NET Institute*

www.NETinst.org

Working Paper #04-03

October 2004

Risk Attitudes and Internet Search Engines: Theory and Experimental Evidence

Aurora García-Gallego, Nikolaos Georgantzís, Pedro Pereira and José C. Pernías-Cerrillo

Universitat Jaume I (Castellón, Spain), Universitat Jaume I, Autoridade da Concorrência (Portugal), and Universitat Jaume I

* The Networks, Electronic Commerce, and Telecommunications ("NET") Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, "virtual networks" comprised of computers that share the same technical standard or operating system, and on network issues in general.

Risk Attitudes and Internet Search Engines: Theory and Experimental Evidence*

Aurora García-Gallego†

Nikolaos Georgantzís‡

Pedro Pereira§

José C. Pernías-Cerrillo¶

30th September 2004

Abstract

This paper analyzes the impact on consumer prices of the size and biases of price comparison search engines. We develop several theoretical predictions, in the context of a model related to Burdett and Judd (1983) and Varian (1980), and test them experimentally. The data supports the model's predictions regarding the impact of the number of firms, and the type of bias of the search engine. The data does not support the model's predictions regarding the impact of the size of the search engine. We identified several data patterns, and developed an econometric model for the price distributions. Variables accounting for risk attitudes improved significantly the explanatory power of the econometric model.

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

JEL Codes: D43, D83, L13.

*Financial support from the NET Institute (http://www.NETinst.org) is gratefully acknowledged.

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

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

§Autoridade da Concorrência, R. Laura Alves, nº 4, 7º, 1050-138 Lisboa (Portugal); phone: 21 7802498; fax: 217802471; e-mail: jppereira@oninetspeed.pt

¶Economics Department, Universitat Jaume I, Castellón (Spain); phone: 34 964728610; e-mail: pernias@eco.uji.es

1 Introduction

1.1 Preliminary Thoughts

In this subsection we layout the motivation of this research.

From the consumers' perspective, one of the more 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.[1] The consumers' enhanced ability of comparing prices would discipline vendors, and put downward pressure on prices.[2]

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 Pereira (2004b) 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 (Pereira (2004b)).

In this paper 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] A *Search Engine* is a program that accesses and reads Internet pages, stores the results, and returns lists of pages, which match keywords in a query. It consists 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. *Price Comparison Search Engines*, also known as shopping agents or shopping robots, are a class of search engines, which crawl commercial Internet sites. In addition to addresses from vendors, they also collect and display other information like prices, or return policies. There are also price comparison directories.

[2] The search literature has no simple prediction about the relation between search costs, price levels, or price dispersion (Pereira (2004a); Samuelson and Zhang (1992)).

1.2 Overview of the Paper

In this subsection we give an overview of the model and the paper's main results.

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)** a 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 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. Pereira (2004b) analyzes these questions. In this paper we take these decisions as given, and focus on the implications for the pricing behavior of vendors.

The theoretical analysis makes several specific 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 counterintuitive 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 hypotheses. This included a parallel experiment, intended to capture two different aspects of risk attitudes: **(i)** the subjects' degree of risk aversion, and **(ii)** whether subjects accepted more risk in exchange for a higher return.

The experimental results confirmed 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 data patterns, such as that prices are lower under biased incomplete coverage than under unbiased incomplete coverage. We also developed an econometric model of the empirical price distributions. This lead us to a very interesting empirical finding concerning risk attitudes. Introducing variables that account for the subjects' risk attitudes improved significantly the explanatory power of the econometric model. The implication of this finding for future research is clear. If behavioral factors are systematically important in these markets, then they should be incorporated explicitly in the modelling assumptions. Otherwise models may generate predictions that contrast with the empirical evidence.[3]

1.3 Literature Review

In this subsection we insert the paper in the literature.

Well before automated price comparison search engines appeared on the Internet, economists had recognized the importance of buyers' search costs for the functioning of oligopolistic markets. For example, models like the one by Baye et al. (1992) demonstrate that any arbitrarily small search cost could yield monopoly pricing by competing sellers.[4] Obviously, there are alternative approaches in the literature which assume buyer heterogeneity as far as search costs are concerned. In the presence of buyer heterogeneity mixed price equilibria may emerge. Such equilibria are often used as the theoretical counterpart of empirical evidence on persistent price dispersion. An interesting special case of such heterogeneity is one in which for a fraction of the consumer population search is costless, whereas the remaining population have identical positive search costs. Under this assumption, models like the one in Stahl (1989) show that mixed strategy equilibrium price distributions range from monopoly to competitive levels as the fraction of zero search cost buyers varies from 0 to 1.

Recently, the use of the Internet and the existence of price comparison search engines largely characterize the properties of the search process followed by individual buyers. Specifically, buyers may be heterogeneous depending on whether they gather price information on the Internet or not. This fact is

[3]Of course the role of behavioral factors can also be magnified in experimental laboratory in comparison with real-world markets.

[4]On this matter, Diamond (1971), Varian (1980), Burdett and Judd (1983) and Stahl (1989) constitute important references.

tested by empirical results reported by Iyer and Pazgal (2000). Furthermore, search engines may offer incomplete or even biased coverage of competing firms' prices, as reported in Kephart and Greenwald (1999).

Baye and Morgan (2001) from a theoretical perspective, as well as Brynjolfsson and Smith (2000) from an empirical point of view, confirm the prediction of persistence price dispersion in Internet markets, even in the presence of homogeneous products.

From an experimental perspective, two laboratory studies have explicitly tested the predictions of theoretical search models with heterogeneous consumers. Cason and Friedman (2003) study markets with costly buyer search. The issue of sample size used in the search process is explicitly addressed allowing for both, human and automated, price search on the demand side. Theoretical predictions in their model are mostly supported by the evidence.

In a different setup, much more similar to ours but assuming complete coverage, Morgan et al. (2004) study the case of a consumer population consisting of Internet searchers and captive clients. Some systematic deviations between theoretical and experimental results are attributed by the authors to uncontrolled idiosyncratic features of their subjects. However, biases of price comparison search engines and the effect of incompleteness, have been left unexplored.

The remainder of the paper 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, A,and C, respectively, include the proofs, the experimental instructions and, and supplementary econometric results .

2 The Model

In this section we present the benchmark 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 $j = 1, \dots, n$, 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 database, i.e., in its in its index, and the prices they charge. The search engine has 1 of 3 types: **(i)** *Complete Coverage*, **(ii)** *Unbiased Incomplete Coverage*, and **(iii)** *Biased Incomplete Coverage*. 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.

2.2.1 Complete Coverage

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 has *Complete Coverage* if it indexes all vendors present in the market: $k = n$.

2.2.2 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*.

2.2.3 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 parameterization, knowing the probability with which vendors are indexed, implies knowing the identity of the indexed vendors.[5]

[5]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 parameterization is that it allows for 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, which differ only with respect to whether they use the price-comparison search engine. *Non-Shoppers*, a proportion λ in $(0, 1)$ 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.[6] 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. Denote by $\Pi_j(p)$, the expected profit of vendor j when it charges price p on $\mathbb{R}_0^+$. Vendors know the behavior 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. 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

[6]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 67% visited an online retailer (Montgomery et al. (2001)). 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)).

$\bar{p}_j$.[7] 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 Characterization of Equilibrium

In this section we construct the equilibrium of the model.

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

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

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:

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

Ignoring ties,[8] the expected profit of a vendor that charges $p \leq 1$ is:[9]

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

[7] As it is well known this game has no equilibrium in pure strategies (Varian (1980)).

[8] Lemma 1: **(ii)** shows that $F_j^\tau(\cdot)$ is continuous.

[9] Under *Unbiased Incomplete Coverage*, a vendor that charges a price $p \leq 1$ sells to shoppers: **(i)** if it belongs to the set of vendors

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.[10] The next proposition characterizes the equilibrium for the model.[11]

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

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

with

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

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

Under *Biased Incomplete Coverage*, vendors $j = k+1, \ldots, n$, are not indexed for sure, and therefore have no access to shoppers. Since these vendors can only sell to non-shoppers, which are captive con-

indexed by price-comparison search engine, which occurs with probability k/n , 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: $(1-\lambda)(k/n)[1-F(p)]^{k-1}$. Under *Biased Incomplete Coverage*, an indexed vendor that charges a price $p \leq 1$ has the lowest price among the indexed vendors with probability $[1-F(p)]^{k-1}$. Thus, the expected share of shoppers of an indexed vendor that charges price $p \leq 1$ is: $(1-\lambda)[1-F(p)]^{k-1}$.

[10] Function $\delta(\cdot)$ is the Dirac delta function.

[11] 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 $[1 - F_i(x_i)]$.

sumers, they charge the reservation price. Vendors $j = 1, \ldots, k$, 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 this subsection, we discuss the case of *Complete Coverage*.

In the case of *Complete Coverage* the model is similar to Varian (1980).[12]

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 looses $(1 - p)$ 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.[13] This has two implications. First, only shifts in the proportion of

[12] See also Rosenthal (1980) and Stahl (1989).

[13] 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 paper.

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 that use 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.

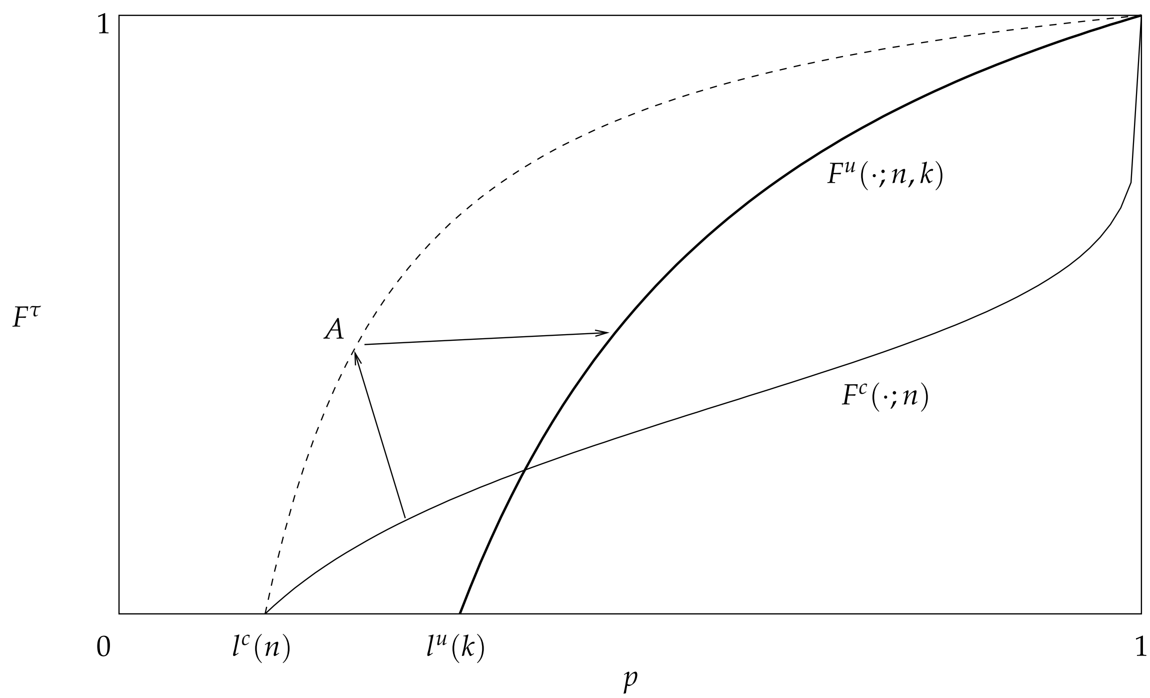
Proposition 2 ***(i)*** *$l^u(k) < l^u(k-1)$;* ***(ii)*** *$\mu^u(n,k) < \mu^u(n,k-1)$ and $\varepsilon^u(n,k) > \varepsilon^u(n,k-1)$;* ***(iii)*** *$F^u(\cdot; n, k) = F^u(\cdot; n+1, k)$.* §

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

FIGURE 1: Unbiased Incomplete Coverage



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 $F^u(\cdot; n, k)$.

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 (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.[14]

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, $\varepsilon^u(n,k) > \varepsilon^u(n,k-1)$.[15] Recall that vendors now charge lower prices with a higher probability. Shoppers and non-shoppers have conflicting interests with respect to *Unbiased Incomplete Coverage*, as compared

[14] See Guimarães (1996) for a related discussion.

[15] See Morgan et al. (2004) for a related discussion.

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 paper. 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)*** *$l^c(n) < l^b(k)$;* ***(ii)*** *$\mu^c(n) < \mu^u(n, k)$ and $\varepsilon^c(n) > \varepsilon^u(n, k)$.* §

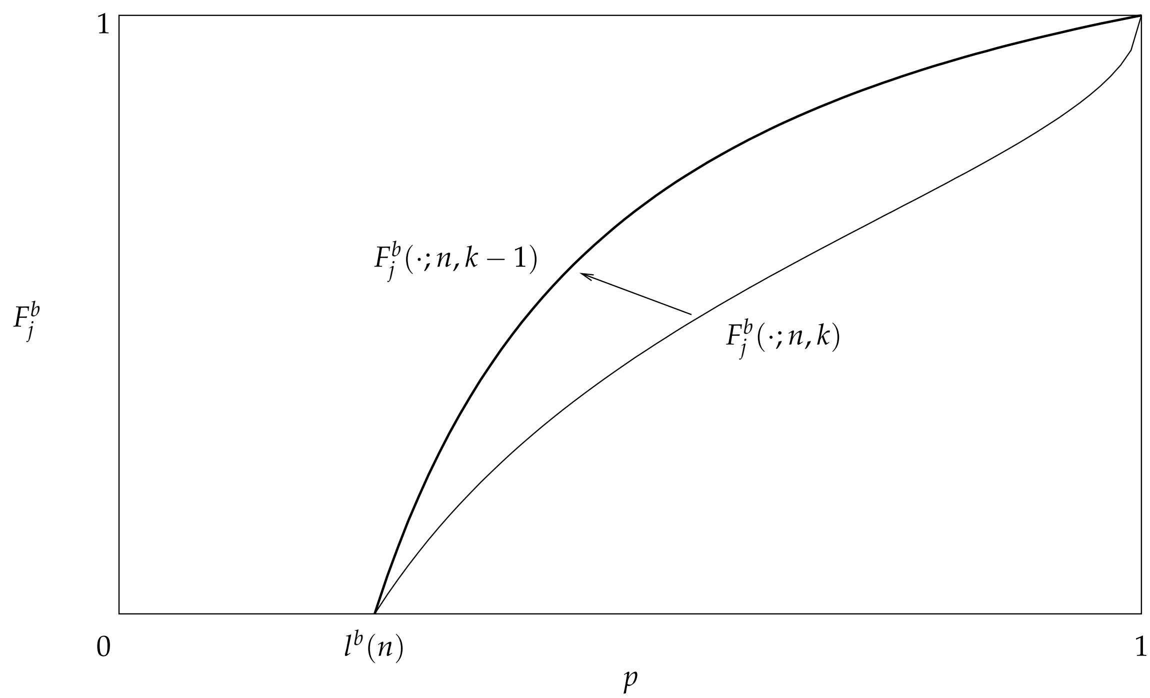
Given that $F^u(\cdot; n, k) = F^c(\cdot; k)$, comparing the price distributions under *Unbiased Incomplete Coverage*, $F^u(\cdot; n, k)$, and under *Complete Coverage*, $F^c(\cdot; n)$, is equivalent to comparing $F^u(\cdot; n, k)$ and $F^u(\cdot; n, n)$, i.e., is equivalent to analyzing the impact 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, $\varepsilon^c(n) > \varepsilon^u(n, k)$.

4.3 Biased Incomplete Coverage

In this subsection, we analyze the case of *Biased Incomplete Coverage*, and compare it with the 2 previous 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.

FIGURE 2: Biased Incomplete Coverage: A Decrease in the Size of the Index



For $j = 1, \ldots, k$ distributions $F_j^b(\cdot; n, k-1)$ are first-order stochastically dominated by distributions $F_j^b(\cdot; n, k)$.

Proposition 3 ***(i)*** *For $j = 1, \ldots, k$, $F_j^b(\cdot; n, k) \leq F_j^b(\cdot; n, k-1)$;* ***(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 $j = 1, \ldots, k$, $F_j^b(\cdot; n+1, k) \geq F_j^b(\cdot; n, k)$;* ***(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 $j = 1, \ldots, k$.* §

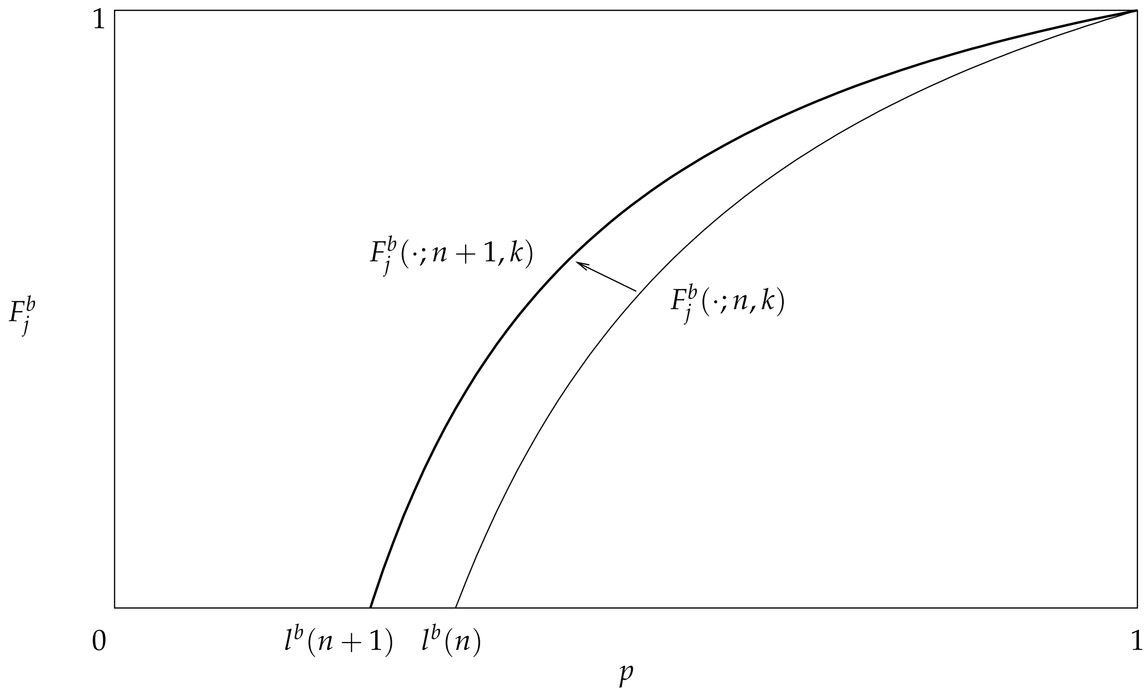
Rewrite (2) as

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

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 2). 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, $\varepsilon_j^b(n, k-1) < \varepsilon_j^b(n, k)$, $j = 1, \ldots, k$, and leaves unchanged the expected price paid by non-shoppers that buy from a non-indexed vendor, $\varepsilon_j^b(n, k-1) = \varepsilon_j^b(n, k)$, $j = k+1, \ldots, n$.

From (5), an increase in the number of vendors in the market, n, leaving fixed the size of a biased

FIGURE 3: Biased Incomplete Coverage: An Increase in the Number of Firms



For $j = 1, \ldots, k$ distributions $F_j^b(\cdot; n+1, k)$ are first-order stochastically dominated by distributions $F_j^b(\cdot; n, k)$.

index, k, reduces the *per Consumer Profit effect*. This leads indexed vendors to reduce the lower bound 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, $F_j^b(p; n+1, k) \geq F_j^b(p; n, k)$ (Figure 3). 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, $\varepsilon_j^b(n+1, k) = \varepsilon_j^b(n, k)$, $j = k+1, \ldots, n$.[16]

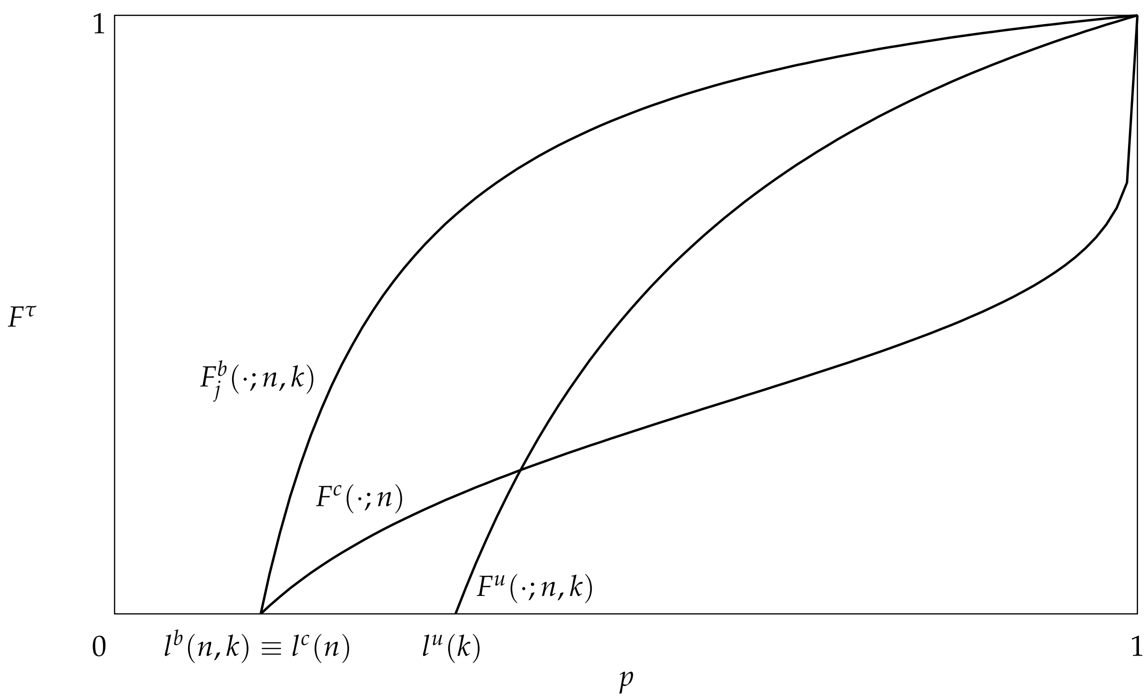
The next corollary compares the case of *Biased Incomplete Coverage*, with the two previous cases.

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 on the *Volume*

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

FIGURE 4: 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 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* $F^u(\cdot;n,k)$ (Figure 4). 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.

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 Main Experiment

In this subsection we describe the experimental design.

The experiment was conducted at the Laboratori d'Economia Experimental, LEE, of the Universitat Jaume I, Castellón, Spain, during the summer 2004. 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. 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). Nothing was said on the aims of the experiment, and the alternative hypotheses. We programmed software using z-Tree (Fischbacher (1999)) in order to organize strategy submission, demand simulation, feedback, and data collection.

Demand was simulated by the local network server. There were 1200 consumers. The consumers' reservation price was normalized to 1000 rather than to 1, for representation and interface reasons, and in order to offer a fine grid for subjects' strategy space. Half of the consumers were shoppers, and the other half were non-shoppers, i.e., we assumed that $\lambda = 1/2$.

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*, the subjects whose prices were indexed were chosen at random. Subjects were informed after each period's prices were set, of whether they where indexed. Under *Biased Incomplete Coverage*, for $k = 2$, with $n = 3$ or $n = 6$, the prices of firms 1 and 2 were indexed, whereas for $k = 4$ and $n = 6$, the prices of firms 1, 2, 3, 4 were indexed. Before each period's prices were set, subjects participating in treatments with *Biased Incomplete Coverage* were informed of whose prices would be indexed. 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]

[17]We allow that subjects observe their rivals' prices for two reasons. First, for realism. Second, because it helps subjects infer the

TABLE 1: Design of Treatments

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.

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.

The 20 initial periods were dropped from the data sets used in the empirical analysis, to eliminate learning dynamics, and guarantee that observations had reached the necessary stability.

5.2 Parallel Experiment

Our theoretical model assumed risk neutrality. However, a similar experiment reported in Morgan et al. (2004) has shown a systematic deviation of the experimental data from the theoretical predictions. The authors conjectured —without explicitly accounting for this in their design— that the subjects' attitudes towards risk should be the cause of this deviation. Risk attitudes are likely to matter in our framework, for two reasons. First, because, similar to Morgan et al. (2004), depending on rival indexed

types of strategies that are adopted and abandoned by the rest of the market, which speeds convergence.

prices, subjects have random sales. Second, because in our unbiased incomplete coverage treatments, subjects are indexed randomly. Following these conjectures, in addition to the main experiment, subjects were faced with a parallel experiment, whose aims and design are described next.

In order to account for risk attitudes, we have used the lottery panel test proposed by Sabater and Georgantzís (2002). Our objective is to use the data obtained from the test as an explanatory factor of any systematic divergence between observed behavior and theoretical predictions. The test is designed to elicit two dimensions of the subjects' risk attitudes. First, the subjects' degree of risk aversion. And second, whether subjects accepted more risk in exchange for a higher expected return.

Following a standard protocol used in the LEE, when registering in our subject pool, subjects had participated previously in the following experiment. They were offered the 4 panels of lotteries in Table 2, involving a probability q of earning a monetary reward of X €, and a probability $1-q$ of earning 0 €. Each one of the 4 panels was constructed using a fixed certain payoff, $c = 1$ €, above which expected earnings, qX, were increased by a ratio h times the probability of not winning, $1-q$, as implied by the formula: $qX = c + h(1-q)$. That is, an increase in the probability of the unfavorable outcome is linearly compensated by an increase in the expected payoff. We used 4 different risk return parameters, $h = 0.1, 1, 5, 10$, implying an increase in the return of risky choices as we moved from one panel to the next. Simple inspection of the panels shows that risk loving and risk neutral subjects would choose $q = 0.1$ in all panels.[18] The more risk averse a subject is, the lower the risk he will assume, i.e., the higher the q he would choose. For risk-averse subjects, their sensitivity to the attraction implied by a higher h was approximated by the difference in their choices across subsequent panels.[19] More specifically, we have considered the sign of transitions across panels as a qualitative variable referring to a subject's tendency to comply, in the case of a negative transition, with the pattern of assuming more risk in the presence of a higher risk return. For labelling purposes, we refer to this pattern of behavior as *Monotonicity*. While measures of local risk aversion are commonly obtained from binary lottery tests,[20] this second aspect of behavior captured by the test concerns a subject's responses to changes in the

[18]Risk neutrality and risk-loving behavior are observationally impossible to distinguish from our test.

[19]In fact, using the utility function $U(x) = x^{1/t}$, it can be shown that the optimal probability corresponding to an Expected Utility maximizing subject is given by $q^{\star} = (1-(1/t))(1+(c/r))$. Apart from illustrating the panel-specific positive relation between the probability chosen and a subject's risk aversion, this would imply that subjects should choose weakly (given the discreteness of the design) lower probabilities as we move from one panel to the next one. However, using more general utility functions or non Expected Utility theories, one can easily construct counterexamples of the aforementioned choice pattern.

[20]For example, Holt and Laury (2002) use binary lottery choice tasks, in order to obtain the parameter r of the utility function, $U(x) = x^{1-r}/(1-r)$.

TABLE 2: Panels for the Parallel Experiment

Panel 1

q	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
X €	1.00	1.12	1.27	1.47	1.73	2.10	2.65	3.56	5.40	10.90
Choice										

Panel 2

q	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
X €	1.00	1.20	1.50	1.90	2.30	3.00	4.00	5.70	9.00	19.00
Choice										

Panel 3

q	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
X €	1.00	1.66	2.50	3.57	5.00	7.00	10.00	15.00	25.00	55.00
Choice										

Panel 4

q	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
X €	1.00	2.20	3.80	5.70	8.30	12.00	17.50	26.70	45.00	100
Choice										

expected profitability of riskier choices.

Subjects were informed that after submitting their four choices, one per panel, a four-sided die would be used to determine which panel was binding. Following this stage, a ten-sided die was used to determine whether the subject would receive the corresponding payoff or not, depending on the odds corresponding to the subject's choice in the panel drawn.

5.3 Testable Implications

In this subsection we present the hypotheses to be tested, expressed in terms of our experimental design.

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; by μ_t the expected minimum price in treatment t, where $t = 1, \ldots, 8$.

We perform the following *Consistency Test*:

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

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

We also test:

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* $\varepsilon_2 > \varepsilon_3^{in}$.

5.4 Results: Descriptive Analysis

In this subsection we analyze the experimental data.

Table 3 summarizes the information and descriptive statistics regarding all treatments. Seven conclusions emerge from the descriptive statistics.

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 Table 4 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, which are shown in Table 1.

TABLE 3: Descriptive Statistics

Treatment	Average prices		Minimum prices		Obs.
	Mean	Std. Dev.	Mean	Std. Dev.	
1	0.535	0.158	0.294	0.103	180
2	0.760	0.109	0.635	0.157	180
3	0.637	0.211	0.475	0.247	180
4	0.700	0.136	0.193	0.161	90
5	0.592	0.149	0.204	0.197	90
6	0.591	0.177	0.208	0.176	90
7	0.758	0.088	0.644	0.258	90
8	0.416	0.214	0.269	0.187	90

Means and standard deviations of the empirical distributions of average and minimum prices. In the biased treatments, 3, 6 and 8, only indexed firms are considered.

TABLE 4: t-tests of Equality of Price Distribution Means to their Theoretical Values

Treatment	Average prices	Minimum prices	d. f.
1	−5.90 [0.00]	−13.14 [0.00]	179
2	25.84 [0.00]	15.68 [0.00]	179
3	11.13 [0.00]	6.35 [0.00]	179
4	−0.58 [0.56]	9.98 [0.00]	89
5	−3.52 [0.00]	−7.17 [0.00]	89
6	0.19 [0.85]	−3.64 [0.00]	89
7	22.52 [0.00]	7.10 [0.00]	89
8	4.06 [0.00]	2.21 [0.03]	89

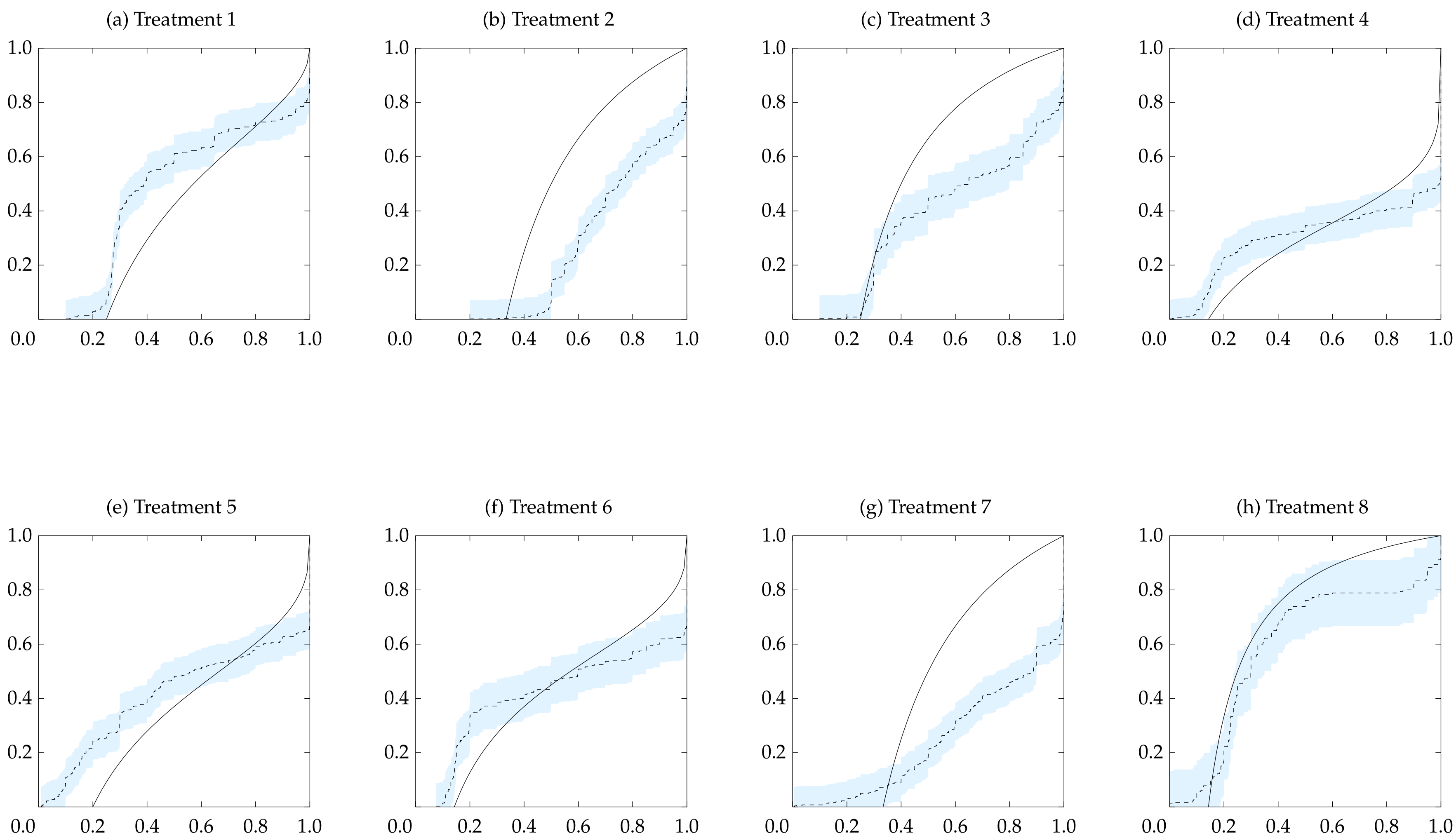
Two-sided t-tests with 'd. f.' degrees of freedom (p-values between brackets). The means of the theoretical distribution of average prices and minimum prices are reported in Table 1. In the biased treatments, 3, 6 and 8, only indexed firms are considered.

Second, this conclusion can also be gleaned from the inspection of Figure 5, 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 5, each empirical price distribution is surrounded with a confidence region built from the 1% critical values of the Kolmogorov-Smirnov test. As none of those confidence regions include the whole theoretical distribution, the observed price distributions are not compatible with the theoretical ones.[21]

The empirical distributions rotate clock-wise compared to the theoretical distributions. In the case of treatments 2 and 3, the empirical distribution "almost" first-order stochastically dominates the theoretical distribution. The rotation of the empirical distributions indicates the presence of more density at both tails of distributions of observed prices, than the theoretical model would have predicted. In fact, 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. And in addition, 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,

[21] As is well known, the Kolmogorov-Smirnov test statistic is the maximum of the absolute value of the difference between the two distributions being compared. Therefore, as none of the theoretical distributions completely lies in the confidence regions of Figure 5, the Kolmogorov-Smirnov test rejects the null hypothesis of equality of the theoretical and empirical distributions in all cases.

FIGURE 5: Theoretical (solid lines) and Empirical (dashed line) Price Distributions.



where $p = 1$ is an accumulation point of the empirical distributions.[22] In line with the way in which the empirical distributions rotate, all of the empirical price distributions have higher standard deviations than the corresponding theoretical ones (see Tables 1 and 3).

We also have 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 many 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 Tables 5 and 6, 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 test, **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 6. This non-trivial result was also obtained by Morgan et al. (2004).

Observation 3 *The data supports the model's predictions regarding the comparison between Unbiased Incomplete Coverage and Biased Complete Coverage.* §

From Tables 5 and 6, 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**. This implies that the average price and the average minimum price 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 7.

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

[22] This finding could be interpreted in terms of the asymmetric equilibrium referenced in footnote 11.

TABLE 5: t-tests of Equality of Means of Average Prices

	H_0	H_1	d. f.	t-test	p-value
HC	$\varepsilon_1 = \varepsilon_4$	$\varepsilon_1 < \varepsilon_4$	268	−8.44	0.00
HU1	$\varepsilon_5 = \varepsilon_4$	$\varepsilon_5 < \varepsilon_4$	178	−5.06	0.00
	$\varepsilon_7 = \varepsilon_4$	$\varepsilon_7 < \varepsilon_4$	178	3.42	1.00
	$\varepsilon_7 = \varepsilon_5$	$\varepsilon_7 < \varepsilon_5$	178	9.11	1.00
	$\varepsilon_2 = \varepsilon_1$	$\varepsilon_2 < \varepsilon_1$	358	15.68	1.00
HU2	$\varepsilon_2 = \varepsilon_7$	$\varepsilon_2 \neq \varepsilon_7$	268	0.11	0.91
HB1	$\varepsilon_6^{in} = \varepsilon_4$	$\varepsilon_6^{in} < \varepsilon_4$	178	−4.64	0.00
	$\varepsilon_8^{in} = \varepsilon_4$	$\varepsilon_8^{in} < \varepsilon_4$	178	−10.60	0.00
	$\varepsilon_8^{in} = \varepsilon_6^{in}$	$\varepsilon_8^{in} < \varepsilon_6^{in}$	178	−5.96	0.00
	$\varepsilon_3^{in} = \varepsilon_1$	$\varepsilon_3^{in} < \varepsilon_1$	358	5.19	1.00
HB2	$\varepsilon_8^{in} = \varepsilon_3^{in}$	$\varepsilon_8^{in} < \varepsilon_3^{in}$	268	−8.07	0.00
HG	$\varepsilon_6^{in} = \varepsilon_5$	$\varepsilon_6^{in} < \varepsilon_5$	178	−0.07	0.47
	$\varepsilon_8^{in} = \varepsilon_7$	$\varepsilon_8^{in} < \varepsilon_7$	178	−14.01	0.00
	$\varepsilon_3^{in} = \varepsilon_2$	$\varepsilon_3^{in} < \varepsilon_2$	358	−6.93	0.00

Except for **HU2**, one-sided t-tests with 'd.f' degrees of freedom. The testable implications of section 5.3 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.

From Tables 5 and 6, 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 Tables 5 and 6, 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 6: t-tests of Equality of Means of Minimum Prices

	H_0	H_1	d.f.	t-test	p-value
HC	$\mu_4 = \mu_1$	$\mu_4 < \mu_1$	268	−6.23	0.00
HU1	$\mu_4 = \mu_5$	$\mu_4 < \mu_5$	178	−0.39	0.34
	$\mu_4 = \mu_7$	$\mu_4 < \mu_7$	178	−14.06	0.00
	$\mu_5 = \mu_7$	$\mu_5 < \mu_7$	178	−12.87	0.00
	$\mu_1 = \mu_2$	$\mu_1 < \mu_2$	358	−24.26	0.00
HU2	$\mu_2 = \mu_7$	$\mu_2 \neq \mu_7$	268	−0.35	0.72
HB1	$\mu_6 = \mu_4$	$\mu_6 < \mu_4$	178	0.59	0.72
	$\mu_8 = \mu_4$	$\mu_8 < \mu_4$	178	2.91	1.00
	$\mu_8 = \mu_6$	$\mu_8 < \mu_6$	178	2.25	0.99
	$\mu_3 = \mu_1$	$\mu_3 < \mu_1$	358	9.09	1.00
HB2	$\mu_8 = \mu_3$	$\mu_8 < \mu_3$	268	−7.00	0.00
HG	$\mu_6 = \mu_5$	$\mu_6 < \mu_5$	178	0.15	0.56
	$\mu_8 = \mu_7$	$\mu_8 < \mu_7$	178	−11.17	0.00
	$\mu_3 = \mu_2$	$\mu_3 < \mu_2$	358	−7.30	0.00

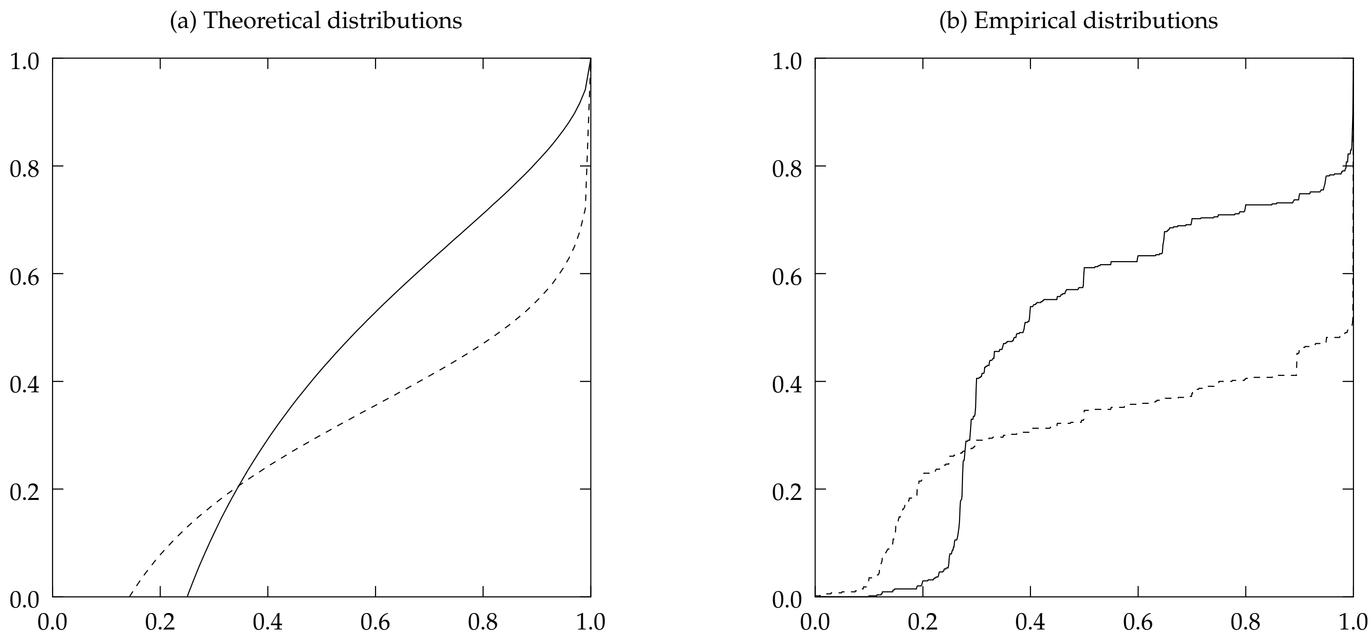
Except for **HU2**, one-sided t-tests with 'd.f' degrees of freedom. The testable implications of section 5.3 correspond to the alternative hypothesis of these tests. For **HU2**, two-sided t-test with 'd.f' degrees of freedom. **HU2** correspond to the null hypothesis of this test.

Tables 5 and 6, 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*.

With respect to the comparison between *Complete Coverage* and *Biased Incomplete Coverage*, from Tables 5 and 6, 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*.

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

FIGURE 6: Comparison of Price Cumulative Distributions: Treatments 1 (solid line) and 4 (dashed line).



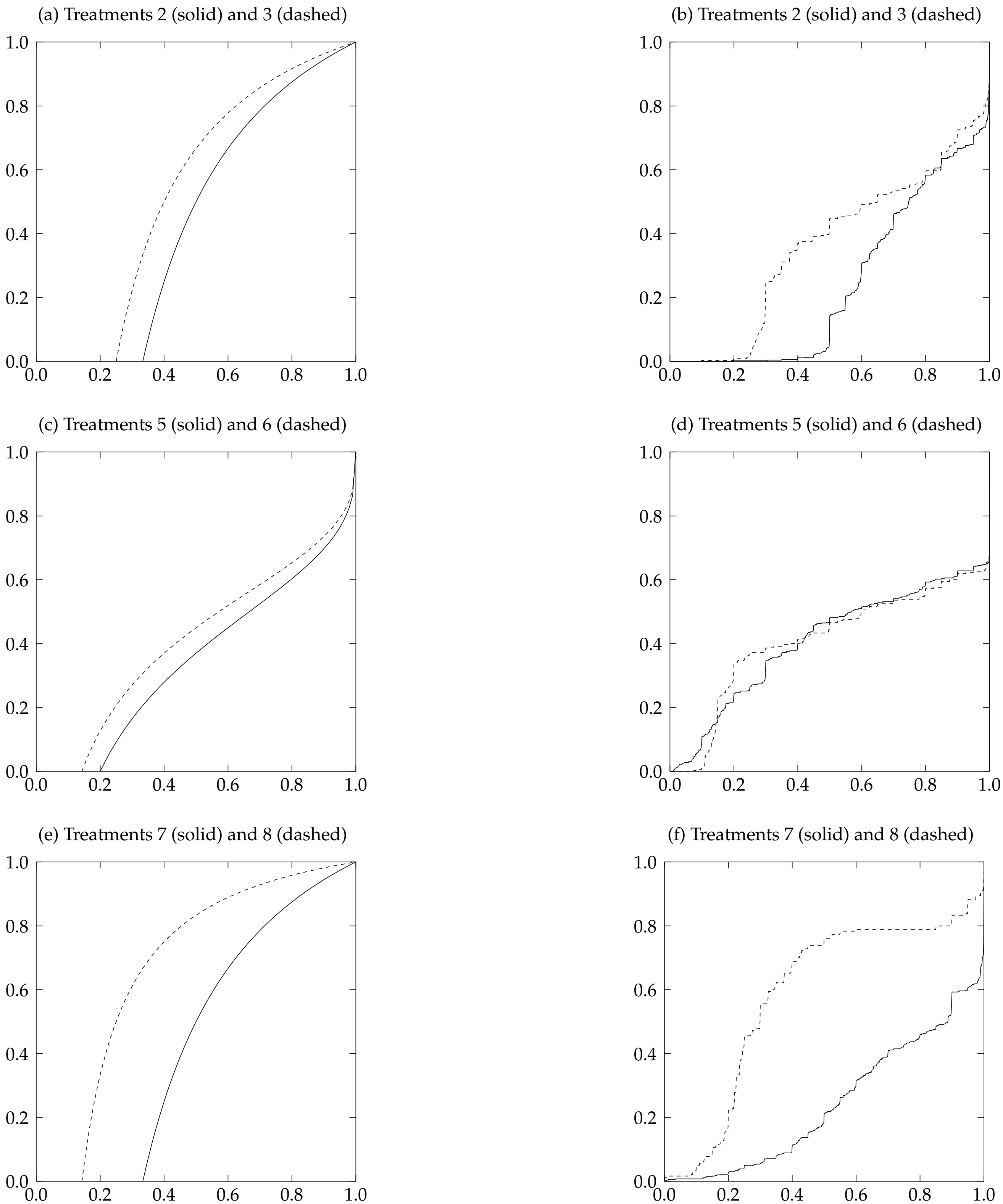
From Table 6 it follows that: **(i)** $\mu_1 < \mu_3$, **(ii)** $\mu_4 = \mu_6$, and **(iii)** $\mu_4 < \mu_8$. Jointly with Observation 3, this imply that shoppers are better off under *Complete Coverage* than under *Incomplete Coverage*.

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 Tables 5 and 6 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).

We interpret these observations as follows. On the basis of descriptive analysis alone, the data supports some of the model's predictions, and rejects others. Given that some of the predictions that the data supports are non-trivial, particularly the consistency test, **HC**, the underlying Burdett and Judd (1983) and Varian (1980) model is sound. However, there is some factor, the model does not account for, that impacts systematically the way subjects play the game. In the case of *Complete Coverage* this unaccounted factor is qualitatively unimportant. However, under *Incomplete Coverage* this unaccounted factor becomes determinant. The model assumes that subjects are risk neutral, expected utility maximizers. The unaccounted factor could be the subjects' risk attitudes.

FIGURE 7: Comparison of Price Distributions



5.5 Results: Accounting for Risk Attitudes

In this subsection we analyze risk attitudes. As already said, for the purposes of our analysis we focus on two aspects of observed behavior: **(i)** the average choice over the four panels, and **(ii)** the "transition" across two subsequent panels. The former is a proxy for a subject's risk aversion, and the latter a qualitative variable reflecting a subject's compliance with the monotonicity property.

Table 7 summarizes the information and descriptive statistics regarding the parallel experiment for all treatments. Next to each treatment's number, we provide the percentage of choices which are compatible with risk neutrality (*Risk Neutr.* column) and the average choice by all subjects in the treatment over all panels (*Avr. Choice*). The remaining columns refer to percentages of subjects fulfilling the monotonicity property at three different levels: *Strict*, requiring that a subject's choices are all compatible with the pattern of a riskier choice in the presence of a higher return to risk; *Average*, requiring that, on average, a subject's choices follow the aforementioned pattern; and *Local*, requiring a negative transition from panel 1 to panel 2 alone.

Now we can formulate two testable hypotheses related to our subjects' risk attitudes:

HRA: *Subjects are risk neutral, i.e., subjects choose* $q = 0.1$*.*

HM: *Subjects will choose an equally risky or a riskier option in the presence of higher returns to risk. That is, across panels, as the return to risk is increased, subjects choose a no higher* q*.*

From Table 7, one concludes the following. Our results confirm previous findings of choice agglomeration around probabilities close to 0.5 (Georgantzís et al. (2003)). Specifically, Table 7 indicates that the average choices exhibit no systematic patterns across treatments. Overall, slightly less than 10% of the decisions would be compatible with risk neutrality in any one of the panels. However, this percentage falls below any significance level if we require that a subject behaved in a way compatible with risk neutrality in all four panels .

It is straightforward to conclude that:

Observation 8 *The data does not support the hypothesis that subjects are risk neutral.* §

As we said, the column labelled *Strict* in Table 7 presents the percentage of subjects whose choice across panels confirms the prediction that a higher risk return induces riskier options. Such percentage varies from one treatment to the other, ranging from 27.7% in treatment 2 to 55.5%, in treatment 4. Overall, only 38% of our subjects follow this pattern. Requiring that this pattern is observed only as an

TABLE 7: Parallel Experiment Results

Treatment	Risk Neutr.	Avr. Choice	Monotonicity compliance measures		
			Strict	Average	Local
1	1.39	0.59	44.44	72.22	77.78
2	15.28	0.46	27.78	77.78	38.89
3	4.17	0.54	33.33	55.56	55.56
4	20.83	0.44	55.56	88.89	83.33
5	5.56	0.59	33.33	77.78	72.22
6	5.56	0.52	33.33	66.67	66.67
7	8.33	0.59	33.33	66.67	55.56
8	18.06	0.50	50.00	77.78	61.11
Average	9.90	0.53	38.89	72.92	63.89

average behavior, over the four panels, because human errors might induce inconsistent transition from a given panel to the next one, raises the percentage of *Monotonicity-compatible* subjects over the whole population up to 72.9% (this *Average* monotonicity compliance measure of Table 7). Following specific utility functions used in expected utility generalizations, one should also look at local behavior (*Local*), which can be done by looking at the transition from panel 1 to 2, in which the minimum increase in the risk-return payoff has been used. The fifth column of the Table shows that only 63.9% of our subjects have fulfilled monotonicity.

Therefore:

Observation 9 *The data does not support the hypothesis that subjects will accept more risk in lottery-panels offering a higher return to risk.* §

5.6 Estimation of Price Distributions

Our theoretical model gives us clear predictions under homogeneity of subjects' utility functions, and risk neutrality. Having established the violation of these hypotheses, we need a tool to assess its impact on experimental results. Our strategy consists of estimating price distributions conditional on treatment-specific or design parameter-specific variates, on one hand, and, individual-specific characteristics related to risky choice, on the other hand. From the estimated price distribution we can derive the distributions of both average and minimum prices and assess the impact of design parameters and risk attitudes on these distributions.

We model the observed prices as coming from beta distributions. The beta distribution is commonly used in the statistical analysis of variables with a bounded support. It is a flexible distribution which can accommodate asymmetries and various distributional shapes: uniform, bell-shaped, U-shaped, J-shaped. As it is a distribution for bounded variables, it naturally incorporates a relation between mean and variance, so that, for unimodal distributions, the variance is lower for beta variables with a mean near the boundaries than for beta variables whose mean lies around the center of the support. Finally, the mathematical tractability of the beta distribution generally puts no excessive burden on the estimation and interpretability of the results.

We allow some degree of heterogeneity by doing the estimation conditional on a set of K covariates, $\boldsymbol{x}_i = (x_{1i}, x_{2i}, \ldots, x_{Ki})'$, where subscript $i = 1, \ldots, N$ indexes the experimental observations. In our analysis, these covariates are related to the design of each treatment and to the risk attitudes of the subjects. We parametrize a price distribution in terms of its conditional mean, $E(p|\boldsymbol{x}_i) = \Theta_i = \Theta(\boldsymbol{x}_i)$, and a function $\Delta_i = \Delta(\boldsymbol{x}_i)$ inversely related to dispersion.[23] Given these definitions, the probability distribution function of prices is

$$f(p|\boldsymbol{x}_i) = \frac{\Gamma(\Delta_i)}{\Gamma(\Delta_i\Theta_i)\Gamma((1-\Theta_i)\Delta_i)} p^{\Delta_i\Theta_i - 1}(1-p)^{(1-\Theta_i)\Delta_i - 1} \tag{6}$$

where $0 \leq p \leq 1$, and $\Gamma(\cdot)$ is the gamma function. We use the following functional forms for the mean and the dispersion:

$$\Theta_i = \Lambda(\boldsymbol{x}_i'\boldsymbol{\theta}) \tag{7}$$

$$\Delta_i = \exp(\boldsymbol{x}_i'\boldsymbol{\delta}) \tag{8}$$

where $\Lambda(s) = 1/(1+e^{-s})$ is the logit function, and $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$ are K-dimensional vectors of unknown parameters. The functional forms (7) and (8) guarantee that Θ_i lies between 0 an 1 and Δ_i is greater than 0, for all possible values of the parameters, $\boldsymbol{\theta}, \boldsymbol{\delta}$, and the covariates, $\boldsymbol{x}_i$. Equations (6)–(8) allow us to write the logarithm of the likelihood function as

$$\ln L = \sum_{i=1}^{N} \ln f(p|x_i) \tag{9}$$

[23] Perhaps, it would be more natural to parametrize the price distribution in terms of its conditional mean and its variance. But this would complicate both the estimation procedure and the interpretation of estimates. This is so, due to the nonlinear relationship between the mean and the upper bound of the variance of a beta distributed variable. In any case, knowing Θ_i and Δ_i we can compute the variance as $\sigma_i^2 = \Theta_i(1-\Theta_i)/(\Delta_i + 1)$. Thus, given Θ_i, higher values of Δ_i imply lower variances.

TABLE 8: Descriptive Statistics of *MON* and *RISK*

Treatment	*MON*	$\bar{q}$ (All subjects)	$\bar{q}$ (*MON*)	$\bar{q}$ (non-*MON*)
1	0.78 (0.43)	0.59 (0.14)	0.61 (0.23)	0.53 (0.31)
2	0.39 (0.50)	0.46 (0.19)	0.51 (0.26)	0.43 (0.24)
3	0.56 (0.51)	0.54 (0.18)	0.62 (0.26)	0.44 (0.23)
4	0.83 (0.38)	0.44 (0.23)	0.39 (0.24)	0.68 (0.51)
5	0.72 (0.46)	0.59 (0.20)	0.55 (0.23)	0.71 (0.42)
6	0.67 (0.49)	0.52 (0.15)	0.57 (0.23)	0.43 (0.21)
7	0.56 (0.51)	0.59 (0.17)	0.62 (0.27)	0.54 (0.28)
8	0.61 (0.50)	0.50 (0.21)	0.50 (0.31)	0.50 (0.23)
All	0.64 (0.48)	0.53 (0.19)	0.54 (0.20)	0.51 (0.19)

Sample means and, between parentheses, standard deviations of the dummy variable *MON*, wich takes value 1 for subjects fulfilling the local monotonicity compliance measure of Table 7, and $\bar{q}$, which is the average choice in the parallel experiment. The last two columns report the sample statistics of $\bar{q}$ for the subsamples defined by the compliance or non-compliance of the monotonicity hypothesis.

The estimation of unknown parameters is performed by maximizing (9) with respect to $\boldsymbol{\theta}$ and $\boldsymbol{\delta}$. These parameters are positively related to the marginal effects of a covariate on the mean and dispersion of prices. Consider a change of covariate x_{li}. It is easy to see that

$$\frac{\partial \Theta_i}{\partial x_{li}} = \Theta_i(1-\Theta_i)\theta_l$$
$$\frac{\partial \Delta_i}{\partial x_{li}} = \Delta_i \delta_l$$

So the marginal effect on the mean (dispersion) varies at different points of the distribution, but the sign is the same as the sign of θ_l (δ_l).[24]

In our analysis, we distinguish between two types of subjects —labelled as *MON* and non-*MON*, respectively— according to the subject's compliance with the prediction that higher risk returns induce riskier options. This is done by checking whether a subject has chosen a weakly lower probability in

[24] The impact on the variance is more complicated to analyze as it depends on two distinct effects which can have different signs:

$$\frac{\partial \sigma^2}{\partial x_{li}} = \sigma^2 \left[(1-2\Theta_i)\theta_l - \frac{\Delta_i}{1+\Delta_i}\delta_l \right]$$

TABLE 9: Estimations of Price Distribution

	Model I	Model II	Model III	
	All subjects	All subjects	*MON*	non-*MON*
θ_1	0.510**	0.779**	0.555**	1.272**
	[0.047]	[0.083]	[0.099]	[0.197]
θ_2	1.315**	1.537**	1.042**	2.326**
	[0.042]	[0.071]	[0.098]	[0.141]
θ_3	0.793**	1.042**	0.858**	1.596**
	[0.056]	[0.085]	[0.117]	[0.173]
θ_4	1.033**	1.235**	1.058**	2.048**
	[0.051]	[0.075]	[0.080]	[0.236]
θ_5	0.696**	0.963**	0.692**	1.904**
	[0.048]	[0.087]	[0.100]	[0.223]
θ_6	0.716**	0.967**	0.664**	1.743**
	[0.057]	[0.087]	[0.109]	[0.179]
θ_7	1.259**	1.528**	1.343**	2.263**
	[0.056]	[0.090]	[0.112]	[0.198]
θ_8	0.056	0.282*	0.124	0.978**
	[0.095]	[0.109]	[0.126]	[0.215]
θ_r		−0.154**	0.016	−0.652**
		[0.040]	[0.048]	[0.098]
δ_1	0.211**	0.262**	0.077	1.182**
	[0.036]	[0.057]	[0.058]	[0.218]
δ_2	0.772**	0.820**	0.924**	1.168**
	[0.032]	[0.047]	[0.084]	[0.128]
δ_3	0.340**	0.391**	0.168*	1.096**
	[0.040]	[0.054]	[0.064]	[0.140]
δ_4	−0.258**	−0.201**	−0.302**	0.464*
	[0.026]	[0.045]	[0.042]	[0.204]
δ_5	−0.243**	−0.182**	−0.331**	0.531**
	[0.025]	[0.047]	[0.049]	[0.182]
δ_6	−0.232**	−0.175**	−0.326**	0.377*
	[0.025]	[0.047]	[0.053]	[0.144]
δ_7	0.339**	0.411**	0.358**	0.948**
	[0.046]	[0.065]	[0.064]	[0.193]
δ_8	0.224*	0.268*	0.300*	0.645**
	[0.099]	[0.105]	[0.124]	[0.231]
δ_r		−0.034	0.039	−0.354**
		[0.025]	[0.026]	[0.092]
RMSE	0.318	0.318	0.315	
$\ln L$	3288.4	3293.1	3347.6	
LR	118.4**	109.0**		
K	16	18	36	

Maximum-likelihood estimates of parameters in equations (10) and (11). White's (1982) robust standard errors between brackets. Sample size is 3600 observations. RMSE is the root mean squared prediction error, $\ln L$ is the value of the likelihood function, LR is the likelihood ratio statistic comparing each Model with Model III, K is the number of parameters estimated in each model. Coefficients marked with ** (*) are significantly different from 0 at the 1% (5%) level.

panel 2 than in panel 1. A more standard feature of a subject's risk attitude (risk aversion in *MON*-terms) is captured by $\bar{q}$, the subject's average choice over the four panels —see descriptive statistics of these variables in Table 8. In Table 9 we present three estimations constructed to assess the impact of *MON*-compliance and risk aversion.[25] Following equations (7) and (8), the specification for the mean and the dispersion of prices is given in the next two equations

$$\Theta_i = \Lambda(\theta_1 T_{1i} + \theta_2 T_{2i} + \cdots + \theta_8 T_{8i} + \theta_r \ln r_i) \tag{10}$$

$$\Delta_i = \exp(\delta_1 T_{1i} + \delta_2 T_{2i} + \cdots + \delta_8 T_{8i} + \delta_r \ln r_i) \tag{11}$$

where T_{ti} is a dummy variable taking a value of 1 if observation i corresponds to treatment t and 0 otherwhise and r_i is a measure of the degree of risk aversion of the setter of price i. We define r_i as the average choice made in the parallel experiment discussed above normalized so that it takes a value of 1 for subjects that choose $q = 0.1$ in the four panels, i. e., $r_i = \bar{q}_i/0.1$. For such subjects we use the term risk neutral throughout the text.[26]

Observation 10 *Subjects' risk attitudes significantly affect the observed pricing behavior.*

Under the heading Model I in Table 9, we present estimates of the simplest model in which only treatment variables, $T_{1i}, \ldots, T_{8i}$, are included. Comparison with Model II of Table 9 shows that the risk aversion variable improves the fit. The θ_r parameter indicates a significant negative relationship between the mean and r_i but a non-significant impact on the dispersion estimate. Model III of Table 9 extends Model II by allowing for differences in the parameters of price distributions for *MON* and non-*MON* subjects. We find significant differences between the two sets of parameter estimates and that the split into two subsamples significantly improves the overall fit. The likelihood ratio tests statistics strongly reject Model I and Model II in favor of Model III. Risk aversion turns out to significantly affect both the mean and the dispersion of price distributions for non-*MON* subjects only, indicating a non-negligible interaction between *MON*-compliance and risk aversion.

Observation 11 ***(i)*** *Risk neutral non-MON subjects tend to set higher prices, which exhibit a lower dispersion, than MON subjects.* ***(ii)*** *For non-MON subjets, a higher degree of risk aversion tends to lower prices and to increase dispersion.*

---

[25]The results of this section are generated using a program written for Ox version 3.30 (Doornik (2002)).

[26]Note that this is done for labelling purposes, given that choosing $q = 0.1$ is also compatible with risk-loving behavior as well as with some, arbitrarily low, degree of risk-aversion.

The first conclusion is based on the rejection of equality of the parameters for *MON* and non-*MON* subjects, and on the observation that $\theta_1, \ldots, \theta_8$ and $\delta_1, \ldots, \delta_8$ parameter estimates take higher values for non-*MON* subjects. The second conclusion follows from the significant negative estimates of θ_r and δ_r parameters for non-*MON* subjects.

The reported heterogeneity of subjects with respect to risk attitudes can explain the rotation of empirical price distributions discussed in Observation 1. The presence of non-*MON* subjects increases probability mass on both high and low prices. On one hand, risk neutral non-*MON* subjects shift probability mass from lower to higher prices. On the other hand, risk averse non-*MON* subjects shift probability mass from higher to lower prices. The varying patterns of rotations found in Figure 5 can be explained by differences across treatments in the proportion of non-*MON* subjects and their degrees of risk aversion.

So far, our analysis concerns treatment dummies. However, treatments are designed using different parameters of the model's variables. Therefore, we can also estimate an alternative model whose explanatory terms are specific to these design variables. We use the following specifications for the mean and the dispersion of price distributions:

$$\Theta_i = \Lambda(\theta_0 + \theta_k^u(1-b_i)\ln k_i + \theta_n^u(1-b_i)\ln n_i + \theta_k^b b_i \ln k_i + \theta_n^b b_i \ln n_i + \theta_r \ln r_i) \quad (12)$$

$$\Delta_i = \exp(\delta_0 + \delta_k^u(1-b_i)\ln k_i + \delta_n^u(1-b_i)\ln n_i + \delta_k^b b_i \ln k_i + \delta_n^b b_i \ln n_i + \delta_r \ln r_i) \quad (13)$$

where k_i is the size of the index, n_i is the number of firms and b_i is a dummy variable taking a value of 1 for observations corresponding to biased treatments and 0 otherwise.[27] All other notation is the same as before. As the *Complete Coverage* case can be thought as a limit case of both *Biased* and *Unbiased Incomplete Coverage* we impose the additional restrictions $\theta_k^u + \theta_n^u = \theta_k^b + \theta_n^b$ and $\delta_k^u + \delta_n^u = \delta_k^b + \delta_n^b$. This restriction guarantees that, when $n = k$, the price distributions for *Unbiased* and *Biased Incomplete Coverage* collapse to a common distribution, corresponding to *Complete Coverage*, but still allows different patterns of influence of n and k on the *Incomplete Coverage* cases.

Table 10 reports estimates of the parameters in equations (12) and (13). Again, likelihood ratio test statistics reject the simpler Model IV and Model V in favor of the more general Model VI, that accounts for differences in the compliance of the monotonicity hypothesis and for varying degrees of risk aversion. Like before, homogeneity between *MON* and non-*MON* subjects is rejected. In Model VI, except

[27]We have tried more complex specifications including squares and interactions of the regressors. But that increased complexity hardly improved the estimation results.

TABLE 10: Price Distribution as a Function of the Model's Parameters

	Model IV	Model V	Model VI	
	All subjects	All subjects	*MON*	non-*MON*
θ_0	0.894** [0.081]	1.339** [0.097]	0.765** [0.142]	2.053** [0.145]
θ_k^u	−0.300** [0.047]	−0.341** [0.047]	−0.213** [0.059]	−0.347** [0.097]
θ_n^u	0.234** [0.062]	0.255** [0.062]	0.320** [0.078]	0.339** [0.106]
θ_k^b	1.056** [0.086]	1.140** [0.087]	0.817** [0.105]	1.446** [0.159]
θ_n^b	−1.122** [0.108]	−1.225** [0.108]	−0.710** [0.136]	−1.454** [0.227]
θ_r		−0.270** [0.040]	−0.071 [0.049]	−0.766** [0.089]
δ_0	1.202** [0.059]	1.279** [0.074]	1.014** [0.102]	1.792** [0.120]
δ_k^u	−0.285** [0.031]	−0.289** [0.032]	−0.244** [0.039]	−0.161* [0.070]
δ_n^u	−0.547** [0.040]	−0.525** [0.041]	−0.495** [0.049]	−0.464** [0.084]
δ_k^b	0.359** [0.096]	0.396** [0.098]	0.310* [0.116]	0.310 [0.177]
δ_n^b	−1.190** [0.112]	−1.210** [0.115]	−1.049** [0.141]	−0.935** [0.218]
δ_r		−0.059* [0.025]	0.022 [0.027]	−0.434** [0.084]
RMSE	0.325	0.323	0.322	
$\ln L$	3192.3	3207.9	3231.0	
LR	77.4**	46.2**		
K	8	10	20	

Maximum-likelihood estimates of parameters in equations (12) and (13). White's (1982) robust standard errors between brackets. Sample size is 3600 observations. RMSE is the root mean squared prediction error, $\ln L$ is the value of the likelihood function, *LR* is the likelihood ratio statistic comparing each Model with Model III, *K* is the number of parameters estimated in each model. Coefficients marked with ** (*) are significantly different from 0 at the 1% (5%) level.

for three estimates, all parameters are significant and have the same sign for both types of subjects. Like in the previous models, our risk measure is only significant and negative for non-*MON* subjects.

The estimated parameters in Table 10 imply that an increase in the number of firms decreases the expected prices of those firms indexed by biased search engines and increases the expected prices of firms that could be indexed by an unbiased search engine. While the former observations agrees with the theoretical predictions, the latter do not. The estimated effects of a change in the index size are in line with the theoretical predictions: an increase of k_i rises the expected price in biased treatments and lowers the expected price in unbiased treatments.

An advantage of our estimation procedure is that we can extrapolate the results in order to predict behavior under different conditions. An example is presented in Tables 11 and 12, where we provide predictions for two interesting scenarios: risk neutrality and "average risk aversion", corresponding to the most frequent choice in the parallel experiment ($q = 0.5$).[28] For risk neutrality, results confirm the comments above: non-*MON* subjects set higher prices with a lower standard deviation. This is also valid for minimum prices. However, these results are often reversed when we calculate moments for risk averse subjects. In three of the eight treatments the means of average and minimum prices, evaluated at the average risk aversion degree, are higher for *MON* than for non-*MON* subjects.

The results in Table 11 also allow us to check the testable hypotheses of section 5.3.[29] A major finding is that some hypotheses are confirmed independently from the subjects' attitudes towards risk. Specifically, hypotheses **HU1 (i)**, **HB2** and **HU2** are confirmed in all cases. The **HC** hypothesis is confirmed strongly for *MON* subjects and weakly for non-*MON* ones. A second set of hypotheses are those which are confirmed by a certain subject type alone. Specifically, **HB1 (i)** is confirmed by non-*MON* subjects and **HB1 (ii)** is confirmed by *MON* subjects. Also, the **HG (i)** hypothesis is not confirmed by *MON* subjects. Finally, **HU1 (ii)** is not confirmed by any type of subjects.

An overview of these tables shows that some general features like, for example our consistency hypothesis, are supported in the majority of the cases. This implies that the predictions of models like Varian (1980), derived under assumptions of risk neutrality are robust enough to survive in the presence of risk-averse subjects. Other results require more qualified testing in that the existence of

[28] Once an estimation of the distribution of prices is obtained, it is straightforward to estimate the mean and the variance of average prices, $\epsilon = \sum p_i/n$, as long as $E(\epsilon) = E(p)$ and $\text{Var}(\epsilon) = \text{Var}(p)/n$. The distribution function for the minimum of n prices is $1 - (1 - F(\cdot))^n$, where F is the cumulative distribution function of p. Using this result, we can calculate the mean and the variance of minimum prices.

[29] Appendix C offers an exhaustive map of predictions concerning our testable hypotheses for different values of n and k.

TABLE 11: Means of Price Distributions

Treatment	Average prices		Minimum prices	
	MON	non-*MON*	*MON*	non-*MON*
(a) Risk neutrality				
1	0.707	0.885	0.436	0.754
2	0.725	0.899	0.568	0.830
3	0.634	0.811	0.447	0.704
4	0.722	0.885	0.230	0.607
5	0.739	0.898	0.371	0.716
6	0.652	0.810	0.241	0.543
7	0.767	0.918	0.613	0.856
8	0.515	0.611	0.288	0.432
(b) Average risk aversion				
1	0.683	0.693	0.407	0.432
2	0.701	0.722	0.539	0.570
3	0.608	0.556	0.418	0.369
4	0.699	0.691	0.206	0.224
5	0.717	0.721	0.342	0.365
6	0.625	0.555	0.215	0.169
7	0.746	0.766	0.585	0.618
8	0.486	0.314	0.260	0.123

Means of average and minimum price distributions calculated from the results of last two columns of Table 10.

TABLE 12: Standard Deviations of Price Distributions

Treatment	Average prices		Minimum prices	
	MON	non-*MON*	*MON*	non-*MON*
(a) Risk neutrality				
1	0.305	0.159	0.290	0.193
2	0.291	0.147	0.300	0.173
3	0.334	0.209	0.317	0.220
4	0.340	0.186	0.253	0.238
5	0.326	0.172	0.311	0.229
6	0.371	0.237	0.281	0.257
7	0.302	0.150	0.331	0.185
8	0.405	0.315	0.342	0.290
(b) Average risk aversion				
1	0.309	0.292	0.281	0.268
2	0.295	0.278	0.297	0.283
3	0.336	0.326	0.310	0.283
4	0.346	0.329	0.235	0.221
5	0.332	0.314	0.298	0.283
6	0.375	0.365	0.263	0.213
7	0.308	0.288	0.330	0.312
8	0.400	0.357	0.327	0.221

Standard deviations of average and minimum price distributions calculated from the results of last two columns of Table 10.

MON-incompatible behavior and lack of neutrality towards risk induces two directions in which observed behavior may deviate from theory. First, a subject alone may not be risk-neutral. Second, even *MON*-compatible risk neutral subjects could be interacting with other types of agents whose behavior -if anticipated by the former- induces everybody in the market act contrary to what the theory predicts. We could summarize and attempt a synthesis of the preceding results, observing that generally speaking, failures of the theoretical model to organize observed behavior relate to *Unbiased Incomplete Coverage*, obviously, because this setting implies an extra source of uncertainty (whether one will be actually indexed or not) for interacting sellers.

6 Conclusion

In this paper we have proposed a theoretical model on the effects of incompleteness and bias of price-comparison search engines on the market outcome. We tested the theoretical predictions in a laboratory experiment whose design is specific to the testable hypotheses emerging from the model. Furthermore, we implemented a parallel experiment, designed to capture two different aspects of our subjects' attitude towards risk: the degree of risk aversion and wheter subjects accept riskier options in exchange for a higher return. The experimental data were used for the estimation of two versions of a specific econometric model. The first version accounts for the fixed effects corresponding to the eight experimental treatments through dummy variables. The second version accounts for the design variables in their quantitative form. Both models are significantly improved when two risk related variables are included as idiosyncratic subject specific shocks. Further hypotheses are tested by simulation results obtained on the estimates of the econometric model.

The theoretical model warns us on possible counterintuitive effects of unbiased coverage of the market on observed prices, e.g., a larger number of firms whose prices are included in the engine may not imply globally lower expected prices. However, experimental results contradict some of the theoretical predictions both in quantitative and qualitative terms, especially, predictions on incomplete coverage. An interesting empirical finding reported here concerns an important improvement in the explanatory power of the empirical model of individual prices, which is obtained when we introduce the variables accounting for our subjects' attitudes towards risk. Specifically, splitting the sample of subjects into two groups according to whether they comply or not with the monotonicity hypothesis improves the explanatory power of the empirical model. Furthermore, a significant negative effect of risk aversion on prices is estimated for those not fulfilling monotonicity. This and the risk-related aspects of our subjects' behavior explain the systematic clock-wise rotation of observed price distributions as compared to the theoretical ones.

Our findings indicate that behavioral factors are important in markets like the ones studied here. Therefore, models assuming monotonicity or risk neutrality may produce predictions which deviate from behavior by real agents.

References

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

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

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

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

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

BURDETT, K. AND K. JUDD (1983): "Equilibrium Price Dispersion." *Econometrica*, vol. 51, pp. 955–969.

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

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

DIAMOND, P. (1971): "A Model of Price Adjustment." *Review of Economic Theory*, vol. 3, pp. 156–178.

DOORNIK, J. A. (2002): *Object-Oriented Matrix Programming using Ox*. Timberlake Consultants Press, London, 3rd ed.

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.

GEORGANTZÍS, N., M. GENIUS, A. GARCÍA-GALLEGO, AND G. SABATER (2003): "Testing for Risk-Return Tradeoffs in Individual Decision Making." Working Paper 43/03, LINEEX, Valencia.

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

HOLT, C. A. AND S. K. LAURY (2002): "Risk Aversion and Incentive Effects." *American Economic Review*, vol. 92, pp. 1644–1655.

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

KEPHART, J. AND A. GREENWALD (1999): *Price-Comparison Site Economics*. IBM Institute for Advanced Commerce.

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

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

MONTGOMERY, A., K. HOSANAGAR, R. KRISHNAN, AND K. CLAY (2001): *Designing a Better Price-Comparison Site*. School of Industrial Administration.

MORGAN, J., H. ORZEN, AND M. SEFTON (2004): "An Experimental Study of Price Dispersion." *Games and Economic Behavior*. Forthcoming.

PEREIRA, P. (2004a): "Do Lower Search Costs Reduce Prices and Price Dispersion?" *Information Economics and Policy*. Forthcoming.

PEREIRA, P. (2004b): *Search and Destroy: Competition between Search Engines and Consumer Prices*. Portuguese Competition Authority.

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

SABATER, G. AND N. GEORGANTZÍS (2002): "Accounting for Risk Aversion in Repeated Prisoners' Dilemma Games." *Journal of Economic Behavior and Organization*, vol. 48, pp. 37–50.

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

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

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

WHITE, H. (1982): "Maximum Likelihood Estimation of Misspecified Models." *Econometrica*, vol. 53, pp. 1–16.

Appendix A Proofs

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

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

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 fom (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 $l_j^\tau \leq p_l < p_h \leq 1$ such that $F(p_l) = F(p_h)$. Suppose also that p_l is the infimum of all prices p, $l_j^\tau \leq p \leq 1$. 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. □

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 defintion 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. (2004); **(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 on economic phenomena. 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 can be characterized as follows:

  (a) a price comparison search engine like the ones on the Internet,

  (b) 3 firms, (**Treatments 4–8:** 6 firms),

  (c) 1 200 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).

- Each session will consists of 50 rounds.

- You are one of the 3 firms (**Treatmens 4–8:** 6 firms) in the industry. Your production costs are zero. Therefore, your profits are equal to your income.

- 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, the *Price Search Engine* lists the price of *all* firms in the industry.

- (**Treatments 2, 5 and 7**) Each period, the *Price Search Engine* lists the price 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' price decision making.

- (**Treatments 3, 6 and 8**) Each period, the *Price Search Engine* lists the price 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' price decision making.

- 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 1 000 UMEX. That is, if the price you fix is higher than 1 000 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 1 000 UMEX. In case of a "tie" (i.e., several firms fix the same minimum price) the consumers are distributed equitatively among the firms with the same minimum price.

- The consumers that do not read the search engine's price list will buy "randomly" from any vendor, so that the total demand of this group of consumers will be distributed equitatively among all firms in the industry.

- In each round, the other 2 firms (**Treatments 4–8:** 5 firms) which (together with you) form the industry, will be randomly assigned among all participants. 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 decide 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: 1 000 000 UMEX = 10 €.

Thank you very much for your participation. Good luck!

Appendix C Additional Results

TABLE 13: Means of price distributions (*MON*, unbiased)

(a) Average prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.725	0.707					
$n=4$	0.743	0.726	0.714				
$n=5$	0.756	0.740	0.728	0.719			
$n=6$	0.767	0.751	0.739	0.730	0.722		
$n=7$	0.776	0.760	0.749	0.740	0.732	0.726	
$n=8$	0.783	0.768	0.757	0.748	0.741	0.734	0.729

(b) Minimum prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.568	0.436					
$n=4$	0.586	0.451	0.351				
$n=5$	0.601	0.464	0.362	0.284			
$n=6$	0.613	0.476	0.371	0.291	0.230		
$n=7$	0.623	0.485	0.380	0.298	0.235	0.187	
$n=8$	0.632	0.494	0.387	0.304	0.240	0.190	0.152

(c) Average prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.701	0.683					
$n=4$	0.720	0.703	0.690				
$n=5$	0.735	0.717	0.705	0.695			
$n=6$	0.746	0.729	0.717	0.707	0.699		
$n=7$	0.755	0.739	0.727	0.717	0.709	0.702	
$n=8$	0.763	0.747	0.735	0.726	0.718	0.711	0.705

(d) Minimum prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.539	0.407					
$n=4$	0.558	0.422	0.324				
$n=5$	0.573	0.434	0.333	0.258			
$n=6$	0.585	0.445	0.342	0.264	0.206		
$n=7$	0.595	0.455	0.350	0.270	0.211	0.165	
$n=8$	0.605	0.464	0.357	0.276	0.215	0.168	0.132

Means of average and minimum price distributions calculated from the estimates of Table 10.

TABLE 14: Means of price distributions (*MON*, biased)

(a) Average prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.634	0.707					
$n=4$	0.586	0.663	0.714				
$n=5$	0.547	0.627	0.680	0.719			
$n=6$	0.515	0.596	0.652	0.692	0.722		
$n=7$	0.488	0.570	0.626	0.668	0.700	0.726	
$n=8$	0.464	0.547	0.604	0.647	0.680	0.707	0.729

(b) Minimum prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.447	0.436					
$n=4$	0.379	0.359	0.351				
$n=5$	0.328	0.300	0.289	0.284			
$n=6$	0.288	0.256	0.241	0.234	0.230		
$n=7$	0.256	0.220	0.203	0.194	0.189	0.187	
$n=8$	0.230	0.192	0.173	0.163	0.157	0.154	0.152

(c) Average prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.608	0.683					
$n=4$	0.558	0.637	0.690				
$n=5$	0.519	0.600	0.655	0.695			
$n=6$	0.486	0.569	0.625	0.667	0.699		
$n=7$	0.459	0.542	0.599	0.642	0.675	0.702	
$n=8$	0.436	0.518	0.576	0.620	0.654	0.682	0.705

(d) Minimum prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.418	0.407					
$n=4$	0.349	0.330	0.324				
$n=5$	0.299	0.272	0.262	0.258			
$n=6$	0.260	0.229	0.215	0.209	0.206		
$n=7$	0.230	0.195	0.179	0.171	0.167	0.165	
$n=8$	0.205	0.168	0.151	0.141	0.136	0.134	0.132

Means of average and minimum price distributions calculated from the estimates of Table 10.

Table 15: Means of price distributions (non-*MON*, unbiased)

(a) Average prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.899	0.885					
$n=4$	0.907	0.895	0.885				
$n=5$	0.914	0.902	0.893	0.885			
$n=6$	0.918	0.907	0.898	0.891	0.885		
$n=7$	0.922	0.911	0.903	0.896	0.890	0.885	
$n=8$	0.925	0.915	0.907	0.900	0.894	0.889	0.885

(b) Minimum prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.830	0.754					
$n=4$	0.841	0.766	0.700				
$n=5$	0.849	0.775	0.709	0.652			
$n=6$	0.856	0.782	0.716	0.659	0.607		
$n=7$	0.861	0.788	0.723	0.665	0.613	0.567	
$n=8$	0.866	0.794	0.729	0.670	0.618	0.571	0.529

(c) Average prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.722	0.693					
$n=4$	0.741	0.713	0.692				
$n=5$	0.755	0.728	0.708	0.692			
$n=6$	0.766	0.740	0.721	0.705	0.691		
$n=7$	0.776	0.750	0.731	0.715	0.702	0.691	
$n=8$	0.783	0.759	0.740	0.725	0.712	0.701	0.691

(d) Minimum prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.570	0.432					
$n=4$	0.590	0.448	0.346				
$n=5$	0.605	0.461	0.356	0.278			
$n=6$	0.618	0.473	0.365	0.285	0.224		
$n=7$	0.628	0.483	0.373	0.291	0.228	0.181	
$n=8$	0.638	0.492	0.381	0.296	0.233	0.184	0.146

Means of average and minimum price distributions calculated from the estimates of Table 10.

TABLE 16: Means of price distributions (non-*MON*, biased)

(a) Average prices, risk neutrality

	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$
$n = 3$	0.811	0.885					
$n = 4$	0.739	0.836	0.885				
$n = 5$	0.672	0.786	0.848	0.885			
$n = 6$	0.611	0.738	0.810	0.855	0.885		
$n = 7$	0.556	0.693	0.774	0.825	0.860	0.885	
$n = 8$	0.508	0.650	0.738	0.795	0.835	0.863	0.885

(b) Minimum prices, risk neutrality

	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$
$n = 3$	0.704	0.754					
$n = 4$	0.602	0.661	0.700				
$n = 5$	0.510	0.574	0.619	0.652			
$n = 6$	0.432	0.495	0.543	0.579	0.607		
$n = 7$	0.365	0.425	0.473	0.511	0.542	0.567	
$n = 8$	0.308	0.363	0.410	0.449	0.481	0.507	0.529

(c) Average prices, average risk aversion

	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$
$n = 3$	0.556	0.693					
$n = 4$	0.452	0.597	0.692				
$n = 5$	0.374	0.517	0.619	0.692			
$n = 6$	0.314	0.451	0.555	0.632	0.691		
$n = 7$	0.268	0.397	0.499	0.579	0.642	0.691	
$n = 8$	0.232	0.351	0.451	0.531	0.596	0.648	0.691

(d) Minimum prices, average risk aversion

	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$	$k = 7$	$k = 8$
$n = 3$	0.369	0.432					
$n = 4$	0.252	0.300	0.346				
$n = 5$	0.174	0.206	0.243	0.278			
$n = 6$	0.123	0.142	0.169	0.197	0.224		
$n = 7$	0.089	0.098	0.117	0.138	0.160	0.181	
$n = 8$	0.066	0.069	0.081	0.096	0.113	0.130	0.146

Means of average and minimum price distributions calculated from the estimates of Table 10.

TABLE 17: Standard deviations of price distributions (*MON*, unbiased)

(a) Average prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.291	0.305					
$n=4$	0.297	0.311	0.321				
$n=5$	0.300	0.314	0.324	0.332			
$n=6$	0.302	0.316	0.326	0.334	0.340		
$n=7$	0.304	0.318	0.328	0.336	0.342	0.347	
$n=8$	0.305	0.319	0.329	0.337	0.343	0.348	0.352

(b) Minimum prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.300	0.290					
$n=4$	0.314	0.307	0.284				
$n=5$	0.324	0.321	0.299	0.270			
$n=6$	0.331	0.331	0.311	0.283	0.253		
$n=7$	0.337	0.340	0.321	0.293	0.263	0.233	
$n=8$	0.342	0.348	0.330	0.303	0.272	0.242	0.213

(c) Average prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.295	0.309					
$n=4$	0.302	0.315	0.325				
$n=5$	0.306	0.320	0.329	0.337			
$n=6$	0.308	0.322	0.332	0.340	0.346		
$n=7$	0.310	0.324	0.334	0.342	0.348	0.353	
$n=8$	0.312	0.326	0.336	0.343	0.349	0.355	0.359

(d) Minimum prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.297	0.281					
$n=4$	0.311	0.299	0.271				
$n=5$	0.322	0.312	0.286	0.255			
$n=6$	0.330	0.324	0.298	0.267	0.235		
$n=7$	0.337	0.333	0.309	0.277	0.245	0.214	
$n=8$	0.342	0.341	0.318	0.287	0.254	0.222	0.193

Standard deviations of average and minimum price distributions calculated from the estimates of Table 10.

TABLE 18: Standard deviations of price distributions (*MON*, biased)

(a) Average prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.334	0.305					
$n=4$	0.367	0.342	0.321				
$n=5$	0.390	0.369	0.349	0.332			
$n=6$	0.405	0.389	0.371	0.355	0.340		
$n=7$	0.416	0.404	0.389	0.374	0.360	0.347	
$n=8$	0.424	0.415	0.402	0.389	0.376	0.364	0.352

(b) Minimum prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.317	0.290					
$n=4$	0.335	0.304	0.284				
$n=5$	0.341	0.307	0.286	0.270			
$n=6$	0.342	0.305	0.281	0.265	0.253		
$n=7$	0.340	0.299	0.273	0.256	0.243	0.233	
$n=8$	0.336	0.292	0.264	0.245	0.231	0.221	0.213

(c) Average prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.336	0.309					
$n=4$	0.367	0.345	0.325				
$n=5$	0.389	0.371	0.353	0.337			
$n=6$	0.403	0.390	0.375	0.360	0.346		
$n=7$	0.412	0.404	0.391	0.378	0.365	0.353	
$n=8$	0.419	0.415	0.404	0.392	0.381	0.369	0.359

(d) Minimum prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.310	0.281					
$n=4$	0.324	0.292	0.271				
$n=5$	0.328	0.292	0.270	0.255			
$n=6$	0.327	0.288	0.263	0.247	0.235		
$n=7$	0.324	0.281	0.254	0.236	0.224	0.214	
$n=8$	0.318	0.272	0.244	0.225	0.211	0.201	0.193

Standard deviations of average and minimum price distributions calculated from the estimates of Table 10.

TABLE 19: Standard deviations of price distributions (non-*MON*, unbiased)

(a) Average prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.147	0.159					
$n=4$	0.148	0.161	0.170				
$n=5$	0.149	0.162	0.171	0.179			
$n=6$	0.150	0.163	0.172	0.179	0.186		
$n=7$	0.150	0.163	0.172	0.180	0.186	0.192	
$n=8$	0.151	0.163	0.173	0.180	0.187	0.192	0.197

(b) Minimum prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.173	0.193					
$n=4$	0.178	0.202	0.213				
$n=5$	0.182	0.209	0.222	0.228			
$n=6$	0.185	0.214	0.229	0.236	0.238		
$n=7$	0.187	0.218	0.235	0.243	0.246	0.245	
$n=8$	0.189	0.222	0.240	0.249	0.252	0.252	0.249

(c) Average prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.278	0.292					
$n=4$	0.283	0.297	0.307				
$n=5$	0.286	0.301	0.311	0.319			
$n=6$	0.288	0.303	0.314	0.322	0.329		
$n=7$	0.289	0.305	0.316	0.325	0.331	0.337	
$n=8$	0.290	0.306	0.318	0.326	0.333	0.338	0.343

(d) Minimum prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.283	0.268					
$n=4$	0.296	0.284	0.257				
$n=5$	0.305	0.297	0.271	0.241			
$n=6$	0.312	0.307	0.283	0.252	0.221		
$n=7$	0.318	0.316	0.293	0.262	0.231	0.201	
$n=8$	0.322	0.324	0.301	0.271	0.239	0.209	0.181

Standard deviations of average and minimum price distributions calculated from the estimates of Table 10.

TABLE 20: Standard deviations of price distributions (non-*MON*, biased)

(a) Average prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.204	0.159					
$n=4$	0.252	0.204	0.170				
$n=5$	0.288	0.242	0.206	0.179			
$n=6$	0.315	0.274	0.237	0.209	0.186		
$n=7$	0.335	0.300	0.265	0.236	0.211	0.192	
$n=8$	0.348	0.321	0.289	0.260	0.235	0.214	0.197

(b) Minimum prices, risk neutrality

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.220	0.193					
$n=4$	0.258	0.231	0.213				
$n=5$	0.279	0.256	0.240	0.228			
$n=6$	0.290	0.270	0.257	0.246	0.238		
$n=7$	0.292	0.276	0.266	0.258	0.251	0.245	
$n=8$	0.288	0.276	0.269	0.263	0.258	0.253	0.249

(c) Average prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.326	0.292					
$n=4$	0.351	0.335	0.307				
$n=5$	0.358	0.360	0.342	0.319			
$n=6$	0.357	0.372	0.365	0.348	0.329		
$n=7$	0.350	0.377	0.379	0.368	0.353	0.337	
$n=8$	0.341	0.377	0.387	0.382	0.372	0.358	0.343

(d) Minimum prices, average risk aversion

	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$
$n=3$	0.283	0.268					
$n=4$	0.271	0.261	0.257				
$n=5$	0.247	0.239	0.239	0.241			
$n=6$	0.221	0.211	0.213	0.217	0.221		
$n=7$	0.197	0.183	0.184	0.189	0.195	0.201	
$n=8$	0.175	0.158	0.157	0.162	0.168	0.175	0.181

Standard deviations of average and minimum price distributions calculated from the estimates of Table 10.