Bias and Size Effects of Price-Comparison Search Engines: Theory and Experimental Evidence*

AURORA GARCÍA-GALLEGO[†]

NIKOLAOS GEORGANTZÍS[‡]

PEDRO PEREIRA[§]

JOSÉ C. PERNÍAS-CERRILLO[¶]

30th January 2007

This article, analyzes the impact on consumer prices of the size and bias of price comparison search engines. we develop a model, related to Burdett and Judd (1983) and Varian (1980), and test experimentally several theoretical predictions. The experimental results confirm the model's predictions regarding the impact of the number of firms, and the type of bias of the search engine, but reject the model's predictions regarding changes in the size of the index.

Keywords: Search engines, incomplete information, biased information, price levels, experiments.

JEL Codes: C91, D43, D83, L13.

Corresponding author: Pedro Pereira: *Autoridade da Concorrência, R. Laura Alves, nº 4, 7º, 1050-138 Lisboa (Portugal); e-mail: jppereira@oninetspeed.pt; phone: 21-7802498; fax: 21-7802471.*

*The authors gratefully acknowledge financial support by the NET Institute, the Spanish Ministry of Science and Technology (BEC2002-04380-CO2-02), and the Generalitat Valenciana (GREC-GRUPOS04/13). This article benefited from helpful comments by H. Raff, J. Brandts, J. Bröcker, T. Cason, S. Hoernig, Martin Sefton, and anonymous seminar participants.

[†]LEE/LINEEX and Economics Department, Universitat Jaume I, Castellón (Spain); e-mail: mgarcia@eco.uji.es; phone: 34-964728604.

[‡]LEE/LINEEX and Economics Department, Universitat Jaume I, Castellón (Spain); e-mail: georgant@eco.uji.es; phone: 34-964728588.

[§]Autoridade da Concorrência, R. Laura Alves, nº 4, 7º, 1050-138 Lisboa (Portugal); e-mail: jppereira@oninetspeed.pt; phone: 21-7802498; fax: 21-7802471.

[¶]LEE and Economics Department, Universitat Jaume I, Castellón (Spain); e-mail: pernias@eco.uji.es; phone: 34-964728612.

1. INTRODUCTION

1.1. Preliminary Thoughts

From the consumers' perspective, one of the most promising aspects of e-commerce was that it would reduce search costs. With search engines, consumers could easily observe and compare the prices of a large number of vendors, and identify bargains. The consumers' enhanced ability of comparing prices would discipline vendors, and put downward pressure on prices.[1]

Price Comparison Search Engines, also known as shopping agents or shopping robots, are a class of search engines, which access and read Internet pages, store the results, and return lists of pages, which match keywords in a query. They consist of three parts: (i) a crawler, (ii) an index, and (iii) the relevance algorithm. The *Crawler*, or spider, is a program that automatically accesses Internet pages, reads them, stores the data, and then follows links to other pages. The *Index*, or catalog, is a database that contains the information the crawler finds. *The Relevance Algorithm* is a program that looks in the index for matches to keywords, and ranks them by relevance, which is determined through criteria such as link analysis, or, click-through measurements. This description refers to crawler-based systems, such as Google or AltaVista. There are also *Directories*, like Yahoo was initially, in which lists are compiled manually. Most systems are hybrid.

Presumably, the larger the number of vendors whose price a search engine lists on its site, and that thereby consumers can easily compare, the more competitive the market becomes. However, there are several technical reasons for search engines to cover only a small subset of the Internet, and to collect and report information biased in favor of certain vendors. This perspective is discussed in Hoernig and Pereira (2007) and documented by several studies (Bradlow and Schmittlein, 1999; Lawrence and Giles, 1998, 1999).

The technology-induced tendency, for search engines to have incomplete and biased coverage, is reinforced by economic reasons. Search engines are profit-seeking institutions, which draw their income from vendors, either in the form of placement fees, sales commissions, or advertising (Hoernig and Pereira, 2007).

In this article we examine, theoretically and experimentally, the impact on consumer prices on electronic markets, of price comparison search engines covering only a small subset of the Internet, and collecting and reporting information being biased in favor of certain vendors.

1.2. Overview

We develop a partial equilibrium search model, related to Burdett and Judd (1983) and Varian (1980), to discuss the implications of price comparison search engines providing consumers with incomplete and biased information. There is: (i) one price comparison search engine, (ii) a finite number of identical vendors, and (iii) a large number of consumers of two types. Shoppers use the price comparison search engine, and buy at the lowest price listed by the search engine. Non-shoppers buy from a vendor chosen at random. Vendors choose prices. In equilibrium, vendors randomize between charging a higher price and selling only to non-shoppers, and charging a lower price to try to sell also to shoppers.

In the benchmark case, the search engine has complete coverage, i.e., lists the prices of all vendors present in the market. We also analyze two other cases. First, the case in which the price comparison search engine has incomplete coverage, and is unbiased with respect to vendors. This case can be thought of as portraying the situation where the search engine is a crawler-based system, which has no placement

1. The search literature has no simple prediction about the relation between search costs, price levels, or price dispersion (Pereira, 2005; Samuelson and Zhang, 1992).

deals with any particular vendor. Second, we analyze the case in which the price comparison search engine has incomplete coverage, and is biased in favor of certain vendors. This case can be thought of as portraying the situation where the search engine is either a crawler-based system or a directory, which has placement deals with certain vendors. The search engines' decisions of how many vendors to index, or the vendors' decisions of whether to become indexed, are obviously of interest. Hoernig and Pereira (2007) analyzes these questions. In this article we take these decisions as given, and focus on the implications for the pricing behavior of vendors.

The theoretical analysis makes several predictions regarding the impact of: (i) the number of vendors, (ii) the size of the search engine, and (iii) the type of bias of the search engine. In addition, the model also draws attention to four general counter intuitive effects. First, there is a conflict of interests between types of consumers, i.e., between shoppers and non-shoppers. This makes it hard to evaluate the welfare impact of these effects. Second, more information, measured by a wider Internet coverage by the price comparison search engine, is not necessarily desirable. It benefits some consumers, but harms others. Third, unbiased information about vendors is not necessarily desirable either, and for the same reason. Fourth, the effects of entry in these markets are complex, and depend on the way entry occurs.

We tested the predictions of the theoretical analysis, in a laboratory experiment designed specifically to evaluate the model's testable propositions.

The experimental results confirm the model's predictions regarding the impact of the number of firms, and the type of bias of the search engine, but reject the model's predictions regarding changes in the size of the index. The analysis of the data indicated several additional patterns, such as, that prices are lower under biased incomplete coverage than under unbiased incomplete coverage.

1.3. Literature Review

The basic search models relevant for our research were developed in the early 1980s by Varian (1980), Rosenthal (1980), and Burdett and Judd (1983). Varian (1980) developed a model where firms are identical and there are two types of consumers. Shoppers, can observe costlessly all prices, while non-shoppers, observe only one price, perhaps because they have a prohibitive search cost. Firms randomize between charging low prices to try to sell to shoppers, and charging high prices to sell only to their share of non-shoppers. Stahl (1989) endogeneized the non-shoppers reservation price through sequential search, and performed several comparative statics exercises. He showed that for this class of models, the equilibrium price distribution ranges between the monopoly and the competitive levels, as the fraction of shoppers varies between 0 to 1.

Recently, the use of the Internet and the existence of price-comparison search engines, renewed the interest in the search literature, e.g., Baye and Morgan (2001), Dinlersoz and Pereira (2007), Hoernig and Pereira (2007), Iyer and Pazgal (2000). Baye and Morgan (2001) developed a model where firms can sell on their local market, or sell beyond their local market through the Internet. Similarly, non-shoppers are consumers buying locally, whereas shoppers are consumers buying over the Internet. The price equilibrium involves mixed strategies. Iyer and Pazgal (2000) developed another model of electronic commerce where firms play mixed strategies, and tested it empirically. Brynjolfsson and Smith (2000) also confirmed empirically the prediction of persistent price dispersion in Internet markets, even in the presence of homogeneous products.

Among the burgeoning experimental literature on search markets, two laboratory experiments are of particular interest to our research. Cason and Friedman (2003) studied markets with costly consumer search. The issue of the size of the sample used in the search process was explicitly addressed allowing for both human and automated price search on the demand side. Theoretical predictions in their model were mostly supported by the evidence. In a setup which corresponds to the full coverage case of our analysis, Morgan et al. (2006) studied the case of a consumer population consisting of Internet shoppers and non-shoppers. The theoretical predictions concerning the impact of firm number and the proportion of uninformed consumers on equilibrium price distributions are mostly confirmed by their results. However, observed price distributions exhibit systematic deviations from the corresponding equilibrium predictions.

The remainder of the article is organized as follows. In Section 2 we present the benchmark model, and in Section 3 we characterize its equilibrium. In Section 4 we conduct the analysis of the model and its variations. Section 5 analyzes the results of the experiment. Section 6 concludes. Appendix A and B include, respectively, the proofs and the experimental instructions.

2. THE MODEL

2.1. The Setting

Consider an electronic market for a homogeneous search good that opens for 1 period. There are: (i) 1 price comparison search engine, (ii) $n \geq 3$ vendors, which we index through subscript $j = 1, \ldots, n$, and (iii) many consumers.

2.2. Price-Comparison Search Engine

The *Price-Comparison Search Engine*, in response to a query for the product, lists the addresses of the firms contained in its database, i.e., in its *Index*, and the prices they charge. Denote by k, the number of vendors the price-comparison search engine indexes. We will refer to k as the *Size of the Index* .The search engine may be of one of the following three types:

(i). *Complete Coverage*: The search engine has *Complete Coverage* if it indexes all vendors present in the market: $k = n$.

(ii). *Unbiased Incomplete Coverage*: The search engine has *Incomplete Coverage* if it does not index all vendors: $1 < k < n$. In addition, the search engine has an *Unbiased Sample* if it indexes each of the n vendors with the same probability: $\binom{n-1}{k-1}/\binom{n}{k} = k/n$. When the search engine has incomplete coverage and an unbiased sample, we say that it has *Unbiased Incomplete Coverage*.

(iii). *Biased Incomplete Coverage*: A search engine with incomplete coverage has a *Biased Sample* if it indexes vendors $j = 1, \ldots, k$, and does not index vendors $j = k+1, \ldots, n$. When the search engine has incomplete coverage and a biased sample, we say that it has *Biased Incomplete Coverage*. For this parametrization, knowing the probability with which vendors are indexed, implies knowing the identity of the indexed vendors.[2]

Denote by τ the type of the search engine, and let 'c', mean *Complete Coverage*, 'u' mean *Unbiased Incomplete Coverage*, and 'b' mean *Biased Incomplete Coverage*, i.e., τ belongs to $\{c, u, b\}$. We will use superscripts 'c', 'u', 'b', to denote variables or values associated with the cases where the search engine has that type.

2. There are alternative ways of introducing sample biasedness, for which knowing the probability with which vendors are indexed does not imply knowing the identity of the indexed vendors. For example, all vendors can be indexed with a non-degenerate probability, which is higher for some vendors than for others. The advantage of our parametrization is that it yields a closed form solution.

2.3. *Consumers*

There is a unit measure continuum of risk neutral consumers. Each consumer has a unit demand, and a reservation price of 1. There are 2 types of consumers, differing only with respect to whether they use the price-comparison search engine. *Non-Shoppers*, a proportion λ on $(0,1)$ of the consumer population, do not use the price-comparison search engine, perhaps because they are unaware of its existence, or perhaps because of the high opportunity cost of their time.[3] The other consumers, *Shoppers*, use the price-comparison search engine.

Consumers do not know the prices charged by individual vendors. Shoppers use the price-comparison search engine to learn the prices of vendors. If the lowest price sampled by the price-comparison search engine is no higher than 1, shoppers accept the offer and buy; in case of a tie they distribute themselves randomly among vendors; otherwise they reject the offer and exit the market. Non-shoppers distribute themselves evenly across vendors, i.e., each vendor has a share of non-shoppers of $1/n$. If offered a price no higher than 1, non-shoppers accept the offer and buy; otherwise they reject the offer and exit the market.

2.4. *Vendors*

Vendors are identical and risk neutral. Marginal costs are constant and equal to zero. Vendors know the functioning rules of the search engine. In particular, under *Unbiased Incomplete Coverage*, vendors know the probability with which they are indexed, but do not observe the identity of the indexed vendors, before choosing prices. Under *Biased Incomplete Coverage*, vendors know the identity of the indexed vendors before choosing prices. In the cases of *Complete Coverage* and *Unbiased Incomplete Coverage*, vendors are identical. In the case of *Biased Incomplete Coverage*, vendors are asymmetric. Denote by $\Pi_j(p)$, the expected profit of vendor j when it charges price p on $\mathbb{R}_0^+$. A vendor's *strategy* is a cumulative distribution function over prices, $F_j(\cdot)$. Denote the lowest and highest prices on its support by $\underline{p}_j$ and $\bar{p}_j$.[4] A vendor's *payoff* is its expected profit.

2.5. *Equilibrium*

A *Nash equilibrium* is a n-tuple of cumulative distribution functions over prices, $\{F_1(\cdot),\ldots,F_n(\cdot)\}$, such that for some $\Pi_j^\star$ on $\mathbb{R}_0^+$, and $j=1,\ldots,n$: (i) $\Pi_j(p)=\Pi_j^\star$, for all p on the support of $F_j(\cdot)$, and (ii) $\Pi_j(p)\leq\Pi_j^\star$, for all p.

When vendors are identical we focus on symmetric equilibria, in which case: $F_j(\cdot)=F(\cdot)$, $\underline{p}_j=\underline{p}$, $\bar{p}_j=\bar{p}$ and $\Pi_j^\star=\Pi^\star$, for all j.

3. To use a search engine consumers might have to download software, learn how to use the search engine's interface, configure the interface, wait for the data to be downloaded, etc. These reasons might dissuade some consumers from using search engines. This perspective agrees with the available evidence, which suggests a very limited use of price-comparison search engines, at least yet. A Media Metrix study found that during July 2000 less than 4% of Internet users used a price-comparison search engine, while over 66% visited an online retailer (Montgomery et al., 2004). Furthermore, a Jupiter Communications survey found that 28% of the respondents where unaware of the existence of price-comparison search engine (Iyer and Pazgal, 2000).

4. As it is well known this game has no equilibrium in pure strategies (Varian, 1980).

3. CHARACTERIZATION OF EQUILIBRIUM

Denote by ϕ_j^τ the probability of firm j being indexed, given that the search engine is of type τ:

$$\phi_j^\tau = \begin{cases} k/n & \text{if } \tau = c, u \\ 1 & \text{if } \tau = b \text{ and } j = 1, \ldots, k \\ 0 & \text{if } \tau = b \text{ and } j = k+1, \ldots, n. \end{cases}$$

Denote by $\hat{p}_{-j}$ the minimum price charged by any indexed vendor other than vendor j, and denote by $\hat{m}_{-j}$ the number of indexed vendors that charge $\hat{p}_{-j}$. The profit function of vendor j when it charges price p_j is:

$$\pi_j(p_j;\tau) = \begin{cases} p_j\left[\dfrac{\lambda}{n} + (1-\lambda)\phi_j^\tau\right] & \text{if } p_j < \hat{p}_{-j} \leq 1 \\ p_j\left[\dfrac{\lambda}{n} + \dfrac{1-\lambda}{\hat{m}_{-j}}\phi_j^\tau\right] & \text{if } p_j = \hat{p}_{-j} \leq 1 \\ p_j\dfrac{\lambda}{n} & \text{if } \hat{p}_{-j} < p_j \leq 1 \\ 0 & \text{if } 1 < p_j. \end{cases}$$

Ignoring ties,[5] the expected profit of a vendor that charges $p \leq 1$ is:[6]

$$\Pi_j(p) = p\frac{\lambda}{n} + p(1-\lambda)\phi_j^\tau[1-F(p)]^{k-1}. \tag{1}$$

Denote by l_j^τ the lowest price vendor j is willing to charge to sell to both types of consumers when the search engine has type τ, i.e., $l_j^\tau[\lambda/n + (1-\lambda)\phi_j^\tau] - \lambda/n \equiv 0$.

The next Lemma states some auxiliary results.

Lemma 1. *For all j:* (i) $l_j^\tau \leq \underline{p}_j \leq \bar{p}_j \leq 1$; (ii) F_j^τ *is continuous on* $[l_j^\tau, 1]$; (iii) $\bar{p}_j = 1$; (iv) $\Pi_j^\star = \lambda/n$; (v) $\underline{p}_j = l_j^\tau$; (vi) F_j^τ *has a connected support.* ♦

From Lemma 1(iv), in equilibrium:

$$p\frac{\lambda}{n} + p(1-\lambda)\phi_j^\tau[1-F_j^\tau(p)]^{k-1} = \frac{\lambda}{n}. \tag{2}$$

Denote by $\delta(p)$, the degenerate distribution with unit mass on p.[7] The next proposition characterizes the equilibrium for the model.[8]

5. Lemma 1(ii) shows that $F_j^\tau(\cdot)$ is continuous.

6. A vendor j that charges a price $p \leq 1$ sells to shoppers: (i) if it belongs to the set of vendors indexed by price-comparison search engine, which occurs with probability ϕ_j^τ , and (ii) if it has the lowest price among the indexed vendors, which occurs with probability $[1-F(p)]^{k-1}$. Thus, the expected share of shoppers of a vendor that charges price $p \leq 1$ is: $(1-\lambda)\phi_j^\tau[1-F(p)]^{k-1}$.

7. Function $\delta(\cdot)$ is the Heavside function, or the cumulative distribution function of the Dirac delta function.

8. The equilibrium described in Proposition 1 is the unique symmetric equilibrium. Baye et al. (1992, Theorem 1, p. 496) showed that there is also a continuum of asymmetric equilibria, where at least 2 firms randomize over $[l, 1]$, with each other firm i randomizing over $[l, x_i)$, $x_i < 1$, and having a mass point at 1 equal to $[1-F_i(x_i)]$.

Proposition 1. (i) *for $\tau = c, u$ and $j = 1, \dots, n$, and for $\tau = b$ and $j = 1, \dots, k$:*

$$F_j^\tau(p;n,k) = \begin{cases} 0 & \Leftarrow p < l_j^\tau \\ 1 - \left[\left(\frac{1}{n\phi_j^\tau} \right) \left(\frac{\lambda}{1-\lambda} \right) \left(\frac{1-p}{p} \right) \right]^{\frac{1}{k-1}} & \Leftarrow l_j^\tau \le p < 1 \\ 1 & \Leftarrow 1 \le p, \end{cases}$$

with

$$l_j^\tau(n) = \frac{\lambda}{\lambda + (1-\lambda)n\phi_j^\tau};$$

(ii) *for $\tau = b$ and $j = k+1, \dots, n$, $F_j^b(p;n,k) = \delta(1)$.* ♦

Under *Biased Incomplete Coverage*, vendors $j = k+1, \dots, n$, are not indexed for sure, and therefore have no access to shoppers. Since these vendors can only sell to non-shoppers, which are captive consumers, they charge the reservation price. Vendors $j = 1, \dots, k$, under *Biased Incomplete Coverage*, and all vendors in the cases *Complete Coverage* and *Unbiased Incomplete Coverage*, are indexed with positive probability. Hence, they face the trade-off of charging a high price and selling only to non-shoppers, or charging a low price to try to sell also to shoppers, which leads them to randomize over prices.

4. ANALYSIS

In this section, we analyze the model for the 3 types of search engine.

4.1. Complete Coverage

In the case of *Complete Coverage* the model is similar to Varian (1980).[9]

Rewrite (2) as:

$$\underbrace{p(1-\lambda)[1-F^c(p)]^{n-1}}_{\text{Marginal Benefit}} = \underbrace{\frac{\lambda}{n}(1-p)}_{\text{Opportunity Cost}}. \tag{3}$$

If a vendor charges a price p lower than the consumers' reservation price, it has the lowest price in the market with probability $[1-F^c(p)]^{n-1}$, sells to $(1-\lambda)$ shoppers, and earns an additional expected profit of $p(1-\lambda)[1-F^c(p)]^{n-1}$: the *Volume of Sales effect*. However, it loses $(1-p)$ per non-shopper, and a total of $(1-p)\lambda/n$: the *per Consumer Profit effect*. The volume of sales effect is the *marginal benefit* of charging a price lower than the consumers' reservation price, and the per consumer profit effect is the *marginal cost*.

Denote by ε, the expected price, i.e., the expected price paid by non-shoppers. And denote by μ, the expected minimum price, i.e., the expected price paid by shoppers.

The next Remark collects two useful observations.

Remark 1. (i) $\lambda\varepsilon^c + (1-\lambda)\mu^c = \lambda$*;* (ii) $\mu^c < \varepsilon^c$. ♦

The first part of Remark 1 says that the average price paid in the market, $\lambda\varepsilon^c + (1-\lambda)\mu^c$, equals the proportion of non-shoppers, λ.[10] This has two implications. First, only shifts in the proportion of

9. See also Rosenthal (1980) and Stahl (1989).

10. Actually, it equals the proportion of non-shoppers times the reservation price: $\lambda \cdot 1$. Also, since marginal cost is 0, and demand is inelastic and unitary, the average price paid in the market equals the average market profits.

non-shoppers change the average price paid in the market. Second, shifts in any other parameter, such as the number of vendors, n, induce the expected prices paid by shoppers and non-shoppers to move in opposite directions. A conflict of interests between types of consumers is a recurring theme of this article.

The second part of Remark 1, says that the expected price paid by shoppers, $\mu^c = l^c + \int_{l^c}^1 (1-F^c)^n dp$, is lower than the expected price paid by non-shoppers, $\varepsilon^c = l^c + \int_{l^c}^1 (1-F^c) dp$. The price-comparison search engine allows shoppers to compare the prices of all vendors in its index, and choose the cheapest vendor. This induces competition among vendors and puts downward pressure on prices, which benefits consumers using search engines.

4.2. *Unbiased Incomplete Coverage*

In this subsection, we analyze the case of *Unbiased Incomplete Coverage*, and compare it with the case of *Complete Coverage*. We show that *Unbiased Incomplete Coverage* compared with *Complete Coverage*, increases the expected price paid by shoppers, and decreases the expected price paid by non-shoppers.

The price distribution for the case in which the market consists of n vendors, and the price-comparison search engine has an unbiased index of size $k \leq n$, is identical to the price distribution for the case in which the price-comparison search engine has *Complete Coverage*, $k = n$, and the market consists of k vendors: $F^u(\cdot; n, k) = F^c(\cdot; k)$. For further reference, we present this observation in the next corollary.

Corollary 1. $F^u(\cdot; n, k) = F^c(\cdot; k)$. ♦

The next proposition analyzes the impact of changes in the size of the index, k, and the number of vendors, n.

Proposition 2. (i) $l^u(k) < l^u(k-1)$; (ii) $\mu^u(n,k) < \mu^u(n,k-1)$ *and* $\varepsilon^u(n,k) > \varepsilon^u(n,k-1)$; (iii) $F^u(\cdot; n, k) = F^u(\cdot; n+1, k)$. ♦

Rewrite (2) as:

$$p(1-\lambda)\left(\frac{k}{n}\right)[1-F^u(p)]^{k-1} = \frac{\lambda}{n}(1-p). \tag{4}$$

From (4), an unbiased decrease in the size of the index has two impacts. First, for indexed vendors, the decrease in the size of the index reduces the number of rivals with which a vendor has to compete to sell to shoppers from $k-1$ to $k-2$. This increases the probability that an indexed vendor will have the lowest price, $(1-F^u)^{k-1}$, which increases the *Volume of Sales effect*. The first impact leads vendors to shift probability mass from higher to lower prices. As a consequence, the price distribution shifts to the left (Figure 1). Second, the decrease in the size of the index reduces the probability that a given vendor is indexed from k/n to $(k-1)/n$, which reduces the *Volume of Sales effect*. The second impact leads vendors to raise the lower bound of the support, and to shift probability mass from lower to higher prices. As a consequence, the price distribution rotates, as shown in Figure 1. The total impact of an unbiased decrease in the size of the index is to cause the price distribution to rotate counter clock-wise.[11]

The increase in the lower bound of the support, $l^u(k) < l^u(k-1)$, raises the expected price paid by shoppers, $\mu^u(n,k) < \mu^u(n,k-1)$. However, from Remark 1(i), the average price paid in the market remains constant and equal to λ. This implies that the expected price by non-shoppers decreases, $\varepsilon^u(n,k) > \varepsilon^u(n,k-1)$.[12] Recall that vendors now charge lower prices with a higher probability. Shoppers

11. See Guimarães (1996) for a related discussion.
12. See Morgan et al. (2006) for a related discussion.

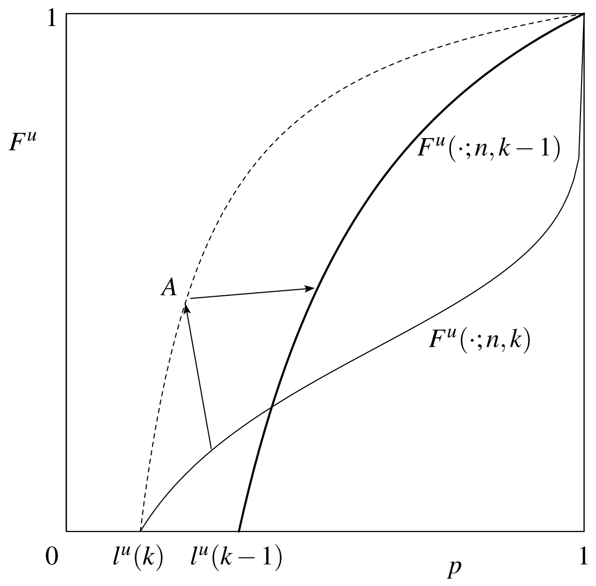


FIGURE 1

Unbiased Incomplete Coverage: A decrease in the size of the index: the first impact causes the distribution to shift from $F^c(\cdot;n)$ to A, and the second impact causes the distribution to shift from A to $F^u(\cdot;n,k)$.

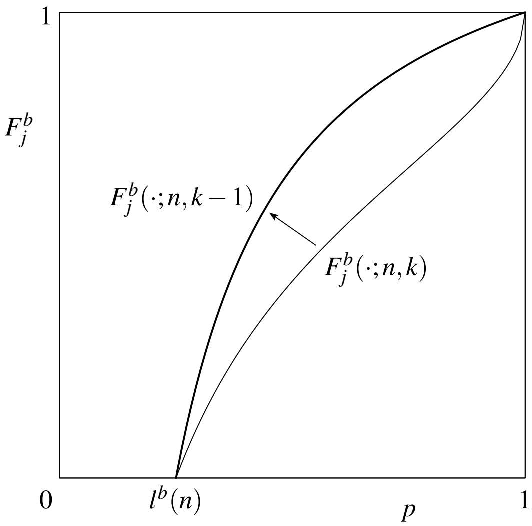
and non-shoppers have conflicting interests with respect to *Unbiased Incomplete Coverage*, as compared with *Complete Coverage*. Shoppers prefer a large to a small unbiased index, and non-shoppers prefer a small to a large unbiased index.

Under *Unbiased Incomplete Coverage*, the equilibrium price distribution does not depend on the number of vendors in the market, $F^u(\cdot;n,k) = F^u(\cdot;n+1,k)$. This result is unexpected. The probability with which a vendor is indexed, k/n, depends on the number of vendors. Besides, each vendor's share of non-shoppers, λ/n, also depends on the number of vendors. But from (4), n cancels out, and only the number of vendors whose price shoppers compare matters. Rosenthal (1980) assumed that the increase in the number of vendors is accompanied by a proportional increase in the measure of non-shoppers. In his setting, an increase in the number of vendors induces first-order stochastically dominating shifts in the price distribution, and therefore higher prices for both types of consumers. The contrast between his and this result illustrates another recurring theme of this article. In this sort of markets, the impact of entry depends critically on the way entry occurs.

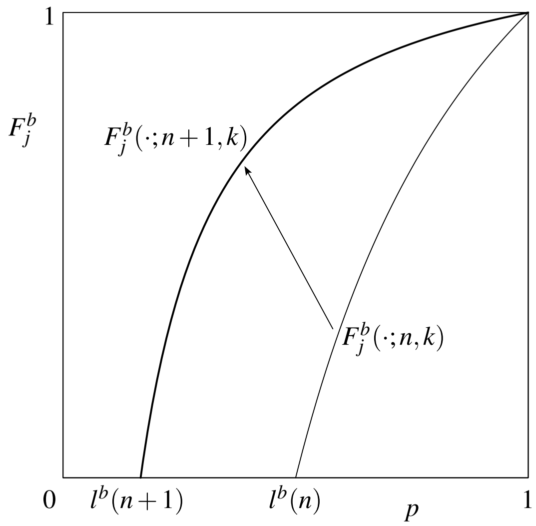
The next corollary compares the cases of *Complete Coverage* and *Unbiased Incomplete Coverage*.

Corollary 2. (i) $l^c(n) < l^u(k)$; (ii) $\mu^c(n) < \mu^u(n,k)$ *and* $\varepsilon^c(n) > \varepsilon^u(n,k)$. ♦

Given that $F^u(\cdot;n,k) = F^c(\cdot;k)$, comparing the price distributions under *Unbiased Incomplete Coverage*, $F^u(\cdot;n,k)$, and under *Complete Coverage*, $F^c(\cdot;n)$, is equivalent to comparing $F^u(\cdot;n,k)$ and $F^u(\cdot;n,n)$, i.e., is equivalent to analyzing the impact of an increase in the size of the index, under *Unbiased Incomplete Coverage*. Thus, compared with *Complete Coverage*, *Unbiased Incomplete Coverage* causes the price-comparison to rotate counter-clockwise, which increases the expected price paid by shoppers, $\mu^c(n) < \mu^u(n,k)$, and decreases the expected price paid by non-shoppers, $\varepsilon^c(n) > \varepsilon^u(n,k)$.



(a) A Decrease in the Size of the Index



(b) An Increase in the Number of Firms

FIGURE 2

Biased Incomplete Coverage: (a) A decrease in the size of the index. For $j = 1,\dots,k$ distributions $F_j^b(\cdot;n,k-1)$ are first-order stochastically dominated by distributions $F_j^b(\cdot;n,k)$. (b) An increase in the number of firms. For $j = 1,\dots,k$ distributions $F_j^b(\cdot;n+1,k)$ are first-order stochastically dominated by distributions $F_j^b(\cdot;n,k)$.

4.3. Biased Incomplete Coverage

In this subsection, we analyze the case of *Biased Incomplete Coverage*, and compare it with the other two cases. We show that *Biased Incomplete Coverage*, compared with both *Unbiased Incomplete Coverage* and with *Complete Coverage*, decreases the expected price paid by shoppers and the non-shoppers that buy from indexed vendors, and increases the expected price paid by non-shoppers that buy from non-indexed vendors.

The next proposition analyzes the impact of changes in the size of the index, k, and the number of vendors, n.

Proposition 3. (i) *For* $j = 1,\dots,k$, $F_j^b(\cdot;n,k) \leq F_j^b(\cdot;n,k-1)$*;* (ii) $\mu_j^b(n,k-1) < \mu_j^b(n,k)$ *and* $\varepsilon_j^b(n,k-1) \leq \varepsilon_j^b(n,k)$*;* (iii) *For* $j = 1,\dots,k, F_j^b(\cdot;n+1,k) \geq F_j^b(\cdot;n,k)$*;* (iv) $\mu_j^b(n+1,k) < \mu_j^b(n,k)$ *and* $\varepsilon_j^b(n+1,k) \leq \varepsilon_j^b(n,k)$*, with strict inequality for* $j = 1,\dots,k$*.* ♦

Rewrite (2) as

$$p(1-\lambda)[1-F_j^b(p)]^{k-1} = \frac{\lambda}{n}(1-p). \tag{5}$$

From (5), a decrease in the size of a biased index, k, increases the probability that an indexed vendor has the lowest price, $(1-F_j^b)^{k-1}$, which increases the *Volume of Sales effect*. This leads indexed vendors to shift probability mass from higher to lower prices. As a consequence, the distribution shifts in the first-order stochastically dominated sense, $F_j^b(p;n,k) \leq F_j^b(p;n,k-1)$ (Figure 2a). This decreases the expected price paid by shoppers, $\mu_j^b(n,k-1) < \mu_j^b(n,k)$, and by non-shoppers that buy from an indexed vendor, $\varepsilon_j^b(n,k-1) < \varepsilon_j^b(n,k)$, $j = 1,\dots,k$, and leaves unchanged the expected price paid by non-shoppers that buy from a non-indexed vendor, $\varepsilon_j^b(n,k-1) = \varepsilon_j^b(n,k)$, $j = k+1,\dots,n$.

From (5), an increase in the number of vendors in the market, n, leaving fixed the size of a biased index, k, reduces the *per Consumer Profit effect*. This leads indexed vendors to reduce the lower bound

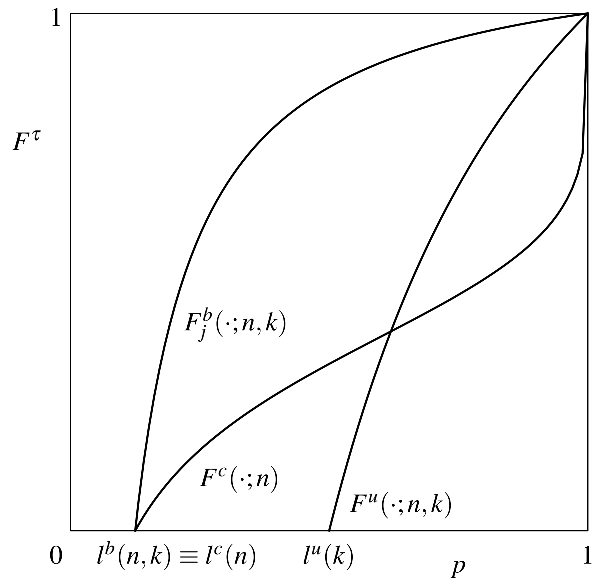


FIGURE 3

Comparison among the three types of coverage. For $j = 1,\ldots,k$ distributions $F_j^b(\cdot;n,k)$ are first-order stochastically dominated by distributions $F^c(\cdot;n)$ and $F^u(\cdot;n,k)$.

of the support, and to shift probability mass from higher to lower prices. As a consequence, the distribution shifts in the first-order stochastically dominated sense, $F_j^b(p;n+1,k) \geq F_j^b(p;n,k)$, as shown in Figure 2b. This decreases the expected price paid by shoppers, $\mu_j^b(n+1,k) < \mu_j^b(n,k)$, and by non-shoppers that buy from an indexed vendor, $\varepsilon_j^b(n+1,k) < \varepsilon_j^b(n,k)$, $j = 1,\ldots,k$, and leaves unchanged the expected price paid by non-shoppers that buy from a non-indexed vendor, $\varepsilon_j^b(n+1,k) = \varepsilon_j^b(n,k)$, $j = k+1,\ldots,n$.[13]

The next corollary compares the case of *Biased Incomplete Coverage*, with the two previous cases.

Corollary 3. (i) $l^b(n) = l^c(n)$; (ii) *For* $j = 1,\ldots,k$, $F_j^b(\cdot;n,k) \geq \max\{F^c(\cdot;n), F^u(\cdot;n,k)\}$, *and for* $j = k+1,\ldots,n$, $F_j^b(\cdot;n,k) \leq \min\{F^c(\cdot;n), F^u(\cdot;n,k)\}$; (iii) $\mu_j^b(n,k) < \min\{\mu^c(n),\ \mu^u(n,k)\}$; (iv) *For* $j = 1,\ldots,k$, $\varepsilon_j^b(n,k) < \min\{\varepsilon^c(n), \varepsilon^u(n,k)\}$, *and for* $j = k+1,\ldots,n$, $\varepsilon_j^b(n,k) > \max\{\varepsilon^c(n),\ \varepsilon^u(n,k)\}$. ♦

For indexed vendors, *Biased Incomplete Coverage* involves only the positive impact of the *Volume of Sales* effect, which leads vendors to shift probability mass from higher to lower prices. Thus, the price distribution of indexed vendors, $F_j^b(\cdot;n,k)$, is first-order stochastically dominated by price distribution under *Complete Coverage*, $F^c(\cdot;n)$, and by the price distribution under *Unbiased Incomplete Coverage* $F^u(\cdot;n,k)$, as shown in Figure 3. Shoppers buy from the cheapest indexed vendor. Thus, the expected price paid by shoppers is smaller under *Biased Incomplete Coverage*, than under either *Complete Coverage*, or *Unbiased Incomplete Coverage*. For non-shoppers that buy from an indexed vendor, the expected price is also smaller. For non-shoppers that buy from a non-indexed vendor, the expected price paid is higher.

13. As $n \to \infty$, $l^b \to 0$, and F_j^b converges weakly to $\delta(1)$, $j = 1,\ldots,k$.

TABLE 1
Treatment Design and Moments of the Theoretical Price Distributions

Treatment	Design parameters				Average prices		Minimum prices	
	τ	n	k	ϕ^τ	Mean	Std. Dev.	Mean	Std. Dev.
1	c	3	3	1.00	0.605	0.143	0.395	0.145
2	u	3	2	0.67	0.549	0.103	0.451	0.113
3	b	3	2	1.00	0.462	0.135	0.359	0.113
4	c	6	6	1.00	0.708	0.125	0.292	0.167
5	u	6	4	0.67	0.648	0.115	0.352	0.158
6	b	6	4	1.00	0.587	0.154	0.275	0.152
7	u	6	2	0.33	0.549	0.073	0.451	0.113
8	b	6	2	1.00	0.324	0.137	0.225	0.099

τ is the type of search engine: 'c' means *Complete Coverage*; 'u' means *Unbiased Incomplete Coverage*, and 'b' means *Biased Incomplete Coverage*. n is the number of firms present in each market. k is the size of the search engine index. ϕ^τ is the probability with which a firm is indexed by the search engine. The last four columns report the mean and the standard deviation of the theoretical distributions of average prices and minimum prices. Note that in treatments with biased sampling, $\tau = b$, only firms with non-null probability of being indexed are considered.

5. EXPERIMENTAL EVIDENCE

In this section, we describe the design and report the results of a laboratory experiment, developed to test the theoretical model in the presence of human subjects.

5.1. Experimental Design

The experiment was conducted at the Laboratori d'Economia Experimental, LEE, of the Universitat Jaume I, Castellón, Spain. A population of 144 subjects was recruited in advance among the students of Business Administration, and other business-related courses taught at this university.

The experiment was run under 8 treatments, each one consisting of a single session with 18 subjects.[14] Table 1 reports the design parameters of each treatment and the moments of the distributions of average price and minimum price holding under the assumptions of the theoretical model presented in the previous sections. Each session consisted of the same setup repeated 50 periods. Each period, depending on the treatment, markets of 3 or 6 subjects were randomly formed. This strangers matching protocol was adopted, in order to maintain the experimental environment as close as possible to the one-shot framework of the theoretical model. Subjects were perfectly informed of the underlying model, and their only decision variable in each period was price. See Appendix B.

Consumer behavior was simulated by the local network server.[15] There were $1,200$ consumers. [16]

14. However, it should be noted that the usual critique to single-session treatments risking confusion between treatment and session effects is not applicable here, as different comparative static predictions can be tested at different levels of each one of the design parameters. For example, differences between biased and unbiased incomplete coverage are tested by differences across treatment pairs T2-T3, T4-T5 and T6-T8. This also is true for differences between complete and incomplete unbiased coverage, studied by comparing between treatment pairs T1-T2, T4-T5 and T4-T7 and so on. Finally, the same is true for comparative statics concerning the number of firms on the index and in the industry. Thus, our econometric model accounts directly for the effect of the design parameters on the results, mitigating possible misinterpretations of session effects as treatment effects.

15. We programmed software using z-Tree (Fischbacher, 1999) in order to organize strategy submission, demand simulation, feedback, and data collection.

16. In order to avoid problems associated with the discreteness of the resulting demands, we use a larger number of

For representation and interface reasons, and in order to offer a fine grid for the strategy space, the consumers' reservation price was normalized to $1,000$ rather than to 1, Half of the consumers were assumed to be shoppers, and the other half were non-shoppers, i.e., $\lambda = 1/2$.

Under *Complete Coverage*, the search engine's index contained the prices of all subjects, i.e., $k = n = 3$ or $k = n = 6$, depending on the treatment. Under *Incomplete Coverage*, the index contained the prices of only a subset of all subjects, i.e., $k = 2$ for $n = 3$, and $k = 2$ or $k = 4$ for $n = 6$. Under *Unbiased Incomplete Coverage*, subjects were informed whether they where indexed *after* each period's prices were set. Under *Biased Incomplete Coverage*, subjects were informed on the composition of the index *before* prices were set. Both in the case of *Unbiased* or *Biased Coverage*, after each period's prices were set, subjects were informed on own and rival prices, as well as own quantities sold and profits earned.[17]

In order to make the earnings of each period equally interesting, subjects' monetary rewards were calculated from the cumulative earnings over 10 randomly selected periods. Individual rewards ranged between 15 € and 50 €. This made the experiment worth participating in, and made trying to have the highest payoff worthwhile.

5.2. Testable Hypothesis

Denote by ε_t, the expected price in treatment t; by ε_t^{in} the expected price of indexed firms in treatment t; by ε_t^{ni} the expected price of non-indexed firms in treatment t; and by μ_t the expected minimum price in treatment t, where $t = 1,\ldots,8$.

Regarding the most well known result of this framework which has also been explicitly tested by Morgan et al. (2006), we test the basic or consistency hypothesis:

HC: *Under* Complete Coverage*, an increase in the number of vendors:* (i) *decreases the expected minimum price:* $\mu_4 < \mu_1$*;* (ii) *increases the average price:* $\varepsilon_4 > \varepsilon_1$.

Regarding *Unbiased Incomplete Coverage* we test:

HU1: *Under* Unbiased Incomplete Coverage*, a decrease in the size of the index:* (i) *increases the expected minimum price:* $\mu_4 < \mu_5 < \mu_7$ *and* $\mu_1 < \mu_2$*;* (ii) *decreases the expected price:* $\varepsilon_4 > \varepsilon_5 > \varepsilon_7$ *and* $\varepsilon_1 > \varepsilon_2$.

HU2: *Under* Unbiased Incomplete Coverage*, the equilibrium price distribution is independent of the number of vendors present in the market:* $\mu_2 = \mu_7$ *and* $\varepsilon_2 = \varepsilon_7$.

Regarding *Biased Incomplete Coverage* we test:

HB1: *Under* Biased Incomplete Coverage*, a decrease in the size of the index:* (i) *decreases the expected minimum price:* $\mu_4 > \mu_6 > \mu_8$ *and* $\mu_1 > \mu_3$*;* (ii) *decreases the expected price of indexed vendors:* $\varepsilon_4 > \varepsilon_6^{in} > \varepsilon_8^{in}$ *and* $\varepsilon_1 > \varepsilon_3^{in}$*;* (iii) *leaves unchanged the expected price of non-indexed vendors:* $\varepsilon_6^{ni} = \varepsilon_8^{ni} = 1$ *and* $\varepsilon_3^{ni} = 1$.

HB2: *Under* Biased Incomplete Coverage*, an increase in the number of vendors present in the market:* (i) *decreases the expected minimum price:* $\mu_3 > \mu_8$*;* (ii) *decreases the expected price of indexed*

consumers than that used by Morgan et al. (2006).

17. Subjects observe their rivals' prices for two reasons. First, for realism. Second, because it helps subjects infer the types of strategies that are adopted and abandoned by the rest of the market, which speeds convergence.

TABLE 2

Descriptive Statistics

Treatment	Average prices			Minimum prices			Obs.
	Mean	Std. Dev.	t-test	Mean	Std. Dev.	t-test	
1	0.535	0.158	-5.90^{**}	0.294	0.103	-13.14^{**}	180
2	0.760	0.109	25.84^{**}	0.635	0.157	15.68^{**}	180
3	0.637	0.211	11.13^{**}	0.475	0.247	6.35^{**}	180
4	0.700	0.136	-0.58	0.193	0.161	9.98^{**}	90
5	0.592	0.149	-3.52^{**}	0.204	0.197	-7.17^{**}	90
6	0.591	0.177	0.19	0.208	0.176	-3.64^{**}	90
7	0.758	0.088	22.52^{**}	0.644	0.258	7.10^{**}	90
8	0.416	0.214	4.06^{**}	0.269	0.187	2.21^{*}	90

Means and standard deviations of the empirical distributions of average and minimum prices and two-sided t-tests of equality of price distribution means to their theoretical values. Test statistics rejecting the null hypothesis at the 5% and 1% significance levels are marked with * and ** respectively. The means of the theoretical distribution of average prices and minimum prices are reported in Table 1. In the biased treatments, 3, 6 and 8, only indexed firms are considered.

vendors: $\varepsilon_3^{in} > \varepsilon_8^{in}$.

We also test the general hypothesis:

HG: (i) *The expected minimum price is smaller under* Biased Incomplete Coverage*, than under* Unbiased Incomplete Coverage*:* $\mu_5 > \mu_6$ *and* $\mu_7 > \mu_8$ *and* $\mu_2 > \mu_3$*;* (ii) *The expected price of indexed vendors is smaller under* Biased Incomplete Coverage*, than under* Unbiased Incomplete Coverage*:* $\varepsilon_5 > \varepsilon_6^{in}$*, and* $\varepsilon_7 > \varepsilon_8^{in}$*, and* $\varepsilon_2 > \varepsilon_3^{in}$*.*

5.3. *Experimental Results*

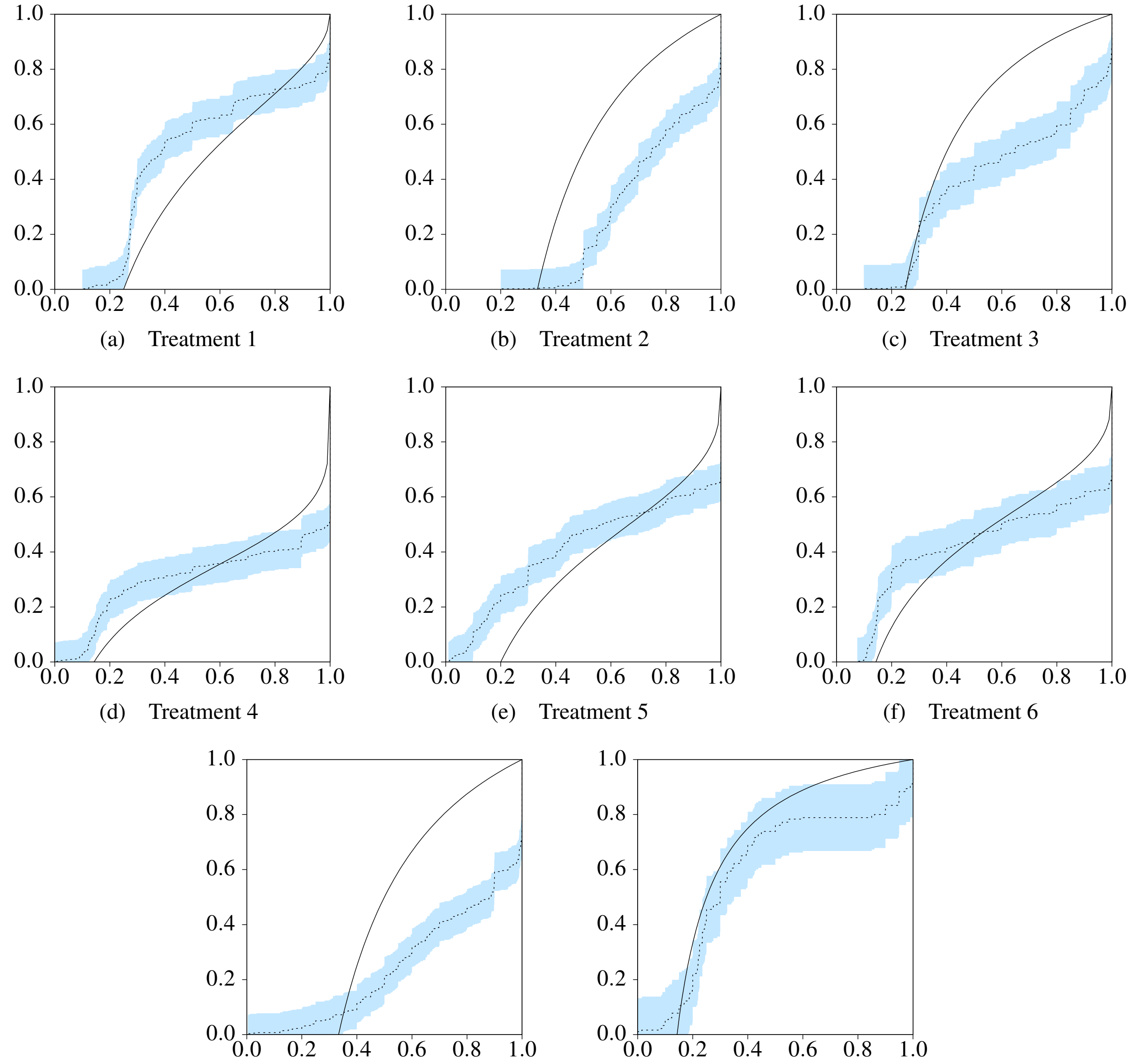
Table 2 summarizes the information and descriptive statistics regarding all treatments. The 20 initial periods were dropped from the data sets to eliminate learning dynamics and guarantee that observations had reached the necessary stability.[18] Seven conclusions emerge from the experimental observations.

Observation 1. *There is a systematic deviation of the empirical results from the theoretical results.* ♦

This conclusion can be reached through at least two alternative ways. First, the inspection of the t-tests in Table 2 shows that with the exception of the average price for treatments 4 and 6, all estimated means of average and minimum prices are significantly different from their theoretical values.

Second, this conclusion can also be gleaned from the inspection of Figure 4, which compares the empirical and theoretical distributions of prices fixed by subjects that had a non null probability of being indexed by the search engine. In Figure 4, each empirical price distribution is surrounded with a

18. We focus here on the comparative static results that can be tested by our data. The dynamic properties of individual behavior are studied in detail by García-Gallego et al. (2006), applying specific theoretical results by Janssen and Moraga-González (2004), Benaïm et al. (2005) and Hopkins and Seymour (2002) on the stability of learning dynamics in the framework of search models.



(a) Treatment 1 (b) Treatment 2 (c) Treatment 3

(d) Treatment 4 (e) Treatment 5 (f) Treatment 6

(g) Treatment 7 (h) Treatment 8

FIGURE 4

Theoretical (solid lines) and Empirical (dashed line) Price Distributions. Empirical distributions are surrounded with a confidence region built from the 1% critical values of the Kolmogorov-Smirnov test.

confidence region built from the 1% critical values of the Kolmogorov-Smirnov test. The Kolmogorov-Smirnov test statistic is the maximum of the absolute value of the difference between the two distributions under comparison. Therefore, as none of the theoretical distributions completely lies in the confidence regions of Figure 4, the data clearly rejects the equality of the theoretical and the empirical distributions in all treatments.

The empirical distributions rotate clock-wise compared to the theoretical ones. In the case of treatments 2 and 3, the empirical distribution "almost" first-order stochastically dominates the theoretical distribution. This rotation indicates the presence of more density on both tails of distributions of observed prices, than the theoretical model would have predicted. On one hand, in some treatments a large number of observations lie below of the infimum of the support of the theoretical distributions, l^{τ}, e.g., in treatments 1, 4, 5 and 6. On the other hand, a large number of observations are at the maximum price, $p_j = 1$. This behavior is specially pronounced in treatments 4, 5, 6, and 7. In line with the way in which the empirical distributions rotate, most of the empirical price distributions have higher standard

TABLE 3

t-tests of Equality of Means of Average Prices

	Average prices			Minimum prices			
	H_0	H_1	t-test	H_0	H_1	t-test	d. f.
HC	$\varepsilon_1 = \varepsilon_4$	$\varepsilon_1 < \varepsilon_4$	-8.44^{**}	$\mu_4 = \mu_1$	$\mu_4 < \mu_1$	-6.23^{**}	268
HU1	$\varepsilon_5 = \varepsilon_4$	$\varepsilon_5 < \varepsilon_4$	-5.06^{**}	$\mu_4 = \mu_5$	$\mu_4 < \mu_5$	-0.39	178
	$\varepsilon_7 = \varepsilon_4$	$\varepsilon_7 < \varepsilon_4$	3.42	$\mu_4 = \mu_7$	$\mu_4 < \mu_7$	-14.06^{**}	178
	$\varepsilon_7 = \varepsilon_5$	$\varepsilon_7 < \varepsilon_5$	9.11	$\mu_5 = \mu_7$	$\mu_5 < \mu_7$	-12.87^{**}	178
	$\varepsilon_2 = \varepsilon_1$	$\varepsilon_2 < \varepsilon_1$	15.68	$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	-24.26^{**}	358
HU2	$\varepsilon_2 = \varepsilon_7$	$\varepsilon_2 \neq \varepsilon_7$	0.11	$\mu_2 = \mu_7$	$\mu_2 \neq \mu_7$	-0.35	268
HB1	$\varepsilon_6^{in} = \varepsilon_4$	$\varepsilon_6^{in} < \varepsilon_4$	-4.64^{**}	$\mu_6 = \mu_4$	$\mu_6 < \mu_4$	0.59	178
	$\varepsilon_8^{in} = \varepsilon_4$	$\varepsilon_8^{in} < \varepsilon_4$	-10.60^{**}	$\mu_8 = \mu_4$	$\mu_8 < \mu_4$	2.91	178
	$\varepsilon_8^{in} = \varepsilon_6^{in}$	$\varepsilon_8^{in} < \varepsilon_6^{in}$	-5.96^{**}	$\mu_8 = \mu_6$	$\mu_8 < \mu_6$	2.25	178
	$\varepsilon_3^{in} = \varepsilon_1$	$\varepsilon_3^{in} < \varepsilon_1$	5.19	$\mu_3 = \mu_1$	$\mu_3 < \mu_1$	9.09	358
HB2	$\varepsilon_8^{in} = \varepsilon_3^{in}$	$\varepsilon_8^{in} < \varepsilon_3^{in}$	-8.07^{**}	$\mu_8 = \mu_3$	$\mu_8 < \mu_3$	-7.00^{**}	268
HG	$\varepsilon_6^{in} = \varepsilon_5$	$\varepsilon_6^{in} < \varepsilon_5$	-0.07	$\mu_6 = \mu_5$	$\mu_6 < \mu_5$	0.15	178
	$\varepsilon_8^{in} = \varepsilon_7$	$\varepsilon_8^{in} < \varepsilon_7$	-14.01^{**}	$\mu_8 = \mu_7$	$\mu_8 < \mu_7$	-11.17^{**}	178
	$\varepsilon_3^{in} = \varepsilon_2$	$\varepsilon_3^{in} < \varepsilon_2$	-6.93^{**}	$\mu_3 = \mu_2$	$\mu_3 < \mu_2$	-7.30^{**}	358

Except for **HU2**, one-sided t-tests with 'd.f' degrees of freedom. The testable implications of section 5.2 correspond to the alternative hypothesis of these tests. For **HU2**, two-sided t-test with 'd.f' degrees of freedom. **HU2** correspond to the null hypothesis of this test. Test statistics rejecting the null hypothesis at the 5% and 1% significance levels are marked with * and ** respectively.

deviations than the corresponding theoretical ones. See Tables 1 and 2.

We also found a difference between the expected and the observed behavior of subjects that knew beforehand that they would not be indexed under *Biased Incomplete Coverage*. In treatments 6 and 8, 8% of the observed prices of these subjects were different from $p_j = 1$. We suspect that most of these observations were mistakes, as many of these subjects only deviated from the degenerate equilibrium strategy once or twice. But in treatment 3, nearly 30% of the prices of these subjects were different from $p_j = 1$, and four of these individuals always choose prices lower than $p_j = 1$. Clearly, the experimental data does not support hypothesis **HB1 (iii)**.

Observation 2. *The data supports the model's predictions regarding changes in the number of firms present in the market.* ♦

From Table 3, it follows that: (i) $\mu_2 = \mu_7$ and $\varepsilon_2 = \varepsilon_7$, (ii) $\mu_3 > \mu_8$ and $\varepsilon_3^{in} > \varepsilon_8^{in}$, (iii) $\mu_1 > \mu_4$ and $\varepsilon_1 < \varepsilon_4$. This implies that the data supports hypotheses: **HC**, **HU2**, and **HB2**.

Consider in particular the consistency hypothesis, **HC**: under *Complete Coverage*, an increase in the number of vendors increases the average price and decreases the expected minimum price. This conclusion can also be gleaned from the inspection of Figure 5. This non-trivial result was also obtained by Morgan et al. (2006).

Observation 3. *The data supports the model's predictions regarding the comparison between Un-*

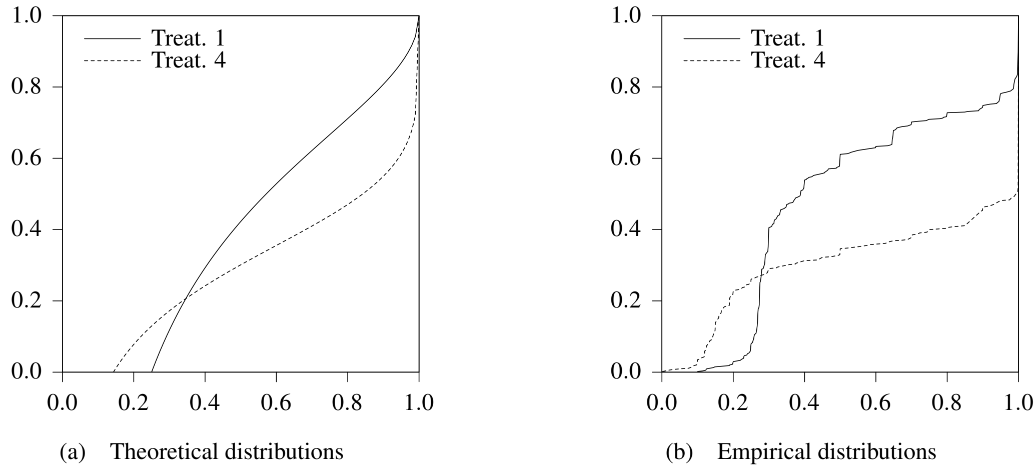


FIGURE 5

Comparison of Price Distributions: Complete Coverage.

biased Incomplete Coverage and Biased Incomplete Coverage. ♦

From Table 3, it follows that: (i) $\mu_3 < \mu_2$ and $\varepsilon_3^{in} < \varepsilon_2$, (ii) $\mu_6 = \mu_5$ and $\varepsilon_6^{in} = \varepsilon_5$, (iii) $\mu_8 < \mu_7$ and $\varepsilon_8^{in} < \varepsilon_7$. The data support hypothesis **HG**: the average and the average minimum prices are weakly lower under *Biased Incomplete Coverage* than under *Unbiased Incomplete Coverage*. Both types of consumers, shoppers and non-shoppers, are better off if the index is biased than if it is unbiased. The same conclusion can be gleaned from the inspection of Figure 6.

Observation 4. *The data does not support the model's predictions regarding changes in the size of the index.* ♦

From Table 3, it follows that: (i) $\mu_1 < \mu_2$ and $\varepsilon_1 < \varepsilon_2$, (ii) $\mu_4 < \mu_7$ and $\varepsilon_4 < \varepsilon_7$, (iii) $\mu_5 < \mu_7$ and $\varepsilon_5 < \varepsilon_7$, (iv) $\mu_4 = \mu_5$ and $\varepsilon_4 > \varepsilon_5$. This implies that the data does not support hypothesis **HU1**.

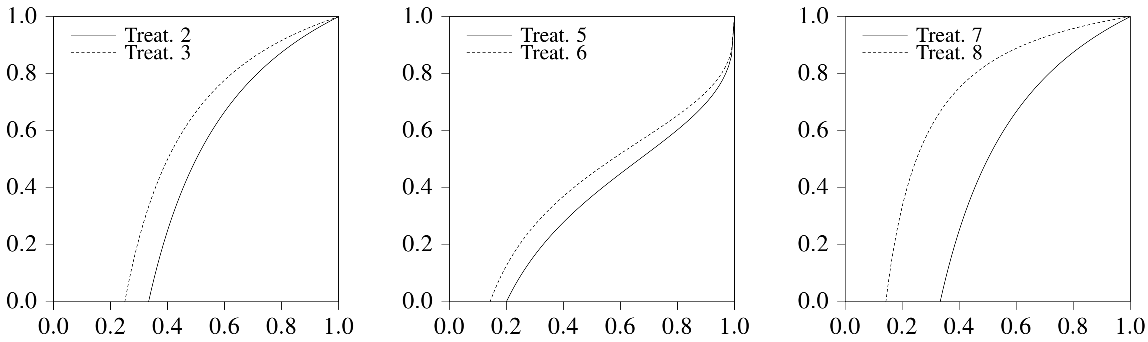
Also from Table 3, it follows that: (i) $\mu_1 < \mu_3$ and $\varepsilon_1 < \varepsilon_3^{in}$, (ii) $\mu_4 < \mu_8$ and $\varepsilon_4 > \varepsilon_8^{in}$, (iii) $\mu_6 < \mu_8$ and $\varepsilon_6^{in} > \varepsilon_8^{in}$, (iv) $\mu_4 = \mu_6$ and $\varepsilon_4 > \varepsilon_6^{in}$. This implies that the data does not support hypothesis **HB1**, either.

Observation 5. *The data does not support the model's predictions regarding the comparison between Complete Coverage and Incomplete Coverage.* ♦

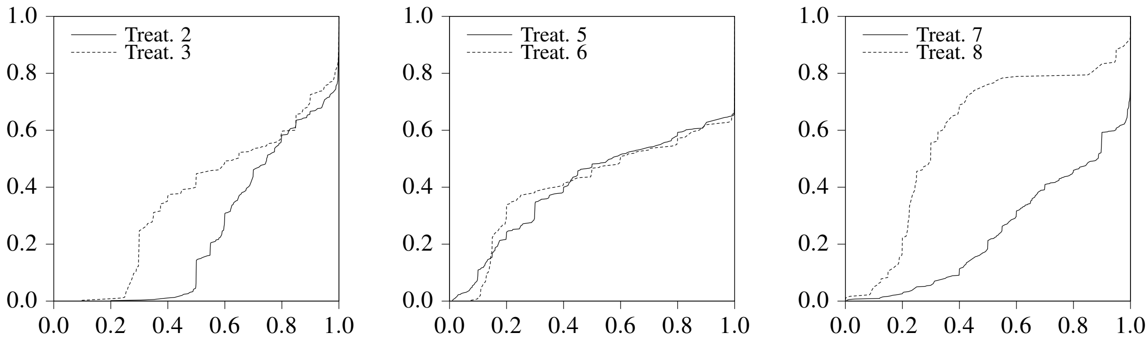
With respect to the comparison between *Complete Coverage* and *Unbiased Incomplete Coverage*, from Table 3, it follows that: (i) $\mu_4 = \mu_5 < \mu_7$ and $\mu_1 < \mu_2$; (ii) $\varepsilon_5 < \varepsilon_4 < \varepsilon_7$ and $\varepsilon_1 < \varepsilon_2$. Only the comparison of minimum prices is weakly compatible with the model's predictions. The data fails to support the predicted comparison between *Complete Coverage* and *Unbiased Incomplete Coverage*, i.e., **HU1**.

With respect to the comparison between *Complete Coverage* and *Biased Incomplete Coverage*, from Table 3, it follows that: (i) $\mu_4 = \mu_6 < \mu_8$ and $\mu_1 < \mu_3$; (ii) $\varepsilon_8^{in} = \varepsilon_6^{in} < \varepsilon_4$ and $\varepsilon_1 < \varepsilon_3^{in}$. The data fails to support the predicted comparison between *Complete Coverage* and *Biased Incomplete Coverage*, i.e., **HB1**.

Observation 6. *The average minimum price is weakly lower under Complete Coverage than under*



(a) Theoretical distributions



(b) Empirical distributions

FIGURE 6

Comparison of Price Distributions: Incomplete Coverage

Biased Incomplete Coverage. ♦

From Table 3 it follows that: (i) $\mu_1 < \mu_3$, (ii) $\mu_4 = \mu_6$, and (iii) $\mu_4 < \mu_8$. Jointly with Observation 3, this implies that shoppers are better off under *Complete Coverage* than under *Incomplete Coverage*.

Observation 7. *Given the type of bias, ratio k/n, and Incomplete Coverage, an increase in the number of firms in the market leads to a lower average price, and a lower average minimum price.* ♦

From Table 3 it follows that: (i) $\mu_5 < \mu_2$ and $\varepsilon_5 < \varepsilon_2$, and (ii) $\mu_6 < \mu_3$ and $\varepsilon_6 < \varepsilon_3$. This observation agrees with the empirical findings of Baye et al. (2003).

6. CONCLUSIONS

In this article we developed a theoretical model of the effects on prices of incompleteness and biasedness of price sampling processes in search models. We tested the model's predictions in a laboratory experiment specifically designed for that purpose.

The theoretical model warns us on possible counter intuitive effects of unbiased coverage of the market on observed prices, e.g., an increase in the number of firms whose prices are included in the search engine may not imply globally lower expected prices. However, the experimental results contradict some of the theoretical predictions, especially, predictions about incomplete coverage.

Generally speaking, it is a rather positive finding of our research that the basic prediction of the Varian

(1980) model concerning the comparative statics with respect to the number of firms is qualitatively confirmed by our data. This positive finding, reinforces the findings by Morgan et al. (2006), which also confirm the aforementioned prediction.[19] Contrary to this positive finding, the proposed extensions of the Varian (1980) model accounting for incomplete price sampling processes receive little if any support by our results. Future research should aim at investigating the behavioral sources of this failure.

APPENDIX A. PROOFS

Proof of Lemma 1. For $\tau = b$ and $j = k+1,\dots,n$ the proofs are obvious, so consider: (a) $\tau = b$ and $j = 1,\dots,k$, and (b) $\tau = c,u$.

(i). For any j, any price $p_j < l_j^\tau$ or $p_j > 1$ is strictly dominated by $p_j = 1$;

(ii). Suppose not, i.e., suppose that F_j^τ has a mass point at price p. Let $\varepsilon > 0$ be arbitrarily small and such that no mass point exists at price $p-\varepsilon$. The expected profits of firm j are:

$$\begin{aligned}\Pi_j(p-\varepsilon) =&(p-\varepsilon)\frac{\lambda}{n} + (p-\varepsilon)(1-\lambda)\phi_j^\tau \operatorname{Prob}[p-\varepsilon < \hat{p}_{-j}] \\ &+ (p-\varepsilon)(1-\lambda)\phi_j^\tau \operatorname{Prob}[p-\varepsilon \le p = \hat{p}_{-j}],\end{aligned}$$

and

$$\Pi_j(p) = p\frac{\lambda}{n} + p(1-\lambda)\phi_j^\tau \operatorname{Prob}[p < \hat{p}_{-j}] + p\frac{(1-\lambda)\phi_j^\tau}{\hat{m}_{-j}} \operatorname{Prob}[p = \hat{p}_{-i}^j].$$

Subtracting the second expression from the first and taking the limit as ε approached zero, one obtains

$$\lim_{\varepsilon\to 0}[\Pi_j(p-\varepsilon) - \Pi_j(p)] = p\alpha(1-\lambda)\phi_j^\tau \left(\frac{\hat{m}_{-j}-1}{\hat{m}_{-j}}\right) \operatorname{Prob}[p = \hat{p}_{-j}] > 0.$$

Hence, vendor j would increase profit by shifting mass from p to an ε neighborhood below p. But this implies that it cannot be an equilibrium strategy to maintain a mass point at p;

(iii). Suppose not, i.e., suppose $\bar{p}_j < 1$. Then

$$\Pi_j(\bar{p}_j) = \bar{p}_j\frac{\lambda}{n} + \bar{p}_j(1-\lambda)\phi_j^\tau[1-F(\bar{p}_j)]^{k-1} = \bar{p}_j\frac{\lambda}{n},$$

since from (ii) there are no mass points at $\bar{p}_j\lambda/n$. However, the payoff from setting a price equal to 1 is $\lambda/n > \bar{p}_j\lambda/n$;

(iv). Follows from (ii) and (iii);

(v). Parts (ii) and (iv) imply that $\underline{p}_j\lambda/n + \underline{p}_j(1-\lambda)\phi_j^\tau = \Pi_j(\underline{p}_j) = \lambda/n$. Hence $\underline{p}_j = l_j^\tau$;

(vi). Suppose not, i.e., suppose there is an interval $[p_l, p_h]$ satisfying $l_j^\tau \le p_l < p_h \le 1$ such that $F(p_l) = F(p_h)$. Suppose also that p_l is the infimum of all prices p, $l_j^\tau \le p \le 1$. Then p_l is in the support of $F(\cdot)$ and, from (ii) $\Pi_j^\star = \Pi_j(p_l) = p_l\lambda/n + p_l(1-\lambda)\phi_j^\tau[1-F(p_l)]^{k-1} < p_h\lambda/n + p_h(1-\lambda)\phi_j^\tau[1-F(p_h)]^{k-1} = \Pi_j(p_h)$, a contradiction. □

19. More recent experimental evidence on this is provided by Orzen (2005)

Proof of Proposition 1. We show constructively that equilibrium exists. Alternatively, existence follows from theorem 5 of Dasgupta and Maskin (1986). (i) Use Lemma 1(iv) to set $\Pi_j(p) = p\lambda/n + p(1-\lambda)\phi_j^{\tau}[1-F(p)]^{k-1} = \lambda/n$. Solving for $F(p)$ the result follows; (ii) Obvious. □

Proof of Remark 1. (i) Follows from the fact that all firms are indifferent between any equilibrium price and the monopoly price. (ii) Follows directly from the definition of $\mu^c = l^c + \int_{l^c}^{1}(1-F^c)^n dp$ and $\varepsilon^c = l^c + \int_{l^c}^{1}(1-F^c)dp$. □

Theorem 1. (i) $\varepsilon^c(n) < \varepsilon^c(n+1)$; (ii) $\mu^c(n) > \mu^c(n+1)$. ♦

Proof of Theorem 1. (i) See Morgan et al. (2006); (ii) Follows from (i) and Remark 1(i). □

Proof of Corollary 1. Obvious. □

Proof of Proposition 2. (i) Obvious; (ii) Follows from Corollary 1 and the Theorem 1; (iii) Obvious. □

Proof of Corollary 2. (i) Obvious; (ii) Follows from Corollary 1 and the Theorem 1. □

Proof of Proposition 3. (i) Obvious; (ii) Follows from (i); (iii) Obvious; (iv) Follows from (iii). □

Proof of Corollary 3. (i) Obvious; (ii) Obvious; (iii) Follows from (ii); (iv) Follows from (ii). □

APPENDIX B. INSTRUCTIONS OF THE EXPERIMENT (TRANSLATED FROM SPANISH)

- The purpose of this experiment is to study how subjects take decisions in specific economic contexts. This project has received financial support by public funds. Your decision making in this session is going to be of great importance for the success of this research. At the end of the session you will receive a quantity of money in cash which will depend on your performance during the session.
- The environment in which the experiment takes place is an industry. This industry has the following characteristics:

  (a) a price comparison search engine like the ones on the Internet,
  (b) 3 firms, (*Treatments 4–8:* 6 firms),
  (c) 1200 consumers.

  Each firm in the industry produces a homogeneous product, and this product is the same for all firms.
- Transactions will take place in *UMEX* (our lab's Experimental Monetary Units).
- This session will consist of 50 rounds.
- You are one of the 3 firms (*Treatments 4–8:* 6 firms) in the industry. Your production costs are zero. Therefore, your profits are equal to your revenue.
- Each round, you and the rest of the firms in the industry have to decide the price at which you want to sell the product. Price is your only decision variable.

- (*Treatments 1 and 4*) Each period, a *Price Search Engine* lists the prices of *all* firms in the industry.
- (*Treatments 2, 5 and 7*) Each period, a *Price Search Engine* lists the prices of 2 firms (*Treatment 5:* 4 firms) in the industry. More precisely, each round, the price comparison search engine randomly chooses 2 firms (*Treatment 5:* 4 firms), whose price will be included in its price list. The identity of the firms whose price will be included in the list of the price search engine, will be announced publicly to the members of the industry *after* the firms' prices are posted.
- (*Treatments 3, 6 and 8*) Each period, a *Price Search Engine* lists the prices of 2 firms (*Treatment 6:* 4 firms) in the industry. More precisely, each round, the price comparison search engine randomly chooses 2 firms (*Treatment 6:* 4 firms), whose price will be included in its price list. The identity of the firms whose price will be included in the list of the price search engine, will be announced publicly to the members of the industry *before* the firms' prices are posted.
- Each consumer wants to buy one unit of the product per round. The maximum willingness to pay of each consumer for a unit of the product is 1000 UMEX. That is, if the price you fix is higher than 1000 UMEX, nobody will buy from you.
- There are two types of consumers. Half of them, i.e., 600 consumers, will read the list of price created by the search engine. The other half do not actually read the list of prices of the search engine (maybe because they are not able to do so).
- The consumers who read the price list of the search engine will buy, each period, from the firm whose price for that period is the lowest *among all prices included in the price list*, if such price does not exceed 1000 UMEX. In case of a "tie" (i.e., several firms fix the same minimum price) the consumers are distributed evenly among the firms with the same minimum price.
- The consumers who do not read the search engine's price list will buy "randomly" from any vendor, so that this group of consumers will be distributed evenly among all firms in the industry.
- In each round, 3 firms (*Treatments 4–8:* 6 firms) forming (together with you) the same industry, will be randomly drawn among the 18 participants of this session. Therefore, the probability of competing with the same 2 firms (*Treatments 4–8:* 5 firms) in 2 different periods is very low (less than 10%).
- Once the participants have been assigned to the industries, you must set your price. The master program in the computer will simulate the consumers' reactions. At the end of each round, you will see on your screen the information about your own sales, your earnings and the prices fixed by your competitors in the market.
- At the end of the session you will be paid in cash. Your reward will be determined taking into account the earnings you accumulate over 10 (randomly selected) out of the total 50 periods. The exchange rate will be: 1000000 UMEX = 10 €.

Thank you very much for your participation. Good luck!

REFERENCES

BAYE, M., GATTI, R., KATTUMAN, P. and MORGAN, J. (2003), "Did the Euro Foster Online Price Competition? Evidence from an International Price Comparison Site", Mimeo, Indiana University.

BAYE, M., KOVENOCK, D. and DEVRIES, C. (1992), "It Takes Two to Tango: Equilibria in a Model of Sales", *Games and Economic Behavior*, **4**, 493–510.

BAYE, M. and MORGAN, J. (2001), "Information Gatekeepers on the Internet and the Competitiveness of Homogeneous Product Markets", *American Economic Review*, **91** (3), 454–474.

BENAÏM, M., HOFBAUER, J. and HOPKINS, E. (2005), "Learning in Games with Unstable Equilibria", ESE Discussion Papers 135, Edinburgh School of Economics.

BRADLOW, E. and SCHMITTLEIN, D. (1999), "The Little Engine that Could: Modeling the Performance of World Wide Web Search Engines", *Marketing Science*, **19** (1), 43–62.

BRYNJOLFSSON, E. and SMITH, M. (2000), "Frictionless Commerce? A Comparison of Internet and Conventional Retailers", *Management Science*, **46** (4), 563–585.

BURDETT, K. and JUDD, K. (1983), "Equilibrium Price Dispersion", *Econometrica*, **51**, 955–969.

CASON, T. N. and FRIEDMAN, D. (2003), "Buyer Search and Price Dispersion: a Laboratory Study", *Journal of Economic Theory*, **112**, 232–260.

DASGUPTA, P. and MASKIN, E. (1986), "The Existence of Equilibrium in Discontinuous Economic Games, I: Theory", *Review of Economic Studies*, **53**, 1–26.

DINLERSOZ, E. and PEREIRA, P. (2007), "On the Diffusion of Electronic Commerce", *International Journal of Industrial Organization*, forthcoming.

FISCHBACHER, U. (1999), "z-Tree: Zurich Toolbox for Readymade Economic Experiments. Experimenter's Manual", Working Paper 21, Institute for Empirical Research in Economics, University of Zurich.

GARCÍA-GALLEGO, A., GEORGANTZÍS, N., JARAMILLO-GUTIÉRREZ, A., PEREIRA, P. and PERNÍAS, J. C. (2006), "On the Evolution of Monopoly Pricing in Internet-Assisted Search Markets", Mimeo, Universitat Jaume I.

GUIMARÃES, P. (1996), "Search Intensity in Oligopoly", *Journal of Industrial Economics*, **44**, 415–426.

HOERNIG, S. and PEREIRA, P. (2007), "Search and Destroy: Competition between Search Engines and Consumer Prices", Mimeo, Portuguese Competition Authority, Lisboa.

HOPKINS, E. and SEYMOUR, R. (2002), "The Stability of Price Dispersion under Seller and Consumer Learning", *International Economic Review*, **43**, 1157–1190.

IYER, G. and PAZGAL, A. (2000), "Internet Shopping Agents: Virtual Co-Location and Competition", *Marketing Science*, **22**, 85–106.

JANSSEN, M. C. W. and MORAGA-GONZÁLEZ, J. L. (2004), "Strategic Pricing, Consumer Search and the Number of Firms", *Review of Economic Studies*, **71**, 1089–1118.

LAWRENCE, S. and GILES, C. (1998), "Searching the World Wide Web", *Science*, **280** (3), 98–100.

LAWRENCE, S. and GILES, C. (1999), "Accessibility of Information on the Web", *Nature*, **400**, 107–109.

MONTGOMERY, A., HOSANAGAR, K., KRISHNAN, R. and CLAY, K. (2004), "Designing a Better Shopbot", *Management Science*, **50** (2), 189–206.

MORGAN, J., ORZEN, H. and SEFTON, M. (2006), "An Experimental Study of Price Dispersion", *Games and Economic Behavior*, **54**, 134–138.

ORZEN, H. (2005), "Concentration and Competition: An Experiment", Discussion Paper 2005-2006, CeDEx.

PEREIRA, P. (2005), "Do Lower Search Costs Reduce Prices and Price Dispersion?", *Information Economics and Policy*, **17** (1), 61–72.

ROSENTHAL, R. (1980), "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price", *Econometrica*, **48**, 1575–1579.

SAMUELSON, L. and ZHANG, J. (1992), "Search Costs and Prices", *Economics Letters*, **38**, 55–60.

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

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