x)} b(w|v')}$$

Since $a(x,w) = 1$ whenever $\mu(x,w) > 0$, we can rewrite condition (ii) as follows. For every $v \in W$ and every $y \neq f(v) = x$,

$$\sum_w \frac{b(w|v)\mu(x,w)}{\sum_{v' \in f^{-1}(x)} b(w|v')} > \sum_w \frac{b(w|v)\mu(y,w)}{\sum_{v' \in f^{-1}(y)} b(w|v')} \tag{3}$$

First, assume that $\beta_x = \beta_y$ for some distinct x, y. Inequality (3) holds for every $v \in f^{-1}(x)$. Add all the inequalities across all these v:

$$\sum_{v \in f^{-1}(x)} \sum_w \frac{b(w|v)\mu(x,w)}{\sum_{v' \in f^{-1}(x)} b(w|v')} > \sum_{v \in f^{-1}(x)} \sum_w \frac{b(w|v)\mu(y,w)}{\sum_{v' \in f^{-1}(y)} b(w|v')}$$

This simplifies into

$$\frac{\alpha(x)}{\alpha(y)} > \sum_w \left(\frac{\sum_{v \in f^{-1}(x)} b(w|v)}{\sum_{v \in f^{-1}(y)} b(w|v)} \right) \beta_y(w) \tag{4}$$

whenever $f(v) = x$. Similarly, we obtain

$$\frac{\alpha(y)}{\alpha(x)} > \sum_w \left(\frac{\sum_{v \in f^{-1}(y)} b(w|v)}{\sum_{v \in f^{-1}(x)} b(w|v)} \right) \beta_x(w) \tag{5}$$

whenever $f(v) = y$. Rearranging these inequalities, and writing $\beta_x(w) = \beta_y(w) = \beta(w)$ for every w, we can see that the problem is to find a collection of real coefficients $(\psi(w))_{w \in W}$ such that

$$\begin{aligned} \sum_w \beta(w)\psi(w) &< 1 \\ \sum_w \frac{\beta(w)}{\psi(w)} &< 1 \end{aligned}$$

where

$$\psi(w) = \frac{\alpha(x) \sum_{v \in f^{-1}(y)} b(w|v)}{\alpha(y) \sum_{v \in f^{-1}(x)} b(w|v)}$$

To see why this is impossible, add up the two inequalities:

$$\sum_w \beta(w) \left[\psi(w) + \frac{1}{\psi(w)} \right] < 2$$

But $\psi(w) + 1/\psi(w)$ attains a minimum at $\psi(w) = 1$, and since $\Sigma_w \beta(w) = 1$, we obtain a contradiction.

Now suppose that $\beta_x \neq \beta_y$ for every distinct x, y. Fix some onto function f. By the definition of b^*, $b^*(w|v) = b^*(w|v')$ whenever $f(v) = f(v')$. Therefore, inequality (3) is simplified into

$$\frac{\alpha(x)}{k_f(x)} > \sum_w \frac{b^*(w|v_x)\alpha(y)\beta_y(w)}{k_f(y)b^*(w|v_y)}$$

whenever $f(v_x) = x$, $f(v_y) = y$. Now plug the definition of b^* as described in the statement of the proposition, and obtain the inequality

$$\sum_w \sqrt{\beta_x(w)\beta_y(w)} = BC(\beta_x, \beta_y) < 1$$

This inequality indeed holds whenever $\beta_x \neq \beta_y$. ■

The key argument in the proof of necessity is that the Battachayyara coefficient of two identical distributions ($\beta_x = \beta_y$) cannot be lower than one. The sufficiency argument exploits the property that $BC(\beta_x, \beta_y) < 1$ whenever $\beta_x \neq \beta_y$, which implies a slack in the equilibrium requirement that a keyword is bought by the firm type with the highest conversion rate. This slack gives us enough freedom in selecting an appropriate broad match function.

These arguments reveal that if we did not require condition (ii) in Definition 1 to involve *strict* inequalities, it would be possible to construct a broad match function that implements an efficient equilibrium for all μ, simply by setting $b(w|v) = \alpha(x)/k_f(x)$ whenever $f(v) = x$, such that $\psi(w) = 1$ for all w. This would imply that all firm types get a conversion rate of $1/n$ from all words. This is highly non-robust to small perturbations in μ.

Let us now turn to the keyword prices induced by the efficient equilibrium we constructed for b^*.

Proposition 4 *Let f be an onto function. Define b^* as in (2), and let $a(x, w) = 1$ whenever $\mu(x, w) > 0$. Then, the keyword prices induced by the efficient market equilibrium (f, a) are given by*

$$p^*(v) = \frac{1}{n \cdot \overline{BC}(f(v))}$$

Moreover, $p^*(w)$ decreases when the matrix $(\beta_x)_{x \in X}$ undergoes Blackwell garbling.

Proof. By definition (f, a subscripts are omitted),

$$p^*(v) = CR(f(v), v) = \frac{\sum_w b^*(w|v)\pi(f(v), w)}{\sum_w b^*(w|v)\sum_y \pi(y, w)}$$

Plugging the definition of b^*, the numerator is

$$\sum_w \frac{\alpha(f(v))\sqrt{\beta_{f(v)}(w)}}{k(f(v))} \cdot \frac{\alpha(f(v))\beta_{f(v)}(w)}{k(f(v)) \cdot \frac{\alpha(f(v))}{k(f(v))} \cdot \sqrt{\beta_{f(v)}(w)}} = \frac{\alpha(f(v))}{k(f(v))}$$

The denominator is

$$\sum_w \sum_x \frac{\alpha(f(v))\sqrt{\beta_{f(v)}(w)}}{k(f(v))} \cdot \frac{\alpha(x)\beta_x(w)}{k(x) \cdot \frac{\alpha(x)}{k(x)} \cdot \sqrt{\beta_x(w)}} = \frac{\alpha(f(v))}{k(f(v))}\sum_x \sum_w \sqrt{\beta_x(w)\beta_{f(v)}(w)}$$

which by the definition of $\overline{BC}(f(v))$ yields the desired expression for $p^*(v)$.

Let us now show that $p^*(w)$ decreases when $(\beta_x)_{x\in X}$ is subjected to a Blackwell garbling. For notational simplicity, we will write $\beta_x(w) = \beta_{ik}$, where i indexes the product x and k indexes the word w. Thus, (β_{ik}) is a stochastic matrix, with $\Sigma_k \beta_{ik} = 1$ for every i. Let

$$\delta_{ik} = \sum_h \beta_{ih} m_{hk}$$

where (m_{hk}) is a $|W| \times |W|$ bi-stochastic matrix. Fix i, j. Then,

$$\sum_k \sqrt{\delta_{ik}\delta_{jk}} = \sum_k \sqrt{\left(\sum_h \beta_{ih} m_{hk}\right)\left(\sum_h \beta_{jh} m_{hk}\right)}$$

By the Cauchy-Schwarz inequality, this expression is weakly greater than

$$\begin{aligned} & \sum_k \sum_h \sqrt{\beta_{ih} m_{hk} \beta_{jh} m_{hk}} \\ = & \sum_h \sqrt{\beta_{ih}\beta_{jh}} \sum_k m_{hk} \\ = & \sum_k \sqrt{\beta_{ik}\beta_{jk}} \end{aligned}$$

Since this inequality holds for every i, j, it follows that

$$\sum_i \sum_k \sqrt{\delta_{ik}\delta_{jk}} \geq \sum_i \sum_k \sqrt{\beta_{ik}\beta_{jk}}$$

Thus, the denominator of $p^*(w)$ increases when $(\beta_x)_{x\in X}$ undergoes a Blackwell garbling, which completes the proof. ■

Thus, under b^*, keyword prices decrease as consumers' queries become less informative. To illustrate the comparative statics, let $X = W$, and consider two extreme cases. First, suppose that $\beta_x(x) = 1$ for all x - i.e., consumers can perfectly describe their wants. Then, $p^*(w) = 1$ for every w under the efficient equilibrium induced by b^*. Second, suppose that $\beta_x(w) \approx 1/n$ for every w, x (the approximate equality is meant to ensure that the condition for existence of an efficient equilibrium is met). In this case, there is virtually no correlation between consumers' favorite product and their query, and we have $p^*(w) \approx 1/n$ for every w.

5.2 Desirable Properties of b^*

The broad match function b^* given by (2) is not the only one that induces an efficient market equilibrium. It is clear from the proof that for a fixed μ, any small perturbation of b^* would also work. However, in this sub-section we show that b^* has certain optimality properties that distinguish it from other broad match functions. It also turns out to be strongly linked to the optimal broad match function when $s > 0$.

Robustness to relative popularity shocks
Condition (ii) in Definition 1 requires that for any pair of products x and y, x (y) firms have a strictly higher willingness to pay for the words in $f^{-1}(x)$ ($f^{-1}(y)$) than y (x) firms. In the proof of Proposition 3, we simplified this constraint into the following inequalities (see (4)):

$$\begin{aligned}
\frac{\alpha(x)}{\alpha(y)} &> \sum_w \left(\frac{\sum_{v\in f^{-1}(x)} b(w|v)}{\sum_{v\in f^{-1}(y)} b(w|v)} \right) \beta_y(w) \\
\frac{\alpha(y)}{\alpha(x)} &> \sum_w \left(\frac{\sum_{v\in f^{-1}(y)} b(w|v)}{\sum_{v\in f^{-1}(x)} b(w|v)} \right) \beta_x(w)
\end{aligned}$$

Now, suppose that when the social planner designs the broad match function b, he has some uncertainty regarding the relative popularity of different products (for the sake of the argument, he has no uncertainty regarding the conditional query distributions). In particular, he is concerned that the ratio on the L.H.S of the above inequality may in fact be equal to $\theta\alpha(x)/\alpha(y)$, where $\theta > 0$ is unknown.

The broad match function b^* has the property that it reduces the R.H.S of both inequalities to $BC(\beta_x, \beta_y)$. Moreover, it minimizes the sum of the R.H.S of the two inequalities. Therefore, it ensures that the broad match function induces an efficient market outcome for the largest possible set of values that θ could get - namely, any $\theta > BC(\beta_x, \beta_y)$. In this sense, b^* maximizes robustness to fluctuations in products' relative popularity.

Firms' waiting time

Although only consumers conduct search in this model, it is interesting to evaluate market outcomes also in terms of the time resources that firms spend in the market, namely, the number of clicks that firms experience before transacting with a consumer. For a firm of type x, the effective conversion rate across all the words it is allocated is

$$CR^*(x) = \frac{\sum_{v \in f^{-1}(x)} \sum_w b(w|v)\pi(x,w)}{\sum_{v \in f^{-1}(y)} \sum_w b(w|v) \sum_y \pi(y,w)}$$

The expected waiting time for x firms is $1/CR^*(x)$. Plugging the definition of π, we obtain that the average waiting time across all firm types is

$$\frac{1}{n}\sum_x \frac{1}{CR^*(x)} = \sum_x \sum_y \sum_w \beta_y(w) \cdot \frac{\alpha(y) \sum_{v' \in f^{-1}(x)} b(x|v')}{\alpha(x) \sum_{v' \in f^{-1}(y)} b(y|v')} \tag{6}$$

Suppose that the planner's objective is to find a broad match function that minimizes (6).

Denote

$$\eta_{x,y}(w) \equiv \frac{\alpha(y) \sum_{v' \in f^{-1}(x)} b(x|v')}{\alpha(x) \sum_{v' \in f^{-1}(y)} b(y|v')}$$

Then, the problem is to find a collection of $\eta_{x,y}(w)$ that minimizes

$$\sum_x \sum_y \sum_w \beta_y(w)\eta_{x,y}(w)$$

Note that $\eta_{x,y}(w)\eta_{y,x}(w) = 1$. Therefore, for every pair x, y, we should minimize

$$\sum_w \left[\beta_y(w)\eta_{x,y}(w) + \beta_x(w)\frac{1}{\eta_{x,y}(w)}\right]$$

The solution is

$$\eta_{x,y}(w) = \sqrt{\frac{\beta_x(w)}{\beta_y(w)}}$$

Hence, b^* is a solution to this problem. It follows that among all broad match functions that induce an efficient market outcome, b^* is the optimal in terms of the firms' average waiting time, as defined by (6).

5.3 Efficient equilibrium under $s > 0$

In the presence of positive search costs, a consumer of type (x, w) will search only if the probability of drawing an x firm when submitting the query w (i.e., $\lambda(x, w)$) is above the search cost s. Such a consumer will search until he finally finds x. His expected payoff is thus equal to his willingness to pay for x (which we have normalized to one) minus the expected search costs, which are equal to the cost per draw divided by the probability of drawing x. Clearly, if we are interested in maximizing social welfare, we can ignore cases in which $\lambda(x, w)a(x, w) = 0$ (we do not want a consumer to waste search resources if he cannot find his desired product, and we do not want a product to be present in a search pool if no one is searching for this product). Therefore, social welfare can be written as follows:

$$\sum_{(x,w)|\lambda(x,w)>0} \mu(x, w)[1 - \frac{s}{\lambda(x, w)}] \tag{7}$$

It is immediately clear from this expression that we can calculate the optimal λ independently for each word.

Let $\lambda^* = (\lambda^*(x, w))_{x \in X, w \in W}$ maximize total surplus. We refer to λ^* as the collection of efficient search pools. This collection must satisfy the following properties. First, if $\lambda^*(x, w) > 0$, then $\lambda^*(x, w) > s$. It follows that if (f, a) induces λ^*, it *must* satisfy condition (i) in Definition 1. To see why this property holds, note that if $\lambda^*(x, w) \leq s$, then a consumer of type (x, w) will not enter the search pool associated with the word w. If nevertheless $\lambda^*(x, w) \in (0, s)$, then the presence of x firms exerts an unnecessary negative externality on consumers who *do* enter this search pool in search of $y \neq x$, since these consumers face a lower probability of finding their product (they sometimes draw an x-firm). Given a vector λ, we denote by $X(w|\lambda)$ the set of products satisfying $\lambda(x, w) \geq s$.

Second, by the first-order conditions, for every word $w \in W$ and product pair

$x, y \in X(w|\lambda^*)$,

$$\frac{\lambda^*(x,w)}{\lambda^*(y,w)} = \sqrt{\frac{\mu(x,w)}{\mu(y,w)}}$$

Since $\sum_{x \in X} \lambda^*(x,w) = 1$, we obtain that for every $w \in W$ and every $x \in X(w|\lambda^*)$,

$$\lambda^*(x,w) = \frac{\sqrt{\mu(x,w)}}{\sum_{y \in X(w|\lambda^*)} \sqrt{\mu(y,w)}} \tag{8}$$

We proceed next to characterize the set of consumer types that decide to enter the efficient search pools as given by λ^*. We begin by noting the following property of efficient search pools.

Lemma 1 *If $\lambda^*(x,w) = 0$ and $\mu(y,w) < \mu(x,w)$, then $\lambda^*(y,w) = 0$.*

Proof. Assume by contradiction that there exists a word w and a pair of products x, y such that $\mu(y,w) < \mu(x,w)$ but $\lambda^*(y,w) > \lambda^*(x,w) = 0$. Consider changing λ^* to λ' where the only difference is that $\lambda'(y,w) = 0$ and $\lambda'(x,w) = \lambda^*(y,w)$. This changes social welfare by the following amount

$$[\mu(x,w) - \mu(y,w)][1 - \frac{s}{\lambda^*(y,w)}]$$

Since $\lambda^*(y,w) > s$ the change is positive, a contradiction. ■

This lemma has the following implication. For each word w, enumrate products in decreasing order of popularity, and denote $\mu_i = \mu(i,w)$, such that $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$. The efficient λ^* has the property that for each w, there exists a cutoff type m^* such that $\lambda_i^* > 0$ for $i \leq m^*$ and $\lambda_i^* = 0$ for $i > m^*$. The efficient cutoff type is characterized as follows.

Lemma 2 *The cutoff m^* is a consumer type with the minimal proportion that still satisfies the inequality*

$$\sqrt{\mu_m} > \frac{2s}{1-s} \sum_{i=1}^{m-1} \sqrt{\mu_i} \tag{9}$$

Proof. With slight abuse of notation, define

$$U(m) \equiv \sum_{i=1}^{m} \mu_i (1 - s \frac{\sum_{j=1}^{m} \sqrt{\mu_j}}{\sqrt{\mu_i}})$$

For any $m \in \{1, \dots, n\}$,

$$U(m) - U(m-1) = \mu_m(1-s) - 2s\sqrt{\mu_m}\sum_{i=1}^{m-1}\sqrt{\mu_i} \tag{10}$$

Type m is the cutoff type if $U(i)-U(i-1) > 0$ for every $i \leq m$ and $U(i)-U(i-1) < 0$ for every $i > m$. Notice that as m increases, μ_m decreases while $\sum_{i=1}^{m-1}\sqrt{\mu_i}$ increases. Hence, if the R.H.S. is negative for some m it would also be negative for any $m' \geq m$. It follows that there exists a maximal index $m^* \in \{1, \dots, n\}$ for which $U(m)-U(m-1) > 0$. By (10), m^* satisfies that for any consumer type with $\mu_m \geq \mu_{m^*}$,

$$\sqrt{\mu_m} > \frac{2s}{1-s}\sum_{i=1}^{m-1}\sqrt{\mu_i}$$

while for any consumer type with $\mu_m < \mu_{m^*}$ this inequality is reversed. ■

Equation (10) illustrates the negative externality that consumer types exert on each other. The first term on the R.H.S represents the welfare gain that is realized when the marginal consumer type m finds his desired product: on his final draw, he gets a net payoff of $1-s$. The second term represents the welfare loss due to the search costs incurred by the marginal consumer as well as the added search costs that he inflicts on other consumers (they now search longer since sometimes they draw m's desired product).

We say that a market equilibrium (f, a) is efficient if it induces the optimal proportions $(\lambda^*(x, w))_{x\in X, w\in W}$. Our next result characterizes the necessary and sufficient conditions for the existence of a broad match function that induces an efficient equilibrium.

Proposition 5 *Let (f, a) be an efficient market equilibrium. There exists a broad match function b^* that induces (f, a) if and only if*

$$\sqrt{\frac{\alpha(x)}{\alpha(y)}} \cdot BC(\beta_x, \beta_y) < 1 \tag{11}$$

for every pair of distinct products x, y in the image of f. In particular, b^* is given by

$$b^*(w, v) = \begin{cases} \frac{\sqrt{\mu(f(v),w)}}{k_f(f(v))} & if \quad \lambda^*(f(v), w) > 0 \\ 0 & if \quad \lambda^*(f(v), w) = 0 \end{cases} \tag{12}$$

and induces the following price function,

$$p^*(v) = \frac{1}{\sum_y \sqrt{\frac{\alpha(y)}{\alpha(f(v))}} BC(\beta_{f(v)}, \beta_y(w))} \tag{13}$$

which decreases when the matrix $(\beta_x)_{x \in X}$ undergoes Blackwell garbling.

Proof. Assume first that there exists a broad match function b^* that induces an efficient equilibrium (f, a). Then, we can plug the definition of $\lambda^*(x, w)$ and the necessary condition for efficiency given by (8) in inequalities (4) and (5) (which represent condition (ii) in Definition 1), and obtain that the following inequalities must hold for every pair of distinct products x, y in the image of f:

$$\begin{aligned} \alpha(x) &> \sqrt{\alpha(x)\alpha(y)} \sum_{w \in W} \sqrt{\beta_x(w)\beta_y(w)} \\ \alpha(y) &> \sqrt{\alpha(x)\alpha(y)} \sum_{w \in W} \sqrt{\beta_x(w)\beta_y(w)} \end{aligned}$$

By the definition of $BC(\beta_x, \beta_y)$ these inequalities may be rewritten as

$$\begin{aligned} 1 &> \frac{\sqrt{\alpha(y)}}{\sqrt{\alpha(x)}} BC(\beta_x, \beta_y) \\ 1 &> \frac{\sqrt{\alpha(x)}}{\sqrt{\alpha(y)}} BC(\beta_x, \beta_y) \end{aligned}$$

which imply the desired condition (11).

For the sufficiency part of the proof, note that by construction, b^* yields the optimal vector of proportions $(\lambda^*(x, w))_{x \in X, w \in W}$. We have already noted that this immediately implies Condition (i) in Definition 1. Plugging the definition of b^* into inequalities (4) and (5) establishes Condition (ii).

It remains to show that b^* as defined in (12) induces the price function given by (13). By definition, the price of word v may be written as

$$p^*(v) = \frac{\sum_w b^*(w, v)\pi(f(v), w)}{\sum_w b^*(w, v) \sum_y \pi(y, w)}$$

Plugging in the expression for b^* and π we obtain that the numerator is equal to

$$\sum_w \frac{\sqrt{\mu(f(v), w)}}{k_f(f(v))} \cdot \frac{\alpha(f(v))\beta_{f(v)}(w)}{k_f(f(v)) \cdot \frac{\sqrt{\mu(f(v), w)}}{k_f(f(v))}} = \frac{\alpha(f(v))}{k_f(f(v))}$$

and the denominator equals

$$\sum_w \sum_y \frac{\sqrt{\alpha(f(v))\beta_{f(v)}(w)}}{k_f(f(v))} \cdot \frac{\alpha(y)\beta_y(w)}{k(y) \cdot \frac{\sqrt{\mu(y,w)}}{k(y)}} = \sum_w \sum_y \frac{\sqrt{\alpha(f(v))\beta_{f(v)}(w)\alpha(y)\beta_y(w)}}{k_f(f(v))}$$

Hence,

$$p^*(v) = \frac{1}{\sum_y \sqrt{\frac{\alpha(y)}{\alpha(f(v))}} \sum_w \sqrt{\beta_{f(v)}(w)\beta_y(w)}} = \frac{1}{\sum_y \sqrt{\frac{\alpha(y)}{\alpha(f(v))}} BC(\beta_{f^*(v)}, \beta_y(w))}$$

The proof that p^* decreases when $(\beta_x)_{x\in X}$ is subjected to Blackwell garbling is the same as in Proposition 4 ■

Condition (11) captures in a succinct way the two key considerations highlighted in the examples. On one hand, a high $BC(\beta_x, \beta_y)$ captures an environment with a high search friction, in the sense that consumers' queries are weak indicators of their true wants. Broad match is meant to address this problem, by diversifying the consumers' access to products. However, since words are allocated to firms via market competition, broad match may increase the risk that a mass-appeal product will crowd out a niche product. This risk increases with the popularity gap between the two products, captured by $\alpha(x)/\alpha(y)$, gets farther away from one.

Note that when all products are equally popular (i.e., $\alpha(x) = 1/n$ for all x), the necessary and sufficient condition for market implementability of efficient outcomes coincides with the $s = 0$ case, over the subset of products $\cup_{w\in W} X^*(w)$. The intuition is as follows. As we saw in Sub-section 5.2, the broad match function b^* is optimal in the sense that it minimizes an "average waiting time" for firms. The average was not weighted by the products' popularity. However, when all products are equally popular, this is the same as minimizing the search time for consumers.

Comment: Canonical b, f
Suppose that we impose the natural restriction $X \subseteq W$ - that is, the name of each product is a keyword. Then, it is also natural to impose $f(x) = x$ and $b(x|x) = 1$. Notice that this does not restrict the implementability of efficient market outcomes. First, our results allow f to be *any* onto function. Second, the equilibrium restrictions on b only impose constraints on the ratio $b(w|v)/b(v|v)$, and therefore we can select $b(v|v) = 1$ w.l.o.g. Of course, this would change the exact formula for $b(w|v)$, $w \neq v$.

6 Conclusion

The general question that this paper has raised is whether *information management* can be decentralized via competitive markets. Before the advent of sponsored online search, the thought that this important activity could be delegated to "the market" would have seemed bizarre. Think about the problem of finding an old receipt while preparing a tax return. Your success in finding it depends on the efficiency of the filing system you employed for storing receipts. Putting receipts in binder folders is one natural method. Preparing duplicates of receipts and filing them in a number of relevant folders is an even more sophisticated method, which is equivalent to associating a set of keywords with each receipt.

However, finding and maintaining an efficient classification system is a complex task, which has to incorporate the costs of storing and retrieving information. Successful decentralization of this task means that the optimal filing system can be simulated by spontaneous market competition for keywords.

Of course, market decentralization demands that for every stored object, there is an agent who is willing to pay for the object's retrieval. This will not always be the case, as in the example of your old receipts. However, in more commercial settings, the exercise performed in this paper suggests that efficient information management could be simulated by a competitive "market for keywords", as long as certain conditions on the distribution of "user types" are met. We hope to explore the question of whether this idea is relevant for other types of information management in future work.

References

[1] Anderson, C. (2008), "The long tail: Why the future of business is selling less of more," Hyperion Books.

[2] Armstrong, M. "Competition in two-sided markets," RAND Journal of Economics 37.3 (2006): 668-691.

[3] Athey, S. and G. Ellison (2011), "Position Auctions with Consumer Search," The Quarterly Journal of Economics, Oxford University Press, vol. 126(3), pages 1213-1270.

[4] Brynjolfsson, E., Yu J., and M. Smith (2006), "From Niches to Riches: Anatomy of the Long Tail." Sloan Management Review 47.4: 67-71.

[5] Caillaud, B. and B. Jullien (2001), "Competing cybermediaries." European Economic Review 45.4: 797-808.

[6] Caillaud, B. and B. Jullien (2003), "Chicken & egg: Competition among intermediation service providers." RAND journal of Economics: 309-328.

[7] Chen, Y., C. He (2011), "Paid Placement: Advertising and Search on the Internet," Economic Journal, Royal Economic Society, vol. 121(556), pages F309-F328, November.

[8] Edelman, B., M. Ostrovsky and M. Schwarz (2007), "Internet Advertising and the Generalized Second-Price Auction: Selling Billions of Dollars Worth of Keywords." American Economic Review 97.1 : 242-259.

[9] Eliaz, K. and R. Spiegler (2011), "A Simple Model of Search Engine Pricing," Economic Journal 121, 329-339.

[10] Hagiu, A. and B. Jullien (2011), "Why do intermediaries divert search?," RAND Journal of Economics, RAND Corporation, vol. 42(2), pages 337-362, 06.

[11] Kranton, R. and D. Minehart (2001), "A Theory of Buyer-Seller Networks," American Economic Review, American Economic Association, vol. 91(3), pages 485-508, June.

[12] Rochet, J. and J. Tirole (2003), "Platform Competition in Two-Sided Markets," Journal of the European Economic Association, MIT Press, vol. 1(4), pages 990-1029, 06.

[13] Spiegler, R. (2000), "Extracting Interaction-Created Surplus," Games and Economic Behavior, Elsevier, vol. 30(1), pages 142-162, January.

[14] Theodoridis. S. and K. Koutroumbas (2008), Pattern Recognition, 4th Edition, Academic Press.