WORKING

PAPERS

Manfred Nermuth Giacomo Pasini Paolo Pin Simon Weidenholzer

Price Dispersion, Search Externalities, and the Digital Divide

December 2009

Working Paper No: 0916



DEPARTMENT OF ECONOMICS

UNIVERSITY OF VIENNA

All our working papers are available at: http://mailbox.univie.ac.at/papers.econ

Price Dispersion, Search Externalities, and the Digital Divide*

Manfred Nermuth[†] Giacomo Pasini[‡] Paolo Pin[§]

Simon Weidenholzer[¶]

December 2009

Abstract

We propose a model of price competition where consumers exogenously differ in the number of prices they compare. Our model can be interpreted either as a non–sequential search model or as a network model of price competition. We show that i) if consumers who previously just sampled one firm start to compare more prices all types of consumers will expect to pay a lower price and ii) if consumers who already sampled more than one price sample (even) more prices then there exists a threshold –the digital divide– such that all consumers comparing fewer prices than this threshold will expect to pay a higher price whereas all consumers comparing more prices will expect to pay a lower price than before.

JEL Classification: D43, D85, L11

^{*}The authors wish to thank Larry Blume, Giacomo Calzolari, Andrea Galeotti, Maarten Janssen, Saul Lach, Marco van der Leij, José Luis Moraga-González and Mario Padula for useful comments on earlier drafts of this paper.

[†]Institut für Volkswirtschaftslehre, Universität Wien. Email: manfred.nermuth@univie.ac.at

[‡]Dipartimento di Scienze Economiche, Università di Venezia.

Email: giacomo.pasini@unive.it

[§]Dipartimento di Economia Politica, Università degli Studi di Siena. Email: pin3@unisi.it

[¶]Institut für Volkswirtschaftslehre, Universität Wien. Email: simon.weidenholzer@univie.ac.at

1. Introduction

What we observe in daily life is that frequently homogenous goods are sold at different prices by different firms, i.e. we observe price dispersion. In a recent survey Baye, Morgan, and Scholten (2006) report on dozens of empirical papers providing ample evidence in both on- and off-line consumer good markets¹ for Varian's (1980) assertion that actually "the law of one price is no law at all".

The classic explanation for persistent price dispersion is provided by models where firms compete on prices and consumers differ (either exogenously or endogenously) in the number of prices they compare. Following the works of Wilde and Schwartz (1979) and Varian (1980), in many of these models there are only two types of consumers:² A fraction of consumers – the informed consumers– compare all prices and buy from the firm offering the lowest price. The remaining fraction –the uninformed consumers– just sample one price at random. The presence of these two groups of consumers implies that firms face a tradeoff between setting low prices in order to attract a high number of informed consumers and setting high prices to rip off the uninformed consumers. These two forces are balanced in a mixed strategy equilibrium where firms randomize on prices, which can be interpreted as the theoretical equivalent of the empirical observation of price dispersion.

In models of consumer search the division of consumers in uninformed consumers and informed consumers arises endogenously as consumers decide on the optimal number of prices to sample. Burdett and Judd (1983) present a model where consumers have homogeneous search costs and engage in optimal non–sequential search. In this setup consumers either find it optimal to sample only one price or randomize between sampling one and two prices. In the latter case consumers endogenously divide themselves into uninformed and (partially) informed consumers. Stahl (1989) presents a model of sequential search where a fraction of consumers has zero search costs and samples all firms in the market, and the remaining fraction has positive search costs and samples only one firm in the market. Janssen and Moraga-González (2004) analyze the case of non–sequential search in the presence of consumers with zero search costs

¹The recent rise of internet price comparison sites has not only eased consumer search but has made it also much easier for researcher to collect data, hence the second wave in the literature.

²The survey by Baye, Morgan, and Scholten (2006) reports on many more theoretical papers on price dispersion and consumer search.

and consumers with positive search costs. Naturally, consumers with zero search costs will always sample all firms in the market. Depending on the search cost, consumers with positive search will either randomize between not sampling (and not buying) and sampling one price (and buying), or sample one price, or randomize between sampling one price and two prices.

In the present paper we study a model that allows for any degree of partially informed consumers. Some consumers will only be visiting one firm, some consumers will be comparing the prices of two firms, some of three firms, and so forth, obtaining a *consumer search distribution*. Burdett and Judd (1983) already show existence of a mixed strategy equilibrium under such a general consumer search distribution. In their model consumers with homogeneous search costs find it optimal to either visit one firm or to randomize between visiting one and two firms.³ We assume that the consumer search distribution is exogenously given and remark that *any* consumer search distribution can be justified as the result of consumers with sufficiently heterogenous search costs engaging in optimal non–sequential search.

The first advantage of looking at more general consumer search distributions is that we are able to provide new insights into the structure of price dispersed equilibria by varying the consumer search distribution. We find that the only determinant for profits and hence also for total consumer welfare is the share of uninformed consumers in the market. That is neither the number of firms nor the exact distribution of the informed consumers matters. For, a firm can always set the monopoly price (i.e. the consumers' valuation for the good) in which case it only sells to uninformed consumers. As in a mixed strategy equilibrium all prices in the support of the equilibrium price distribution have to yield the same expected profits it follows that these expected profits will only depend on the share of uninformed consumers.

In line with the previous literature on consumer search we find that lowering the fraction of uninformed consumers tightens competition and lowers expected prices for all different types of consumer. One might be tempted by this result to think that "more search" is always beneficial to *all consumers*. We show that this is not the case if already informed consumers start to compare more prices. In particular we show that if already informed consumers start to compare even more prices then there exists a threshold – *the digital divide* – such that everybody who is

³ Burdett and Judd (1983) also analyze the case of noisy search, where consumers decide whether to receive an unknown number of price observations. See Armstrong, Vickers, and Zhou (2009) for a model of noisy search analyzing the welfare effects of price caps and advertisement.

comparing fewer prices than this threshold will expect to pay a higher price whereas everybody comparing more prices will expect to pay lower prices than before. In this sense, more search by informed consumers inflicts a negative (pecuniary) *search externality* upon relatively poorly informed consumer whereas it causes a positive search externality for relatively well informed consumers. Figure 1 illustrates this point by means of an example.

A further advantage of considering general consumer search distributions is that it allows us to put forward an alternative interpretation of the search model at hand as a model of price competition on a network in the following sense:⁴ Firms and consumers represent nodes in a bipartite network, i.e. a network where there are only links between firms and consumers. A link between a consumer and a firm in this network indicates that a consumer observes the price of this firm and may buy from this firm and that the firm may sell to this consumer. If we assume that firms only know the degree distribution of consumers, i.e. the the probabilities that a given consumer has one link, two links, and so forth, we obtain a network game of incomplete information.⁵ Alternatively, we could also assume that the network is symmetric in the sense that every firm faces the same consumer degree distribution and analyze this game of complete information. Note that the number of potential consumers a firm can attract (i.e. its own degree) may be different across firms, though.

Within this network interpretation the consumer degree distribution essentially plays the role of the consumer search distribution in the consumer search model. In this sense, the equilibrium we find in the search context and the comparative static exercise of varying the consumer search distribution translate into our network context. In particular, the comparative static exercise of introducing more search can be interpreted as increasing the density of the network. We then have that adding links to consumers who are only linked to one firm decreases expected prices for all consumers, whereas adding links to consumers with already more than one link amounts to a redistribution of social welfare from consumers who just have a few links to well connected consumers.

This second interpretation of our model is related to the literature of competition on networks.

⁴Galeotti (2009) presents a model of search and word of mouth communication, where a network determines the communication process.

⁵This incomplete information setup stipulates the use of the Bayesian-Nash equilibrium concept for networks, as e.g. in Jackson and Yariv (2007) or Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2009).

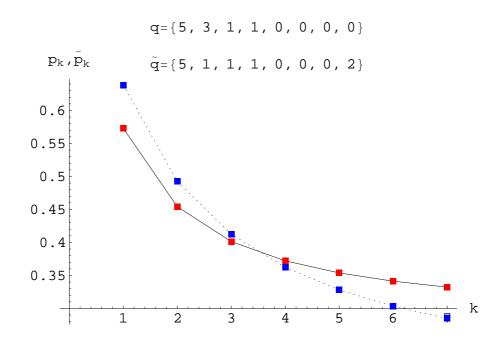


Figure 1: The effect of already informed consumers engaging in even more search on the expected prices paid by different types of consumers. The *k*-th component q_k of the 'search distribution' $q = (q_1, \ldots, q_8)$ is proportional to the number of consumers sampling exactly *k* prices; and p_k denotes the expected price paid by such a consumer. \tilde{q} is obtained from q if two thirds of the consumers who previously sampled two firms change to sampling eight firms. The solid line connects the expected prices p_k under the original distribution q and the dashed line connects the expected prices \tilde{p}_k under \tilde{q} . After this shift consumers who sample three or fewer prices will expect to pay higher prices while consumers sampling more than three prices expect to pay a lower price.

Kranton and Minehart (2001) and Corominas-Bosch (2004) analyze bargaining situations between fully informed agents on exogenously given non–regular networks. These models however become soon very complex, untractable and with multiple equilibria, as the size of the network grows. Blume, Easley, Kleinberg, and Tardos (2009) present instead a model with uniqueness of equilibria, with sellers, buyers, and traders, where traders play the role of mediators and may set different prices to different buyers and sellers. Lever (2008) studies duopolistic price competition under complete information on a network, where of course consumers may at most have degree two.

The rest of the paper is organized in the following way: Section 2 spells out the model. Section 3 presents the results and section 4 concludes. Proofs are in the Appendix.

2. The Model

We consider a market for a homogenous good with *N* firms and *M* households (consumers), and write $\mu = M/N$ for the number of households per firm. Each firm can produce the good at constant marginal cost, without fixed cost, and sets the price at which it offers the good (all firms set their prices simultaneously). Each household demands one unit of the good, up to a given willingness to pay (assumed greater than the cost). Without loss of generality we normalize the cost to 0 and the willingness to pay to 1 (the same for all firms, respectively households).

Households differ in the *information* they have about the firms' prices. A household of *type k* observes the prices of *k* firms and buys from the cheapest (randomizing with equal probabilities in case of ties), provided the price does not exceed its willingness to pay. We denote by q_k the fraction of households of type *k*. The *information structure* is represented by the vector $q = (q_1, \dots, q_N)$, where of course $q_k \ge 0$ for all *k* and $\sum_{k=1}^N q_k = 1$. We also refer to *q* as the consumer *search distribution* or the *degree distribution*.

As already mentioned in the Introduction, there are two interpretations for this kind of model: (1) *non–sequential search:* the *k* firms which a household of type *k* observes are a random sample (with equal probabilities) from the set of all *N* firms; and (2) a *bipartite network* of firms and households: a household and a firm can trade only if they are connected by a link. A household of type k (degree k) has links to k different firms, and each firm faces the same degree distribution among its potential customers (i.e. the households to which it is linked). One justification of this particular kind of symmetry among firms is that the network itself might be symmetric in the sense that each firm faces the same degree distribution but may in principle attract a different number of potential consumers. Alternatively it could also be the case that firms do not know the degree of each consumer they are linked to but instead only know the degree distribution in the overall network.⁶ This incomplete information setup seems to be a realistic approximation of situations where firms normally do not know the shopping habits of each individual consumer, but rely their strategy on survey data representing the shopping habits of average customers instead.

We will use both – the search and the network – interpretation interchangeably; in either case, we obtain a strategic market game among the N firms, where we can take without loss of generality the strategy set of each firm to be the unit interval [0,1] (a price below 0 would generate losses, and at a price above 1 nobody would buy). Trivial cases apart, we will see that this game has equilibria only in mixed strategies, thus generating *price dispersion*. Our main interest is the dependence of these price dispersion equilibria, and hence consumer welfare, on the information structure q.

First we introduce some more terminology and notation. A household of type k = 1 is called *uninformed* (or *locked in* in the network case), households of types $k \ge 2$ are called *informed* (also searchers or shoppers). The average number of searches (or links) per household is $\kappa = E_q[k] = \sum_k kq_k$; it is a measure for the intensity of search in the market (density of the network); the number of links per firm is $\mu \kappa$. For $q = (q_1, \dots, q_N)$, define the auxiliary functions

$$\varphi(x) := \sum_k kq_k x^{k-1}$$
 for $x \in [0,1]$

and if $q_1 > 0$

$$\psi(x) := \frac{1}{q_1} \varphi(x) \quad \text{for} \quad x \in [0, 1]$$

We have $\psi(0) = 1$, $\psi(1) = \kappa/q_1$, and ψ is continuous and strictly increasing in *x*. The following known results are stated here for easy reference.

⁶See footnote 5.

Proposition 2.1. A market game with information structure $q = (q_1, ..., q_N)$ has the following equilibria:

- (a) if all households are uninformed $(q_1 = 1)$, the only equilibrium is the monopoly outcome, *i.e.* all firms charge the households' willingness to pay (= 1).
- (b) if all households search $(q_1 = 0)$, the only equilibrium (in which all firms are active) is the competitive outcome, i.e. all firms charge the competitive price (= marginal cost = 0).
- (c) If $0 < q_1 < 1$, there is no equilibrium in pure strategies, but there exists a unique symmetric equilibrium in mixed strategies: each firm chooses its price at random according to a continuous distribution F(p) with support $[p_{min}, p_{max}]$, where

$$0 < p_{min} = \frac{1}{\psi(1)} = \frac{q_1}{\kappa} < 1, \qquad p_{max} = \frac{1}{\psi(0)} = 1$$

and

$$F(p) = 1 - \psi^{-1}\left(\frac{1}{p}\right) \quad for \quad p_{min} \le p \le p_{max}$$

(and of course F(p) = 0 for $p \le p_{min}$, F(p) = 1 for $p \ge p_{max}$). Moreover, the equilibrium profit per firm is $\pi = \mu q_1$, and the average selling price (average household expenditure) is $p_{av} = q_1$.

These results are known at least since Burdett and Judd (1983) (see especially the proof of Lemma 2 there); for the reader's convenience, we give a brief sketch of the proof of Proposition 2.1(c) in the Appendix. From now on, we focus on price dispersion equilibria and assume always $0 < q_1 < 1$. The model is capable of generating a large variety of price distributions: the density f(p) = F'(p) can be increasing, decreasing, constant, U-shaped, bell-shaped, W-shaped, etc., depending on the information structure q. Fig. 2 exhibits a few examples⁷ (cf. also Example 2.3).

Remark 1. The price distribution F does not depend on the numbers M, N of households and firms, but only on the function ψ , i.e. (the nonzero components of) the search distribution q.

⁷ In labeling this and some later figures, we give 'search distributions' q that are not normalized to sum to one; replacing such q by $q/\sum_k q_k$ converts them into probability vectors. Note also that the function ψ is homogeneous of degree zero in q, so that the normalization of q does not matter in the formula for the price distribution F.

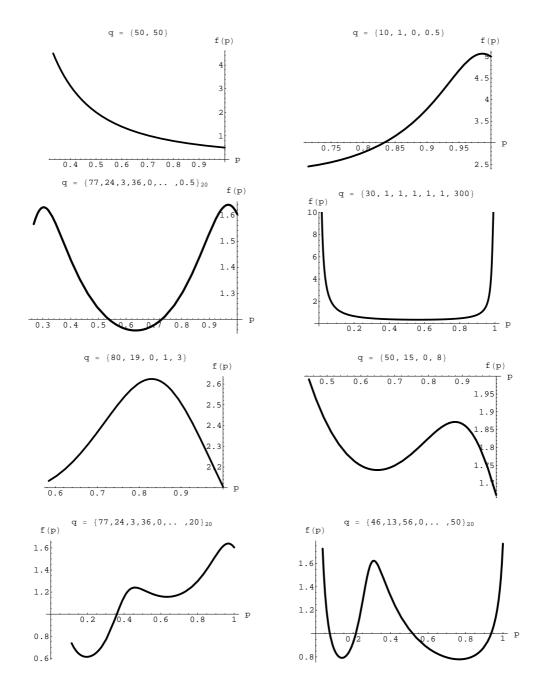


Figure 2: Some densities f(p) for various $q = (q_1, ..., q_K)$, with support $[p_{min}, 1]$. The label above each picture represents the search distribution (without the normalization $\sum_k q_k = 1$). Dots "..." represent omitted zeros; in this case the length K of q is indicated by a subscript (e.g. K = 20 in the last picture).

The effective size of the market is given by $K = \max\{k | q_k > 0\}$, the highest type actually occurring in the population. Thus for a town with N = 50 gas stations, with q such that nobody compares the prices of more than K = 10 stations, this model predicts the same equilibrium price distribution as if there were only N = 10 firms (only the profit per firm will be lower for larger N). Intuitively, since every customer of a firm observes at most K - 1 other prices, each firm has only K - 1 effective competitors, and the model behaves like an oligopoly with K firms. Thus, for given q (hence given K), we may take the total number N of firms as arbitrary (as long as $N \ge K$). Of course, this conceptual separation between changes in N and changes in q, while analytically useful, is somewhat artificial: in practice, an increase in the number of firms may lead to a change in search behavior.⁸ This is often assumed in the literature. For example in models where there are households with zero search costs who always sample *all* firms, we must have K = N, and a change in N implies a corresponding change in q.

Remark 2. The average selling price in the market depends only on q_1 , the share of uninformed consumers. Therefore, changes in the search behavior of the shoppers can only have redistributive effects (benefitting some types and hurting others), as long as q_1 remains constant. One can ask what the precise form of these redistributive effects is. Theorem 3.2 below gives a partial answer. At this point we note only that it is hard to predict in general: for example, the change from $q = \frac{1}{60}(10, 27, 20, 0, 0, 3)$ to $\tilde{q} = \frac{1}{60}(10, 38, 0, 0, 0, 1)$ changes the equilibrium price distribution in such a way that types k = 1, 2, 6 pay lower expected prices and types k = 3, 4, 5 pay higher expected prices (note that it is not clear whether q or \tilde{q} represents "more search"). See Figure 3.

The *expected price* p_k paid by a household of type k is given by

$$p_k = \int_{p_{min}}^1 p \, dF_k(p) \qquad (k = 1, \dots N)$$

where $F_k(p) = 1 - [1 - F(p)]^k$ is the distribution of the minimum of a sample of size k from the distribution F. It is easy to see⁹ that the expected price p_k can also be written $p_k = \int_{p_{min}}^1 [1 - F(p)]^k dp$. This implies the well-known fact that p_k is a strictly decreasing, convex

⁸Lach and Moraga-González (2009) empirically analyze the effects of an increase in the number of gas stations in the Netherlands and find that all types of consumers benefit from such an increase.

 $^{{}^9}p_k = \int p \cdot F'_k(p) dp = \int pk[1 - F(p)]^{k-1}F'(p)dp = \int pk[1 - F(p)]^{k-1} dF(p) = \int [1 - F(p)]^k dp$, where the last equality follows by partial integration.

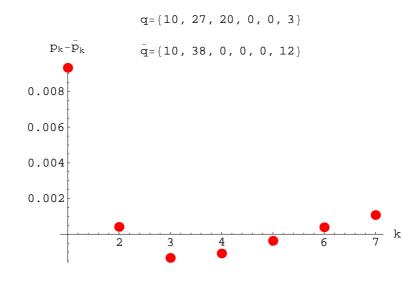


Figure 3: p_k resp. \tilde{p}_k is the expected price paid by type k under search distribution q resp. \tilde{q} . The picture shows the difference $p_k - \tilde{p}_k$. The change from q to \tilde{q} hurts the middle types.

function of k (cf. Burdett and Judd (1983), p. 961). We write $p = (p_1, \dots, p_N)$ for the list of expected prices paid by the various types. Note that if N > K, we have expected prices p_k even for some types k > K, although there are actually no such consumers. Such a p_k is simply what a (hypothetical) consumer would expect to pay if he sampled k prices from the given distribution F.

The next lemma will turn out to be very useful for the comparative statics exercise presented in the next section.

Lemma 2.2. The expected price of a consumer of type k is given by

$$p_k = \int_0^1 \frac{1}{\psi(x)} dx^k = \int_0^1 \frac{kx^{k-1}}{\psi(x)} dx.$$
 (1)

Proof: see Appendix.

Using this lemma, we can check that the average of the p_k 's is indeed equal to the average selling price $p_{av} = q_1$:

$$q \cdot p = \sum_{k} q_{k} p_{k} = \int_{0}^{1} \frac{\sum_{k} q_{k} k u^{k-1}}{\psi(u)} du = \int_{0}^{1} \frac{\varphi(u)}{\psi(u)} du = q_{1} = p_{av}$$

We conclude this section with an analytically tractable example (where formally $K = N = \infty$).

Example 2.3. Suppose the information structure q is such that $q_k = C \frac{\alpha^k}{k}$, for k = 1, 2, ..., with $0 < \alpha < 1$, and $C^{-1} = \sum_k \frac{\alpha^k}{k} = -\log(1-\alpha)$. Then $\psi(x) = (1-\alpha x)^{-1}$. By Proposition 2.1(c) the price distribution is

$$F(p) = \frac{1}{\alpha} [p - (1 - \alpha)] ,$$

with $p_{min} = 1 - \alpha$, $p_{max} = 1$, and the density $f(p) = F'(p) = \frac{1}{\alpha}$ is uniform on its support. The expected price for a household of type k is $p_k = 1 - \frac{k}{k+1}\alpha$, by (1), and the average selling price is $p_{av} = \alpha / \sum_k \frac{\alpha^k}{k} = 1 / \sum_k \frac{\alpha^{k-1}}{k}$. If α increases, also the probabilities that customers have more connections increase, and expected prices decrease.

3. The Digital Divide

The boundary between households who have access to the internet and those who do not is sometimes referred to as the "digital divide", and there is evidence that households above the divide (with many links) pay lower prices on average than those below (see e.g. Baye, Morgan, and Scholten (2003)). This is trivially true in our model, too, whereever we put the "divide", simply because the expected prices p_k decrease with k.

In what follows, we will demonstrate the existence of a much less obvious, but perhaps even more deplorable kind of divide (which, for lack of a better term, we also call "digital"): if some of those consumers who already search (have two or more links) begin to search even more (in a certain well-defined sense, see Definition 1 below), then the equilibrium price distribution changes in such a way that all types below a certain threshold (our "digital divide") face *higher* expected prices than before (in particular, the uninformed households always suffer), while those above the divide face *lower* prices than before (Theorem 3.2).

That is, (a certain way of) increasing the information in the market (making the network denser) favors only the high types, and harms the low types. Nobody here searches less than before, but the increased activity of some produces a negative (pecuniary) externality for the others (the low types who do not change their behavior). This is made precise in the following.

Consider two different search distributions $q = (q_1, ..., q_N)$ and $\tilde{q} = (\tilde{q}_1, ..., \tilde{q}_N)$, and denote the associated equilibrium quantities by $F, \tilde{F}, p_k, \tilde{p}_k$, etc.

Definition 1. \tilde{q} has fewer low types and more high types than q if there exists a threshold type ℓ , $1 \le \ell < N$, such that

$$\tilde{q}_k \leq q_k \text{ for } k < \ell, \quad \tilde{q}_\ell < q_\ell, \quad \tilde{q}_k \geq q_k \text{ for } k > \ell.$$

In this case, we write $\tilde{q} \succ_{\ell} q$ or simply $\tilde{q} \succ q$, and say that households *search more* under \tilde{q} than under q.

REMARK. The concept just introduced is stronger than first-order stochastic dominance, i.e. $\tilde{q} \succ q$ implies \tilde{q} FOSD q, but not vice versa. Here \tilde{q} FOSD q means that $\tilde{q}_1 + \cdots + \tilde{q}_k \leq q_1 + \cdots + q_k$ for all $k = 1, \ldots N$. It is easy to see that if \tilde{q} FOSD q and $\tilde{q} \neq q$, then $\tilde{\kappa} > \kappa$ and hence $p_{min} > \tilde{p}_{min}$. Therefore this last inequality is also implied by $\tilde{q} \succ q$.

We want to see what happens to the expected prices p_k if households search more in the sense of Definition 1, and consider the following two cases in turn. First, we consider the case when uninformed consumers start to compare more prices, and second, we consider the case when already informed consumers compare more prices.

Uninformed Consumers begin to search. Consider first the case that only the uninformed households (type 1) search more. This means that $\tilde{q}_1 < q_1$ and $\tilde{q}_k \ge q_k$ for all $k \ge 2$, i.e. $\tilde{q} \succ_1 q$. Here we find, in line with the previous literature:

Theorem 3.1. Assume $\tilde{q}_1 < q_1$ and $\tilde{q}_k \ge q_k$ for all $k \ge 2$, i.e. $\tilde{q} \succ_1 q$. Then $\tilde{p}_k < p_k$ for all k.

Proof. Obviously the assumption implies that $\tilde{q}_k/\tilde{q}_1 \ge q_k/q_1$ for all k, with strict inequality for some k, so that $\tilde{\psi}(x) > \psi(x)$ for all $x \in (0,1)$. This in turn implies $\tilde{p}_k < p_k$ for all k by equation (1).

That is, if some uninformed households begin to search, the expected prices for all types go down. The average selling price, $p_{av} = q_1$ also goes down. Intuitively, there is more search and the market becomes more competitive.

Informed Consumers search more. Consider next the case that only the informed households (types $k \ge 2$) search more. This means that $\tilde{q}_1 = q_1$ and $\tilde{q} \succ_{\ell} q$ for some $\ell \ge 2$. Here the situation is not so transparent: there is also "more search", but the average selling price $p_{av} = q_1 = \tilde{q}_1$ remains the same. This suggests (but does not prove) that not *all* prices p_k can go down. Our main result is the following.

Theorem 3.2. Assume $\tilde{q}_1 = q_1$ and $\tilde{q} \succ_{\ell} q$ for $\ell \ge 2$. Then there exists a number d, d > 1 such that $\tilde{p}_k > p_k$ for k < d and $\tilde{p}_k < p_k$ for k > d.

Proof: see Appendix.

We call *d* the *digital divide*. All types below the digital divide pay higher expected prices, and all types above it pay lower expected prices. In particular, the uninformed households (k = 1) always suffer because d > 1. For example, in Figure 4 we can take d = 10.5. It may be that d > K, i.e. there are actually no consumers above *d*. In this case, all types *k* of consumers with $q_k > 0$ will expect to pay a higher price, see example 3.3.

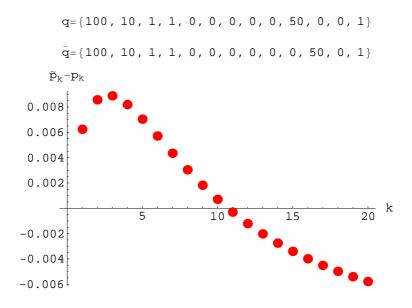


Figure 4: The Digital Divide: p_k resp. \tilde{p}_k is the expected price paid by type k under search distribution q resp. \tilde{q} . The picture shows the difference $\tilde{p}_k - p_k$. We have $\tilde{q}_1 = q_1$, $\tilde{q} \succ q$, and the change from q to \tilde{q} causes all prices p_k with k < d = 10.5 to increase, and all p_k with k > d = 10.5 to decrease. Note that the size of these price changes need not be monotonic: e.g. p_3 increases more than both p_1 and p_5 .

It is important to realize that firms' profits and hence also total consumer welfare depend only on the fraction of uniformed consumers. More search by already informed consumers does not affect total consumer welfare. However, it amounts to a redistribution of consumer welfare from relatively uninformed consumers to relatively informed consumers. In this sense, more search by informed consumers imposes a *negative externality* on relatively uninformed consumers and may impose a *positive externality* on highly informed consumers.

As noted above, we find that the number of firms does not affect consumer welfare either (as long as the search distribution q does not change). Increasing the number of firms just results in lower profits per firm, leaving the industry's profits and hence also consumer welfare unaffected. This is in contrast to previous work by Rosenthal (1980) who considers a model with uninformed and fully informed consumers and shows that if new firms enter the market all types of consumers will expect to pay a higher price. The driving force behind this surprising result is the assumption that each new firm brings new uninformed consumers to the market. As the number of firms increases also the fraction of uninformed consumers in the market increases and competition becomes less tough, like in our Theorem 3.1. Morgan, Orzen, and Sefton (2006) and Waldeck (2008) also consider models with uninformed and fully informed consumers and show that increasing the number of firms will decrease expected price for informed buyers and increase expected prices paid by uninformed consumers. The main reason behind this result is the assumption that the informed consumers always sample all firms in the market. As the number of firms increases also the number of prices sampled by the informed consumers increases. Morgan, Orzen, and Sefton (2006) and Waldeck (2008) show that in this setup uninformed consumers will expect to pay a higher price and the fully informed consumers will expect to pay a lower price, which can be readily interpreted in the light of Theorem 3.2.

We will now present an example where some consumers who previously compared two prices start to compare three prices.

Example 3.3. Consider the case where consumers only sample at most three firms, i.e. K = 3 and consider the consumer search distribution q = (5,5,10). Using Lemma 2.2 we can numerically compute the expected prices paid by the different types of consumers, obtaining $p_1 \simeq 0.3761$, $p_2 \simeq 0.2408$ and $p_3 \simeq 0.1915$. Suppose now that one consumer who previously sampled two firms samples now three firms, obtaining a new consumer search distribution $\tilde{q} = (5,4,11)$. Computing expected prices under this new consumer search distribution we now

obtain that $\tilde{p}_1 \simeq 0.3832 > \tilde{p}_1$, $\tilde{p}_2 \simeq 0.2434 > p_2$ and $\tilde{p}_3 \simeq 0.1919 > p_3$. In addition, we find that a consumer of type 4 would actually pay a lower price (i.e. $\tilde{p}_4 < p_4$), but there are no such consumers represented in the population.

At a first glance, the previous example seems to be in contradiction with the observation that as the number of uninformed consumers is the same under both consumer search distributions also the expected selling price has to be constant. This contradiction is however easily resolved if one takes into account that the consumers who switched from sampling two firms to three firms now pay a lower price, i.e. $\tilde{p}_3 < p_2$.¹⁰

It is clear from the above that if some shoppers search more they can exert a negative externality on other households. Perhaps surprisingly, it is even possible that all consumers who search more hurt themselves:

Example 3.4. starting from the initial search distribution q = (1, 20000, 0, 0, 0, 10000), suppose that all type 2 households switch to become type 3 (begin to compare three prices instead of two), resulting in the new search distribution $\tilde{q} = (1, 0, 20000, 0, 0, 10000)$. We have $\tilde{q} \succ q$ and $\tilde{q}_1 = q_1$, i.e. the average selling price $p_{av} = \tilde{p}_{av}$ remains the same, and one can check that $\tilde{p}_k > p_k$ for $k \le 5$ and $\tilde{p}_6 < p_6$ (the digital divide is d = 5). But even more is true: the expected price \tilde{p}_3 paid by the new type 3 households under distribution \tilde{q} is not only higher than p_3 , but even higher than p_2 , the expected price paid by the same households before the change. Thus the increased search by these former type 2 households benefits only the highest type 6 in the population, and actually makes those who search more worse off than before.

Note that the concept of "more search" introduced in Def. 1 is stronger than First Order Stochastic Dominance. Regarding FOSD, we have:

Corollary 3.5. Assume $q_1 = \tilde{q}_1$, $\tilde{q} \neq q$, and \tilde{q} FOSD q. Then there exist thresholds $1 < \underline{d} \leq \overline{d}$, such that all types below \underline{d} (resp. above \overline{d}) pay higher (resp. lower) expected prices under \tilde{q} than under q.

Proof: see Appendix.

In particular, the uninformed (type k = 1) always suffer from an FOSD improvement among the shoppers. We conjecture that Cor. 3.5 holds with $\underline{d} = \overline{d}$, i.e. the digital divide result in Th. 3.2

¹⁰Similar observations are known as Simpson's paradox in statistics.

remains true if "more search" is defined by FOSD, but have not been able to prove this stronger result.

Finally, if we assume only \tilde{q} FOSD q (but not necessarily $\tilde{q}_1 = q_1$), it is still true that sufficiently high types pay lower prices under \tilde{q} , because the additional change from q_1 to a (necessarily smaller) \tilde{q}_1 lowers all prices, by Th. 3.1. The effect on the uninformed of such a general FOSD change is ambiguous: the reduction in q_1 tends to lower p_1 , but the "digital divide" tends to raise p_1 .

4 Conclusion

It has become widely accepted in economics that consumer welfare can be increased by making the comparison of prices easier. If previously uninformed consumers start to compare prices this is definitely the case. However, if already informed consumers compare more prices total consumer welfare stays unaffected. Moreover, it will result in a redistribution of consumer welfare from relatively uniformed to informed consumers.

The internet – and with it the onset of price comparison sites – has been largely praised by economists and policy makers as a way to ease consumer search and thereby increase competition. Consequently, policies to improve internet accessibility, and to promote the use of price comparison sites, are also seen as tools to increase consumer welfare. Whether such policies are indeed beneficial to all consumers is however questionable. Suppose there are three types of consumers: those who are uninformed consumers, those who sample just a few prices, e.g. by using newspapers ads, and those who rely on internet price comparison sites. If the uninformed consumers are not affected by measures promoting search, total consumer welfare will remain constant. Moreover, if partially informed consumers decide to become fully informed, then the uninformed will expect to pay a higher price.¹¹ Consequently, when trying to enhance competition by easing the burden of comparing prices uninformed consumers should be targeted first.

Our findings are at the heart of the current political debate on broadband stimulus which turns

¹¹As seen in Example 3.4, it is not even certain that the partially informed consumers will benefit from such a shift.

around the question whether existing internet connections should be made faster or the internet should be made accessible to a broader audience. *The Economist* argues that:

... But the case for large-scale government investment in broadband is not as strong as its proponents claim. When it comes to promoting economic activity, it is easy to see why having broadband is better than not having it, but most benefits are likely to come from wiring people up in the first place rather than making existing connections hum faster. ... ("Broadband Stimulus," *The Economist*, January 31st 2009, Vol. 380, Number 8616, p12.)

Our model supports this viewpoint, as making existing connections faster would merely result in a redistribution of welfare from poorly informed consumers to better informed consumers, whereas wiring up new consumers would increase total consumer welfare.

5. Appendix

Sketch of Proof of Prop. 2.1(c).

First one establishes that the equilibrium price distribution F must be continuous and strictly increasing on a support of the form $[p_{min}, p_{max}]$ with $0 < p_{min} < p_{max} = 1$ (otherwise, one can easily find profitable deviations). Given this, we observe that every price in the support must give the same (expected) payoff to a firm. If the firm charges $p_{max} = 1$ it gets only its share of uninformed consumers, so $\pi = p_{max} \cdot \mu q_1 = \mu q_1$ is the equilibrium profit. If it charges p_{min} it gets all households to which it is linked, so $\pi = p_{min} \cdot \mu \kappa$, hence $p_{min} = q_1/\kappa$. Total profits are $N\pi$ and must be equal to total household expenditure $M p_{av}$, so $p_{av} = (N/M)\mu q_1 = q_1$.

A firm charging any price $p \in [p_{min}, 1]$ is observed by $\mu q_k k$ households of type k and makes a sale to such a household iff the other (k-1) prices observed by the household are higher than p, which occurs with probability $[1 - F(p)]^{k-1}$. Thus the firm's profit is

$$p.\sum_{k=1}^{N} \mu q_k k [1 - F(p)]^{k-1} = p.\mu q_1. \Psi[1 - F(p)] = \pi = \mu q_1$$

This implies

$$p.\psi[1 - F(p)] \equiv 1 \tag{2}$$

or
$$F(p) = 1 - \psi^{-1}(1/p)$$

Proof of Lemma 2.2.

The function $x = F_k(p) = 1 - [1 - F(p)]^k$ is one-to-one from $[p_{min}, 1]$ to [0, 1]. Solving for p gives the inverse function $p = F_k^{-1}(x)$ as follows: $(1 - x)^{1/k} = 1 - F(p) = \psi^{-1}(1/p)$ (by Prop. 2.1), hence $1/p = \psi[(1 - x)^{1/k}]$ or

$$p = F_k^{-1}(x) = \frac{1}{\psi[(1-x)^{1/k}]}$$

It is geometrically obvious that the areas in the unit square to the left and to the right of the graph of F_k (which is the same as the graph of F_k^{-1}) sum to one, i.e. $\int_0^1 F_k^{-1}(x) dx + \int_{p_{min}}^1 F_k(p) dp = 1$. By a well-known formula for RS-integrals we have

$$\int_{p_{min}}^{1} p \, dF_k(p) + \int_{p_{min}}^{1} F_k(p) \, dp = p \cdot F_k(p) |_{p_{min}}^{1}$$

The leftmost term in this expression is p_k , and the RHS is equal to 1, hence

$$p_k = 1 - \int_{p_{min}}^{1} F_k(p) \, dp = \int_0^1 F_k^{-1}(x) \, dx = \int_0^1 \frac{1}{\psi[(1-x)^{1/k}]} \, dx$$

The change of variable $x = f(u) := 1 - u^k$, with $f'(u) = -ku^{k-1}$, transforms the last integral to

$$p_{k} = \int_{0}^{1} \frac{ku^{k-1}}{\psi(u)} du = \int_{0}^{1} \frac{1}{\psi(x)} dx^{k}$$

The proof of Theorem 3.2 is preceded by some lemmas.

Lemma 5.1. Assume $\tilde{q}_1 = q_1$ and $\tilde{q} \succ_{\ell} q$, $2 \le \ell < N$. Then there is a number $b \in (0,1)$ such that $\tilde{\psi}(x) < \psi(x)$ for $x \in (0,b)$ and $\tilde{\psi}(x) > \psi(x)$ for $x \in (b,1]$.

Of course, by continuity, $\tilde{\psi}(b) = \psi(b)$, i.e. $\tilde{\psi}$ and ψ cross only once in (0,1], at the point *b*.

Proof of Lemma 5.1: Since $\tilde{q}_1 = q_1 > 0$, it suffices to prove the assertion for $\varphi = q_1 \psi$. By assumption, $q_k - \tilde{q}_k \ge 0$ for $k < \ell$, $q_\ell - \tilde{q}_\ell > 0$, and $q_k - \tilde{q}_k \le 0$ for $k > \ell$.

We know that $\varphi(0) - \tilde{\varphi}(0) = q_1 - \tilde{q}_1 = 0$ and $\varphi(1) - \tilde{\varphi}(1) = \kappa - \tilde{\kappa} < 0$. For $0 < x \le 1$ we have

$$\varphi(x) - \tilde{\varphi}(x) = \sum_{k=1}^{\ell} k(q_k - \tilde{q}_k) x^{k-1} + \sum_{k=\ell+1}^{N} k(q_k - \tilde{q}_k) x^{k-1} = \frac{1}{x^{\ell-1}} [A(x) - B(x)]$$

where

$$A(x) = \sum_{k=1}^{\ell-1} k(q_k - \tilde{q}_k) \frac{1}{x^{l-k}} + \ell(q_\ell - \tilde{q}_\ell)$$

is nonincreasing in x and $\geq \ell(q_k - \tilde{q}_k) > 0 \ \forall x$, and

$$B(x) = \sum_{k=\ell+1}^{N} k(\tilde{q}_k - q_k) x^{k-\ell}$$

is strictly increasing in x (because at least one coefficient $k(\tilde{q}_k - q_k)$ must be positive), and tends to zero for $x \to 0$. Therefore the function f(x) := A(x) - B(x) is strictly decreasing on (0, 1], positive for x near zero, and negative for x = 1 (because $\varphi(1) - \tilde{\varphi}(1) < 0$). This implies the assertion.

Lemma 5.2. Let g(x) be a nonnegative continuous function on the interval [0,1] which is strictly positive except in at most finitely many points, and choose b with 0 < b < 1. For i = 1, 2, ... define

$$A_i := \int_0^b g(x) dx^i = \int_0^b g(x) ix^{i-1} dx, \quad B_i := \int_b^1 g(x) dx^i = \int_b^1 g(x) ix^{i-1} dx.$$

Then

$$\frac{B_i}{A_i} < \frac{B_{i+1}}{A_{i+1}}$$
 for $i = 1, 2, ...$

(obviously A_i , B_i are always positive).

Intuitively, the distribution $H_i(x) = x^i$ on the interval [0,1] has more weight on the right if *i* increases, hence B_i should increase relative to A_i (one might also think that $B_i - A_i$ increases in *i*, but one can show by examples that this is not true in general).

Proof of Lemma 5.2: Fix $b \in (0,1)$ and $i \ge 1$. For any $n \ge 2$ the integrals A_i and B_i can be written as Riemann sums

$$A_{i} = \sum_{j=1}^{n} g(\eta_{j}) i \eta_{j}^{i-1} dy_{j}, \quad B_{i} = \sum_{j=1}^{n} g(\zeta_{j}) i \zeta_{j}^{i-1} dz_{j}$$

where

$$0 = y_0 < y_1 < \dots < y_n = b, \quad b = z_0 < z_1 < \dots < z_n = 1$$

$$\eta_j \in (y_{j-1}, y_j), \quad \zeta_j \in (z_{j-1}, z_j), \quad dy_j = y_j - y_{j-1}, \quad dz_j = z_j - z_{j-1} \quad \forall j$$

Moreover, the numbers y_j , η_j , z_j , ζ_j can be chosen so that the summands

$$\alpha_{ij} := g(\eta_j)i\eta_j^{i-1}dy_j = \int_{y_{j-1}}^{y_j} g(x)ix^{i-1}dx$$

are all equal, i.e. $\alpha_{ij} = A_i/n$ for all j = 1, ..., n; and the summands

$$\beta_{ij} := g(\zeta_j) i \zeta_j^{i-1} dz_j = \int_{z_{j-1}}^{z_j} g(x) i x^{i-1} dx$$

are also all equal, i.e. $\beta_{ij} = B_i/n$ for all j = 1, ... n.

Now, using these *same* numbers y_j , η_j , z_j , ζ_j , we can *approximate* A_{i+1} , B_{i+1} by Riemann sums of the form

$$\tilde{A}_{i+1} = \sum_{j=1}^{n} \alpha_{i+1,j}, \quad \tilde{B}_{i+1} = \sum_{j=1}^{n} \beta_{i+1,j}$$

where

$$\alpha_{i+1,j} := g(\eta_j)(i+1)\eta_j^i dy_j, \quad \beta_{i+1,j} := g(\zeta_j)(i+1)\zeta_j^i dz_j.$$

By choosing *n* is sufficiently large, we can make \tilde{A}_{i+1} resp. \tilde{B}_{i+1} arbitrarily close to A_{i+1} resp. B_{i+1} (because *g* is continuous and all possible Riemann sums converge to the same integral if the grid size goes to zero). Then

$$\frac{\beta_{i+1,j}}{\alpha_{i+1,j}} = \frac{g(\zeta_j)(i+1)\zeta_j^i dz_j}{g(\eta_j)(i+1)\eta_j^i dy_j} = \frac{\beta_{ij}}{\alpha_{ij}} \cdot \frac{\zeta_j}{\eta_j} = \frac{B_i}{A_i} \cdot \frac{\zeta_j}{\eta_j}$$
(3)

By construction, $\zeta_j/\eta_j > z_{j-1}/y_j \ge \min_{1\le i\le n} \{z_{i-1}/y_i\} =: \lambda > 1$ (recall $n \ge 2$). Moreover, if we proceed to a finer partitioning, replacing *n* by n' = n + 1 (with the same properties as above, i.e. all summands α'_{ij} resp. β'_{ij} are equal), then $y'_j < y_j$ and $z'_j > z_{j-1}$ for j = 1, ..., n. Therefore a fortiori $\zeta'_j/\eta'_j > z'_{j-1}/y_j \ge \lambda > 1$.

Equation (3) implies $\beta_{i+1,j} > \lambda (B_i/A_i) \alpha_{i+1,j}$, hence

$$\tilde{B}_{i+1} = \sum_{j} \beta_{i+1,j} > \lambda \frac{B_i}{A_i} \sum_{j} \alpha_{i+1,j} = \lambda \frac{B_i}{A_i} \tilde{A}_{i+1}$$

Therefore $\tilde{B}_{i+1}/\tilde{A}_{i+1} > \lambda B_i/A_i$ and, since $\tilde{B}_{i+1}/\tilde{A}_{i+1} \to B_{i+1}/A_{i+1}$ as $n \to \infty$, we obtain $B_{i+1}/A_{i+1} \ge \lambda B_i/A_i > B_i/A_i$.

Proof of Theorem 3.2.

Let $\tilde{q}_1 = q_1$, $\tilde{q} \succ_{\ell} q$, $2 \le \ell < N$, and denote by p_k resp. \tilde{p}_k the expected price paid by a type k household under the search distribution q resp. \tilde{q} , for k = 1, 2, ...N.

Consider the function

$$h(x) := \frac{1}{\psi(x)} - \frac{1}{\tilde{\psi}(x)}.$$

By Lemma 5.1, there is a point *b*, 0 < b < 1, such that h(x) < 0 for 0 < x < b and h(x) > 0 for b < x < 1. Therefore, by equation (1):

$$p_k - \tilde{p}_k = \int_0^1 h(x) dx^k = B_k - A_k$$

where $A_k = \int_0^b |h(x)| dx^k$, $B_k = \int_b^1 |h(x)| dx^k$. The function g(x) := |h(x)| satisfies the assumptions of Lemma 5.2, hence B_k/A_k increases strictly with k. Thus, if $p_m - \tilde{p}_m = B_m - A_m > 0$ for some m, so that $B_m/A_m > 1$, we have also $B_k/A_k > 1$, hence $p_k - \tilde{p}_k = B_k - A_k > 0$ for all $k \ge m$. In other words, if the expected price decreases for some type m, then also for all higher types.

Denote by *d* the lowest type for whom the expected price does *not* increase, i.e. *d* is the first index such that $p_d \ge \tilde{p}_d$.

We know that $\tilde{p}_{min} < p_{min}$, therefore $\tilde{p}_N < p_N$ for N sufficiently large (because a household who searches long enough must find a price arbitrarily close to the minimum); thus $d < \infty$.

It remains to show that d > 1, i.e. not all prices go down. Using vector notation, write $q \cdot p = \sum_k q_k p_k$. Then the expected average prices satisfy

$$\tilde{q}_1 = \tilde{p}_{av} = \tilde{p} \cdot \tilde{q} = q_1 = p_{av} = p \cdot q \tag{4}$$

Moreover, the components of $p = (p_1, \dots, p_N)$ are strictly decreasing and \tilde{q} FOSD q. Therefore $\tilde{q} \cdot p < q \cdot p$. If all prices \tilde{p}_k were less than p_k , we should have $\tilde{p} \leq p$, hence $\tilde{q} \cdot \tilde{p} \leq \tilde{q} \cdot p < q \cdot p$, contradicting (4).

Proof of Cor 3.5.

A search distribution $q = (q_1, \dots, q_N)$ can equivalently be represented by its cumulative distribution $Q = (Q_1, \dots, Q_N)$, where $Q_k = q_1 + \dots q_k$ for $k = 1, \dots, N$ (of course $Q_1 \le Q_2 \le \dots Q_N = 1$). Write $\tilde{Q} \succ Q$ iff $\tilde{q} \succ q$. Now assume $\tilde{q}_1 = q_1$, $\tilde{q} \neq q$, and \tilde{q} FOSD q. i.e. $\tilde{Q}_k \leq Q_k \forall k$. Write $Q = (Q_1, \ldots, Q_N)$, $\tilde{Q} = (\tilde{Q}_1, \ldots, \tilde{Q}_N)$ and define $Q^1 = Q$, and $Q^{\ell} = (\tilde{Q}_1, \ldots, \tilde{Q}_{\ell-1}, Q_\ell, \ldots, Q_N)$ for $2 \leq \ell \leq N$. Then $Q^N = \tilde{Q}^N$ and each Q^ℓ is obtained from $Q^{\ell-1}$ by shifting the nonnegative mass $\delta_\ell = Q_\ell - \tilde{Q}_\ell$ of consumers from type $\ell - 1$ to type ℓ , i.e. $Q^\ell \succ Q^{\ell-1}$ (or $Q^\ell = Q^{\ell-1}$) for $\ell = 2, \ldots N$. Each move from $Q^{\ell-1}$ to Q^ℓ either leaves all prices unchanged (if $Q^\ell = Q^{\ell-1}$), or it creates a digital divide $d_\ell > 1$, raising expected prices for the types below d_ℓ and lowering prices for the types above d_ℓ , by Th. 3.2. Since $\tilde{q} \neq q$, such a change occurs at least once. Putting $\underline{d} = \min(d_\ell)$, $\overline{d} = \max(d_\ell)$ gives the assertion of the Corollary.

References

- ARMSTRONG, M., J. VICKERS, AND J. ZHOU (2009): "Consumer Protection and the Incentive to Become Informed," *Journal of the European Economic Association*, 7(2-3), 399–410.
- BAYE, M. R., J. MORGAN, AND P. SCHOLTEN (2003): "The Value of Information in an Online Consumer Electronics Market," *Journal of Public Policy and Marketing*, 22(1), 17–25.
- (2006): "Information, Search, and Price Dispersion," in *Handbook on Economics and Information Systems*, ed. by T. Hendershott, pp. 323–376. Elsevier.
- BLUME, L. E., D. EASLEY, J. KLEINBERG, AND V. TARDOS (2009): "Trading networks with price-setting agents," *Games and Economic Behavior*, 67(1), 36–50.
- BURDETT, K., AND K. L. JUDD (1983): "Equilibrium Price Dispersion," *Econometrica*, 51(4), 955–69.
- COROMINAS-BOSCH, M. (2004): "Bargaining in a network of buyers and sellers," *Journal of Economic Theory*, 115(1), 35–77.
- GALEOTTI, A. (2009): "Talking, Searching and Pricing," *International Economic Review*, forthcoming.

- GALEOTTI, A., S. GOYAL, M. O. JACKSON, F. VEGA-REDONDO, AND L. YARIV (2009): "Network Games," *Review of Economic Studies*, forthcoming.
- JACKSON, M. O., AND L. YARIV (2007): "Diffusion of Behavior and Equilibrium Properties in Network Games," *American Economic Review (Papers and Proceedings)*, 97(2), 92–98.
- JANSSEN, M. C. W., AND J. L. MORAGA-GONZÁLEZ (2004): "Strategic Pricing, Consumer Search and the Number of Firms," *Review of Economic Studies*, 71(4), 1089–1118.
- KRANTON, R. E., AND D. F. MINEHART (2001): "A Theory of Buyer-Seller Networks," *American Economic Review*, 91(3), 485–508.
- LACH, S., AND J. L. MORAGA-GONZÁLEZ (2009): "Asymmetric price effects of competition," IESE Research Papers D/797, IESE Business School.
- LEVER, C. (2008): "Price Competition on a Buyer-Seller Network," Mimeo, Department of Economics, Stanford University.
- MORGAN, J., H. ORZEN, AND M. SEFTON (2006): "An experimental study of price dispersion," *Games and Economic Behavior*, 54(1), 134–158.
- ROSENTHAL, R. W. (1980): "A Model in Which an Increase in the Number of Sellers Leads to a Higher Price," *Econometrica*, 48(6), 1575–79.
- STAHL, D. O. (1989): "Oligopolistic Pricing with Sequential Consumer Search," American Economic Review, 79(4), 700–712.
- VARIAN, H. R. (1980): "A Model of Sales," American Economic Review, 70(4), 651-59.
- WALDECK, R. (2008): "Search and price competition," *Journal of Economic Behavior & Or*ganization, 66(2), 347–357.
- WILDE, L. L., AND A. SCHWARTZ (1979): "Equilibrium Comparison Shopping," *Review of Economic Studies*, 46(3), 543–53.