12 Price search and obfuscation: an overview of the theory and empirics

Sara Fisher Ellison

1 INTRODUCTION

The cost of price search is an important and salient feature of how many retail markets function. Take the market for new cars as a concrete example. Prices are typically not posted at dealerships. When they are, it is understood by most market participants that they are negotiable, so the posted price really is just a maximum price. Calling a dealership for a price quote is often futile: customers are told to ‘come in, test drive the car, and we can talk price at that point’. There are many ways to upgrade and customize each car, making price comparisons across dealerships difficult. All-in prices may also include additional fees, taxes, destination charges, and so forth, which may neither be reflected in the quote nor be particularly transparent. Also, of course, a visit to a dealership to obtain a price quote can be a very time-consuming activity. It might involve a fair amount of sitting around, perhaps totaling a few hours. Dealerships are typically spread out geographically, so obtaining a number of reliable price quotes could easily be a two- or three-weekend proposition. Clearly, the high cost of price search in that market can have a large effect on consumer behavior (how many customers end up visiting one dealership and buying from them?) and can have a large effect on dealers’ equilibrium prices as well. Furthermore, it can certainly seem as though dealers go out of their way to make price search a very difficult and time-consuming process. In 1961, George Stigler published his seminal article ‘The economics of information’, which started the formal conversation in economics about costly price search. Industrial organization, in an attempt to understand retail pricing behavior, has seen a recent renaissance of some of the ideas first mentioned in Stigler (1961). Stigler does not provide a game-theoretic model of price search, where both sellers and consumers behave rationally. He does, though, make a number of important observations which have helped frame the conversation on price search. He ties together price dispersion and imperfect information about prices by consumers, quite closely as it turns out: ‘Price dispersion is a manifestation – and, indeed it is the measure – of ignorance in the market’ (Stigler 1961, p. 214). That statement augurs an important branch of the search literature highlighting equilibrium price dispersion as a result of search costs. He points out that price search typically has a decreasing marginal value in the number of quotes already obtained. He observes that search may be more valuable in markets where consumers make repeated purchases (provided there is positive correlation in prices over time). He mentions situations where one-time buyers would be expected to pay higher average prices than repeat purchasers. Much modern empirical work leverages these insights. Stigler also discusses market mechanisms that arise to mitigate market failures or inefficiencies caused by costly price search, such as specialized brokers and centralized market platforms.
My focus here is on price search, both the theory and empirics of the broad topic first motivated by Stigler (1961). I also discuss a growing literature on obfuscation, the actions that firms take to make price search more costly. Finally, I touch on a more recent topic that has grown out of the study of price search, in particular how consumer price search has changed with advances in technology. One broad goal is to lay out a useful theoretical framework for assessing markets in which price search costs are an important feature. A second is to discuss the empirical evidence on price search costs in retail markets. Along the way, I also indicate the different styles of empirical work and discuss their roles in improved understanding of these phenomena.

Section 2 discusses four main types of search models, starting with the basic result known as the Diamond paradox. I then discuss heterogeneous search, fixed-sample search and clearinghouse models, and, finally, models of obfuscation. Section 3 discusses reduced-form empirical studies, which can be thought of as tests of comparative statics of specific search models, and structural empirical studies, which assume specific features of consumer and firm behavior and then exploit those assumptions to back out primitives, such as search costs. Section 4 discusses empirical evidence on price obfuscation. Section 5 deals with empirical evidence on how technologies such as the Internet and smart phones are changing price search from the way it was conducted in more traditional brick-and-mortar markets. Section 6 concludes.

It is worth noting that search models have been applied in a variety of contexts in economics, not just to retail price search. Perhaps the most prominent example is job search. In particular, Stigler himself moved on fairly quickly to applications of search theory to the labor market (Stigler 1962), where it has received significant popular and academic attention over the past five decades, resulting, even, in a Nobel Prize awarded to Diamond, Mortensen, and Pissarides in 2010. (Stigler also won a Nobel Prize, in 1982.) The role that job search and match quality play in individuals’ likelihood of finding a job, as well as the role they play in how macro-level measures such as unemployment evolve, has been extensively studied. Some of these search models developed and studied by labor economists are applicable to price search in retail markets as well as to job search in labor markets, but others exploit the particular structure or institutional detail of the labor market. Our focus will be on the models most often found useful to the analysis of price search.

Our focus is even sharper, though, specifically on retail price search. Search models and ideas have been applied to price search in other contexts: wholesale markets, or markets for services, for instance. Because the settings are so similar, results in those contexts are often relevant to the questions we study here, and can inform our knowledge of retail price search in useful ways. I refer to such results occasionally as they can enhance our understanding of retail price search.

2 THEORY

2.1 The Most Basic Model of Price Search

We start with the model that serves as the basis for our theoretical discussions of price search, Diamond (1971). First, though, it is important to place the topic of price search
in the broader context of industrial organization (IO). Among other things, IO seeks to provide answers to the question posed by the Bertrand paradox: what are the factors that allow firms to escape marginal cost pricing? When teaching IO, we offer our students a number of possibilities. We present the Cournot model, which offers one, fairly unsatisfactory, answer. We mention product differentiation. We teach students that firms can collude, tacitly or explicitly. Finally, we offer the presence of search costs as one potential answer. Our baseline model, due to Diamond (1971), makes that point quite emphatically. It highlights the important role that even small search frictions can play in the functioning of markets, and, indeed, firms’ ability to escape Bertrand competition.

We start with \( N \) firms producing a homogeneous good. They share a common and constant marginal cost \( c \). There is a continuum of consumers, each having demand \( D(p) \), where \( p \) is the price of the good. Denote the monopoly price \( p^m \). Assume that \((p - c)D(p)\) is concave.

Suppose firms simultaneously choose prices \( p_1, \ldots, p_N \). Consumers then search optimally, getting some number of price quotes before stopping and purchasing \( D(p) \) from the firm with the lowest price among those they have sampled. Note that the assumptions mimic those for price competition a la Bertrand, but for the addition of a cost of price search.

**Proposition:** Suppose each consumer has a cost \( s \) of obtaining a price quote with \( 0 < s < CS(p^m) \), where \( CS(.) \) is the consumer surplus which occurs at that particular price. Then the unique Nash equilibrium is \( p_1^* = p_2^* = \ldots = p_N^* = p^m \).

**Proof:** To see that this is a Nash equilibrium, note that if \( p_1^* = p_2^* = \ldots = p_N^* = p^m \), then optimally searching consumers will just get one price quote. Hence, if a firm moves price away from \( p^m \), it will get no extra customers, implying that \( p^m \) is optimal.

To check whether there are other equilibria, suppose \((p_1^*, \ldots, p_N^*)\) is a pure-strategy Nash equilibrium. Assume, without loss of generality, that \( p_1^* \) is the lowest price. If \( p_1^* \) is not equal to \( p^m \), then firm 1 can deviate and change its price to

\[
\min \left\{ p_1^* + \frac{s}{2D(p_1^*)}, p^m \right\}.
\]

Any consumer who samples firm 1 will buy, as the possible gain from another price quote is less than or equal to \( \frac{3}{2} \). To see that the gain is at most \( \frac{3}{2} \), note that buying \( D(p_1^*) \) units at the higher price is only \( \frac{3}{2} \) more costly than the best possible outcome of another search: getting \( D(p_1^*) \) units at \( p_1^* \). Hence, firm 1 sells to the same number of consumers and makes more on each.

If \( p_1^* = p^m \) and some other firm charges more, that firm benefits from cutting its price to \( p^m \).

There also cannot be a mixed equilibrium where firm \( i \) mixes on \([p_i, \bar{p}]\). Again, the firm with the lowest \( p_i \) benefits from charging

\[
\min \left\{ p_i + \frac{s}{2D(p_i)}, p^m \right\}
\]

whenever it was going to charge in a neighborhood of \( p_i \).
The result is striking: the tiniest of search frictions can cause a perfectly competitive market to flip to one in which all firms charge the monopoly price. This result makes the point that search frictions are important in determining how markets function, but perhaps too emphatically. There are at least three aspects of the result that do not seem consistent with empirical observation. First, the knife-edge quality, or lack of continuity, in how search frictions alter the market seems implausible. Second, the model provides an alternative ‘Law of one price’, this time at the monopoly price instead of marginal cost, but still predicts no price dispersion in equilibrium. Price dispersion, even for quite homogeneous goods, is ubiquitous and perhaps larger in markets where price search is costly. Third, the lack of consumer search in equilibrium is inconsistent with casual observation.

This casual empiricism suggests that the Diamond model would be difficult to use as a basis for empirical work. Another set of observations suggests a direction in which to enrich it. We all know consumers who engage in price search, some quite extensively, and who actually enjoy it, instead of viewing it as a cost or burden. At the same time, it is also true that many of consumers engage in very little price search, perhaps viewing it as costly and hoping to free-ride off of the more favorable equilibrium outcomes that result from the fraction of consumers who do engage in price search. These observations point toward an alternative model which accommodates consumer heterogeneity.

With these observations in mind, we move to discussion of models that incorporate these alternative assumptions and whose results may better reflect our intuition.

2.2 Heterogeneous Search

The Diamond model provides a useful base case, but its empirical applicability is limited. Both the prediction of one price (the monopoly price) and the knife-edge quality to the result that the equilibrium flips with the introduction of an epsilon cost of price search put this model at odds with casual (and formal) empiricism. An important insight, that there could be heterogeneity in the costs that consumers bear for price search, seems consistent with common observations. In fact, there could be consumers who bear no cost of price search, or even a negative cost. The most influential papers to formalize this idea are Stahl (1989, 1996). Related are Reinganum (1979) and MacMinn (1980), which derive qualitatively similar results about equilibrium prices by using firm heterogeneity, instead of consumer heterogeneity.

I now discuss the Stahl (1989) model in more detail. We start with the same assumptions as the Diamond model above. Suppose, in addition, that a fraction $\mu$ of the consumers have negative search costs, ‘shoppers,’ and a fraction $1 - \mu$ have search costs distributed on the same interval $[s, \overline{s}]$ with $0 < s < \overline{s} < \text{CS}(p^*)$, ‘non-shoppers’.6

Proposition: (a) The model described above has no pure-strategy Nash equilibrium. (b) There exists a symmetric mixed-strategy Nash equilibrium in which firms choose prices from an atomless distribution $F$ whose support does not contain $c$.

This model has a number of interesting features that could be relevant for empirical testing and application. First, there is a smooth relationship between the fraction of consumers who are ‘shoppers’ and the degree of equilibrium price dispersion. As the percentage of ‘shoppers’ increases, price converges to marginal cost, the range of prices shrinks, and the Bertrand paradox obtains. As the percentage of ‘shoppers’ decreases, price converges to the monopoly price, the range of prices again shrinks, and the Diamond paradox obtains. In other words, as the fraction of shoppers varies from zero to one, the amount of price dispersion varies continuously, from zero to positive dispersion back to zero. Expected prices vary continuously from marginal cost up to monopoly price.

Second, there is a related comparative static involving the level of search costs of the ‘non-shoppers’. As search costs decrease, price converges to marginal cost and the range of prices shrinks. As search costs increase, price dispersion remains because the ‘shoppers’ still provide incentive for firms to price low with positive probability.

Finally, it is worth emphasizing the model’s prediction that firms employ mixed strategies in pricing, resulting in equilibrium price dispersion for identical products in intermediate cases. In addition, in a common simplification of the model often considered, there are only two types of consumers, ones with a common negative search cost and ones with a common positive search cost, instead of a continuum of ‘non-shoppers.’ In that simplification, only the ‘shoppers,’ who have a negative search cost, actually search in equilibrium.

Many of the main issues we had with the empirical salience of Diamond’s results are addressed with this model. The knife-edge quality of the equilibrium is gone, replaced by smooth transitions as search costs change or the fraction of ‘shoppers’ changes. We can obtain both Bertrand and Diamond results as limiting cases of the Stahl model, with a smooth bridge between them. Also, we obtain equilibrium price dispersion, consistent with decades of empirical evidence. Finally, we have search being conducted in equilibrium and, in fact, different levels of search performed by different types of consumers.

As a result of these attributes, Stahl’s model has proven attractive to both theorists and empiricists since its publication.

2.3 Sequential versus Fixed Sample Search and Clearinghouse Models

So far, we have seen models where consumers are faced with a per-price-quote cost of price search and they decide after each price quote obtained whether to continue searching. Sometimes the search technology is at odds with that type of model. Consider, for instance, buying Consumer Reports and seeing prices, product characteristics, and product reviews all in one place. In general, we might think that there is a two-part cost associated with price search, a fixed cost paid once and an additional per-price-quote cost. The models we have seen so far are at one extreme: the fixed cost is zero, and per-quote cost is non-zero. At the other extreme, we could imagine that consumers pay one up-front cost and see all (or a predetermined number of) prices simultaneously. In other words, the per-quote price is zero.

Certainly price search on the Internet often has more of a fixed-cost-with-no-per-quote-cost flavor to it. Consumers can often find clearinghouses or price search websites listing most or all of the relevant prices in one place. Sometimes consumers can even simply click on a live link to make their purchase after they have performed their price
search and chosen a product to buy. Even in such an environment, though, some consumers may be unaware of those clearinghouses or simply feel more comfortable visiting a seller’s website directly.

A number of papers have captured this clearinghouse idea (for instance, Salop and Stiglitz 1977; Rosenthal 1980), but Baye and Morgan (2001) provide a particularly good example. In contrast to much of the literature, they endogenize not only the decisions of the firms and consumers to use the clearinghouse, but also the fees charged by the clearinghouse. They show that an equilibrium with price dispersion exists even with \textit{ex ante} homogenous firms and consumers. That is, they obtain results qualitatively similar to those from heterogeneous search models without using heterogeneity as a starting point.

Baye and Morgan (2001) start with the assumption that each firm has a ‘home’ market, which it would serve as a monopolist if no consumers had access to the clearinghouse price information. There is also a monopolist clearinghouse, maximizing its own profits. It may charge fees to firms to list prices (and potentially attract consumers outside their ‘home’ markets) and may charge fees to consumers to access price information for all listing firms. Given assumptions on consumer behavior, they show that the clearinghouse sets fees to consumers low enough that all will access the price information, and sets fees to firms at a level where each firm lists its price, drawn from a distribution, with a positive probability. Therefore, in equilibrium, this model also generates price dispersion, suggesting its potential empirical relevance. Furthermore, this clearinghouse model and other similar ones provide a useful alternative to heterogeneous search models for empirical applications because the assumptions for a clearinghouse-type model may simply fit certain markets better than the assumptions for a heterogeneous search model.

Note that Baye et al. (2006) offer an excellent review of much of the relevant theory literature in the categories of search, heterogeneous search, and clearinghouse models. The focus of their chapter is on models that provide equilibrium price dispersion for homogenous goods, and many such models are ones involving costly price search. Interested readers may refer to that chapter for a detailed treatment of that literature and a more comprehensive list of related papers.

2.4 Models of Obfuscation

Comparative statics on the foregoing literature make one thing clear: firms benefit from higher price search costs and consumers benefit from lower price search costs. The cost of price search was treated as exogenous in the literature discussed above, but interest groups will have an incentive to raise or lower it, depending on their position. On the one hand, consumer groups may push for legislation to require public posting of prices or transparent description of prices. They might push for adoption of technology that would aid in price search. On the other hand, sellers could engage in a number of obfuscation techniques that would make price search more difficult. They could encourage manufacturers to proliferate model numbers, making comparison of essentially identical models across retailers difficult. They might make prices hard to find on a website or in a physical store. They might engage in strategies such as add-on pricing and upselling, where the efficient quality to be selling to most consumers is actually an upgrade of the advertised product and therefore potentially more difficult to search for. Though it may seem obvious that firms would have a collective interest in obfuscating to raise search
costs, it is not clear that all of these strategies would be individually rational for firms to engage in.

A recent branch of the theory literature on search endogenizes the search cost, recognizing the incentive that firms have to obfuscate, and also addresses the question of individual versus collective interest. Ellison (2005) was an early example, presenting a model of add-on pricing. The idea is that firms offered a base good, for which search and price comparison are easy, and an add-on or upgrade to the base good, for which search and price comparison are costly. Even when consumers anticipate the add-ons, firms are still able to raise prices, on average, with this strategy. A similar model is presented in Ellison and Ellison (2009), who note that ‘add-on’ could also be interpreted as ‘quality upgrade’, and the base good is the lowest quality good. The model has a number of predictions. First, the low-quality, easy-to-compare goods have low margins. Second, the higher-quality, hard-to-compare goods have higher margins. Third, firms use the easy-to-compare goods to attract customers. Fourth, the existence of differential quality in the market creates an adverse-selection problem: firms that price too low will attract more ‘cheapskates’ unwilling to upgrade to higher quality. Finally, in equilibrium, firms do not fully compete away higher margins on upgraded products by selling the easy-to-compare good too far below cost. It is interesting to note that improved technologies available to firms and consumers in online retailing, for instance, are typically assumed to aid in price search. To the extent that obfuscation strategies such as add-on pricing could be automated by these technologies, this could be a counterexample of that accepted wisdom.

Another approach to modeling obfuscation is taken by Ellison and Wolitzky (2012) and Wilson (2010). Both models start with homogeneous goods and heterogeneity in consumer time costs, as in Stahl, but firms are able to choose the level of search costs for the non-shoppers. Wilson constructs a three-stage game whereby duopolists first choose the cost that consumers will have to pay to find their prices. Consumers and firms all observe each firm’s search cost. Then firms choose mixtures over prices, and consumers (with unit demands) decide the extent and order of their search. Wilson shows that an equilibrium exists where one firm chooses a high search cost for its product and the other firm does not. Intuitively, obfuscation by one firm induces non-shoppers to search the rival first and makes them less willing to search the obfuscating firm second. This situation gives the rival incentive to raise its price, softening competition for the shoppers. This result indicates that creating positive search costs can be individually rational for firms and that those search costs can also be asymmetric.

Ellison and Wolitzky (2012) provide another model in which firms rationally engage in obfuscation. In their model, search costs are unobservable and they affect consumers by reducing the likelihood of future search. They find equilibria similar to Stahl’s but with higher prices. In fact, the upper bound of the price distribution is the monopoly price in all equilibria of the standard model. The conclusion is that obfuscation hurts consumers in two ways: consumers incur higher search costs and also pay higher prices.

Ireland (2007) provides a model that explores an alternative mechanism for obfuscation. In settings of imperfect price information and incomplete search, firms can benefit from listing multiple prices that consumers do not know are from the same firm. These firms enjoy increased monopoly power, and average prices increase. One could imagine such strategies becoming increasingly prevalent in Internet retail owing to the ease with which firms can proliferate brands online. (This observation provides an example of
another way in which improved technology might be hindering, not aiding, price search.)
And a final mechanism, limiting comparability, is proposed by Piccone and Spiegler (2012). In their model, firms can choose a ‘format’ in which to present their price, and, with some probability, consumers are unable to compare prices across different formats and simply choose to purchase from their default firm. They find that when consumers’ abilities to make comparisons are limited, firms can typically secure positive profits in a setting which is, otherwise, like Bertrand.

Based on the assumptions and the obfuscation mechanisms considered, these models could all have applicability in various empirical settings. One useful role of empirical work on search models is to derive testable predictions from these models and then perform tests of their relevance in settings where the particular assumptions or mechanisms seem to fit. This general style of empirical work is addressed in the next section.

3 EMPIRICAL EVIDENCE ON PRICE SEARCH

3.1 Reduced-Form Tests of the Theory

We started by recalling the theoretical framework for thinking about equilibrium effects of costly price search. We saw the stark result of the Diamond model that even very small search frictions can result in no equilibrium search, monopoly pricing, and no price dispersion. Although this result reinforces an important intuition – that even small search frictions can have large effects on how markets function – it does not offer predictions that are empirically salient. We also saw that search models where some fraction of consumers like to search, such as Stahl’s, generate more empirically plausible and useful results, in particular where average markups and the degree of price dispersion are related to search costs. These results arise from firms playing mixed-strategy equilibria.

With that framework in mind, I turn to discussion of various strands of empirical literature on price search. An early direction for empirical studies on price search were studies which established reduced-form associations between characteristics of price distributions and variables which measured intensity or ease or benefit of price search. One could also think of such papers as testing comparative statics of models like Stahl: as the percentage of searchers increases, price approaches cost and price dispersion shrinks.

An excellent example is Sorensen’s (2000) paper on prescription drugs. Exploiting a New York state law mandating the clear posting of prices for the top 150 prescription drugs at pharmacies, he gathered data and pharmacy characteristics from all of the pharmacies operating in two small, isolated towns in upstate New York. He also gathered drug-level data, such as therapeutic class to proxy for user demographics, and whether the drug is taken for chronic or acute conditions. From this latter information, Sorensen constructs a variable for purchase frequency to use as a proxy for the benefit of price search. If the drug is typically taken for a chronic condition, recurring purchases would have to be made, and the benefit of price search would be higher. The analysis is very straightforward: price range for a particular drug across pharmacies in a town is regressed on a variable measuring purchase frequency, controlling for factors such as wholesale costs and presence of generic drugs. Sorensen does, indeed, find that the observed price range of drugs purchased monthly is 28 percent smaller than the observed
price range for those purchased for one-time use, consistent with models of heterogeneous search. Recall, though, that the theory pertains to price levels, not just dispersion, suggesting an additional analysis using retail margins in place of price range. Using a published wholesale price list for drugs, Sorensen computes retail margins and does find that they are 37 percent lower for drugs purchased monthly versus one-time use.

Some settings, such as retail pharmaceuticals sold in physical pharmacies, suggest models of sequential search. Models of fixed sample search or clearinghouse models have gained more attention recently, though, as the Internet becomes a more important platform for retail commerce. In many cases, these models are a more accurate reflection of how consumers find prices online. Baye et al. (2004) is an early and important paper documenting price dispersion online and also providing reduced-form evidence in favor of a specific search model. They document significant price dispersion for over 1000 products; that dispersion is largely stable over time. Their other main empirical finding is that their measure of dispersion, the gap between the lowest and next lowest price, varies systematically with the number of firms listing prices, decreasing as \( N \) increases. They rationalize the patterns they observe with a clearinghouse-type model, where some fraction of consumers have access to all prices (through a clearinghouse like shopper.com) but some fraction do not. This type of a search model may be more relevant than a sequential search model for many online retail markets, especially, given the ubiquity of price search and comparison websites, ones where many price quotes can be obtained from a visit to one website.

De Los Santos et al. (2012) also provide reduced-form tests of sequential search with a known price distribution using data on web browsing and purchasing for books. They reject this model in favor of a model of fixed-sample search.

### 3.2 Structural Empirical Studies

Reduced-form studies are important for both testing existing models and informing new directions for the theory on search. More recently, though, the empirical search literature has pivoted toward an approach which relies more on theoretical models as opposed to providing tests of them: assuming rational search, typically a specific model, and using structural techniques, one can back out consumer search costs and other primitives of the system from firm pricing behavior, sometimes supplemented with quantity or market share data.

Villas-Boas (1995) straddles the line between a reduced-form test of search models and something more structural, relying more on specifics of a particular model. He bases his empirical strategy on a model by Varian (1980), which seeks to explain price dispersion by allowing for search costs to vary by proximity of the customer to the store as well as by ex ante heterogeneity in how informed customers are about price. The Varian model implies specific price distributions and Villas-Boas compares those with the samples of prices he observes for coffee and saltine crackers. Sorensen (2001), however, provides one of the earliest examples of true structural analysis of search models. Like his previous reduced-form empirical study, this study also uses data he gathered on retail prescription drug prices supplemented with transaction-level data from a marketing research firm covering the same two isolated towns in upstate New York. Using a sequential model of consumer search where consumers have a per-quote search cost drawn from a distribution, he
estimates that the mean search probability is about 9 percent, with essentially no price search for low price, one-time-use drugs and up to 30 percent price search for expensive drugs purchased repeatedly. Furthermore, the mean search cost is estimated to be about $15.

Hong and Shum (2006) provide a general treatment of identifying search costs from price distributions, an important piece linking theoretical models of search and empirical observation. They observe that, given a specific model of consumer search, the equilibrium conditions from that model can be exploited to infer consumer search costs from observed price distributions alone. That is, if you have an empirical price distribution of a homogenous good and are willing to make assumptions about how consumers engage in price search for that good, you can back out a distribution of consumer search costs. Depending on the model, it might even be possible to do so non-parametrically.

Hong and Shum (2006) gather data on prices of four different textbooks from numerous online vendors. They have no quantity data, but they do assume that positive sales are made at each price. (In other words, all of the prices they observe are ‘legitimate’ prices.) They consider two paradigms, fixed-sample search and sequential search. For fixed-sample search, they assume that the observed price distribution is the symmetric equilibrium mixed strategy of a game played between a continuum of firms and consumers. They can obtain a non-parametric estimate of the distribution of consumer search costs by exploiting the consumer’s optimality condition on searching and the firm’s optimality condition on pricing. The idea behind their method begins with the following observation: the expected savings that a consumer receives from obtaining an additional price quote is just the expected difference between the lowest out of \((i+1)\) price quotes and \(i\) price quotes. That sequence of marginal expected savings is non-increasing in \(i\) for any price distribution, while, by assumption, the cost per search is constant. Therefore, the sequence of marginal expected savings can also be interpreted as the search costs of the consumers indifferent between searching \((i+1)\) and \(i\) stores. Since we can obtain the former sequence from the empirical distribution, we can then back out the proportion of consumers with particular numbers of price quotes. Finally, the firm’s optimality condition allows us to obtain the distribution of search costs from these proportions.

In the case of sequential search, Hong and Shum posit a model such as Stahl (1989), with heterogeneous search, which would generate a non-degenerate equilibrium price distribution. They cannot estimate the distribution non-parametrically in this case and must assume a family of parametric distributions. For their empirical application, they use a gamma distribution for its flexible shape.

The estimates Hong and Shum (2006) obtain from the two paradigms vary considerably. The largest discrepancy for any of the titles they estimate is nearly tenfold: they estimate that the median per-quote search cost is $2.32 for a fixed sample search, versus $29.40 per quote in the case of sequential search. Since they are unable to test between the two paradigms, it may be sensible for the researcher to choose a paradigm based on his knowledge of how price search is carried out in a particular market.

Building on this work, Honka and Chintagunta (2015) develop a framework for testing between simultaneous and sequential search. Their method does have additional data requirements beyond those for Hong and Shum – they use data on prices supplemented with information on consumers’ consideration sets. They also provide a new estimation approach for the sequential search model. Honka and Chintagunta present an applica-
tion to purchasing auto insurance, somewhat afield of a traditional retail market, but their methods are broadly applicable.

Subsequent papers using similar methodology or extending it, and examining other applications include Moraga-González and Wildenbeest (2008) (memory chips), Baye et al. (2009) (personal digital assistants), Kim et al. (2010) (camcorders), Brynjolfsson et al. (2010) (books), De los Santos et al. (2012) (books), and Moraga-González et al. (2013) (memory chips). In addition, Koulayev (2013) carries out a similar analysis of hotel bookings. Strictly speaking, one might not categorize services such as hotel stays as part of retail commerce. They are, however, similar enough in terms of their economic characteristics that results on consumer search for these services help inform our broad knowledge of consumer search in retail markets.

Similarly, an interesting related paper is Hortaçsu and Syverson (2004). They study price dispersion in S&P 500 index funds, again not a retail product, narrowly defined. They find an 8 to 1 ratio between prices (fees to investors) of the ninetieth percentile fund to the tenth percentile fund in a market where the funds are all merely trying to match an index. Using both price and market-share data, they develop a methodology to separately identify the contribution of product differentiation and search costs to the observed price dispersion. Investors care about characteristics such as fund age and family, but after accounting for that differentiation, only fairly small search costs are needed to rationalize the observed prices and market shares.

A retail market with an extremely homogeneous product, which is nevertheless characterized by price search and price dispersion, is the market for gasoline. Noel discusses these and other issues in the context of gasoline markets in Chapter 16 of this volume.

Finally, another interesting example following on results in Hong and Shum is Spence (2014). In a study of a traditional retail market, textbooks, Spence uses a structural approach based on Hong and Shum (2006) but also exploits cross-sectional differences to identify the effects of search costs, as in Sorensen (2001). Sorensen exploits differences in the product type, drugs for chronic or acute conditions, to proxy for the consumer’s benefit from search. Spence exploits cross-sectional differences in searchers, not products, and he uses those differences to proxy for the cost of search, instead of the benefit. In particular, Spence studies the textbook searching and purchasing behavior of college students, using their year in school (freshman, sophomore, junior, or senior) as a measure of their experience, a proxy for their cost of search. His analysis is structural, and he is able to estimate characteristics of the search cost distributions as well as derive welfare estimates for market participants.

4 PRICE SEARCH AND OBFUSCATION

In addition to covering the classic models of price search, the earlier theory section discussed models that made a quite different point. If decreasing search costs can lead to Bertrand-like competition, firms, wanting desperately to escape that competition, have a strong incentive to take actions to increase search costs. These actions are typically called obfuscation. We have all experienced obfuscation in many forms – inefficiently low-quality goods being heavily advertised, fake sales, upselling strategies – so casual empiricism abounds.
A number of empirical papers over the years have documented obfuscation strategies in one way or another. I do not provide an exhaustive list, but I offer a few recent examples. Hackl et al. (2014) document firm pricing strategies, which are consistent with obfuscation and which exploits boundedly rational consumers. Li and Dinlersoz (2012) find that the design of shipping menus of Internet book retailers is consistent with an add-on-pricing-style obfuscation. McDonald and Wren (2014) study the online auto-insurance market. They find brand proliferation and patterns of pricing across brands within firms that are consistent with search obfuscation. Although auto insurance is not typically categorized as a retail product, those findings are likely to be relevant for traditional retail markets as well.

Formal empirical evidence – testing models of obfuscation or demonstrating evidence of the benefits of these strategies – has been scarce, but examples are appearing. An early paper of this type of more formal empirical evidence is Ellison and Ellison (2009). The empirical setting is an online market for computer components dominated by a single price search engine. The firms engage in a type of obfuscation where they offer multiple products in any particular product category. Searching for the lowest-priced product in that category is very easy, but searching for a good price on a product in that category that is of higher quality – has been pre-tested, or has better warranty terms, or is faster or more reliable – becomes very difficult. The search engine cannot create different categories for every dimension of differentiation among the products, and firms, knowing this, engage in cutthroat competition for the low-quality offerings in a particular category, use those products to attract customers, and then persuade the customers to upgrade when they visit the website with the intention to buying the low-priced product. (This strategy has elements of both the familiar loss-leader and bait-and-switch strategies.) The key is that price search is very easy for customers who want low quality but potentially tedious for customers who want higher quality precisely because the search engine cannot create a category called ‘128MB PC100 memory modules with decent warranty terms and reasonable quality’.

Using a year’s worth of data from a price search website (as well as detailed data provided by one firm), Ellison and Ellison (2009) are able to estimate demand for multiple quality levels of multiple products. Based on the demand estimates, they then derive matrices of estimated elasticities. Table 12.1 provides one such matrix, that of estimated demand elasticities for low-, medium-, and high-quality modules for 128MB PC100 memory modules. The diagonal of the matrix has own-price elasticities. The off-diagonals are cross-price elasticities, where, for instance, the number at the top of the middle column, −12.5, is the elasticity of demand for the medium quality with respect to the price of the low quality.

There are a number of important observations to take Table 12.1. First, demand is extremely elastic for the low-quality product, where search is easy. In fact, that estimated elasticity would predict a margin of 4 percent, not large enough for a firm to survive if it only sold products where search was easy. Second, demand is less elastic for the higher-quality products – the estimated elasticities suggest margins on the order of 12–15 percent. Third, there are large, negative, and significant cross-price elasticities on medium- and high-quality products with respect to low-quality price. This finding is striking given that the products are very close substitutes, and naïve economic models would suggest positive cross-price elasticities in that situation. These elasticities are
strong evidence in favor of firms engaging in an add-on pricing strategy. When the price of the low-quality module is low, customers are attracted, some of whom upgrade to medium or high quality.

The estimated elasticities tell us that firms are using the low-quality product to attract customers, some of whom upgrade to the higher-quality products. They also tell us that firms face an adverse-selection problem. We turn to the estimated parameters from the demand function to clarify this point.

Table 12.2 shows estimates from three separate regressions where the dependent variables are sales volumes in different quality levels for 128MB PC100 modules. The reported numbers are the estimated coefficients on the rank of the low-quality version on the price search website. First, note that the low-quality version has the largest (in absolute value) estimated effect, and the high-quality version has the smallest. We see evidence of the adverse selection that the firm faces: when a firm lowers the price of its low-quality version, it experiences a larger proportional increase in low-quality sales than in medium- or high-quality sales. In other words, lowering price attracts more ‘cheapskates’. The estimates suggest that moving from rank 1 to rank 7 (lowest price for the low quality modules to seventh lowest price) decreases sales of low-, medium-, and high-quality modules by 83 percent, 66 percent, and 51 percent, respectively.

Finally, using data on wholesale acquisition costs of goods and very detailed information about operational and other costs, Ellison and Ellison (2009) calculate markups and find slightly negative markups on low-quality products and substantial positive markups on higher-quality products, showing that the adverse selection the firms face provides an incentive for them not to lower the price of the (advertised) low-quality good so much as to compete away profits on the higher-quality counterparts.
Recalling the predictions of theoretical models of obfuscation, we find that this empirical evidence is consistent with those predictions. First, the low-quality, easy-to-search-for goods have extremely elastic demand, which implies razor-thin margins. Second, the higher-quality, hard-to-search-for goods have less-elastic demand, consistent with higher margins. Third, firms use the easy-to-search-for goods to attract customers, consistent with the negative cross-price elasticities in Table 12.1. Fourth, the existence of differential quality in the market creates an adverse-selection problem: firms that price too low will attract more ‘cheapskates’ unwilling to upgrade to higher quality. Estimated coefficients in Table 12.2 provide evidence for this prediction. Finally, in equilibrium, firms do not fully compete away higher margins on upgraded products by selling the easy-to-search-for good too far below cost.

Another paper providing more formal empirical evidence of obfuscation and its effects is Muir et al. (2013). They collected detailed data on auto driving schools in Portugal, a setting where standardized reporting of prices is rare, and issues such as price search and obfuscation might loom large. (Again, although attending a driving school is not a retail transaction, insights from consumer and firm behavior in this market are likely to be quite relevant for retail markets where price complexity could arise.) They estimate a model of demand where consumers have limited price information and where, in fact, their degree of price information is a function of measures of price complexity and search costs. They find results consistent with complexity significantly limiting consumer price information, suggesting that price complexity obfuscates.

Finally, Kalaycı and Potters (2011) offer a different type of evidence of the effects of price obfuscation. They carry out an experiment in which sellers decide on the number of attributes of their good and then set prices. The number of attributes of the good affects neither the cost to the sellers nor the value to the buyers but simply makes prices more complex (and more difficult to compare). Their results indicate that buyers make more suboptimal choices and that prices are higher when the number of attributes of the goods is higher.

5 TECHNOLOGY AND PRICE SEARCH

Throughout the chapter, I have alluded to characteristics of offline and online price search and suggested that online price search might have some important differences from what we might call traditional price search. In fact, a number of the empirical papers I have discussed use price data scraped from websites, and those prices typically are for goods purchased online. So they have de facto studied the mechanics and effects of price search online. I have skirted the issue, though, of how technologies such as the Internet could be altering price search. This last section will discuss a small number of papers which deal explicitly with this question.¹¹

At the dawn of e-commerce, academics, as well as consumers and media outlets, went in search of hard facts as to whether retail prices online were higher or lower than offline retail prices. These questions were not necessarily motivated by the academic questions about the mechanics of price search, but their answers were still relevant to our discussion here. One early study, Bailey and Brynjolfsson (1997), collected prices from samples of online and offline retailers in 1997 and reports that online prices for software, books,
and CDs were 3 percent to 13 percent higher on average than offline prices. Later studies such as Brynjolfsson and Smith (2000), Clay et al. (2001), and Ellison and Ellison (2009), however, tend to find that online prices are lower than offline prices. One exception is Ellison and Ellison (2014), which finds that used-book prices are often significantly higher online, with offline prices at the twenty-sixth percentile of online price distributions, on average. All of these studies note that there was substantial dispersion in online prices. Baye et al. (2004) offers an early look at electronics markets online. Although they were not explicitly comparing online prices to offline prices, they were responding to considerable public optimism at the time that the Internet would lead to widespread marginal cost pricing. They found instead that significant price dispersion was an equilibrium phenomenon and was here to stay.

Explicit comparisons of markets over time as search-aiding technologies have been introduced are rare. One can imagine why: gathering detailed enough data offline and online or before and after a technology is introduced is a difficult and daunting task. At least two papers, though, offer interesting examples of this type of study.

Jang (2014) observes a technology-driven change over time in the cost of price search. In particular, he analyzes the Korean retail market for gasoline as the government makes comprehensive daily gasoline price data increasingly available and easy to use through the launching of a website, development of iOS and Android apps, and smartphone penetration. He provides descriptive evidence that the increasing availability to consumers of price information in this market does not result in lower prices or in less price dispersion. He then rationalizes this somewhat counterintuitive result by estimating a two-type consumer search model – informed and uninformed – and showing that equilibrium price dispersion and levels can increase as more consumers become informed. After this, he performs counterfactuals with this structural model to calculate the likely effects of continued increases in consumer price awareness, as we would expect to occur as smartphone usage increases and apps to find low gas prices improve. Those counterfactuals indicate an eventual fall in price levels and dispersion.

Finally, Ellison and Ellison (2014) study the market for used books with an unusual data set covering both offline and online prices. They have approximately 300 titles found in brick-and-mortar used-book stores and the corresponding title-matched set online. Furthermore, there are cross-sectional differences across titles in how important internet search technologies should be. For instance, books with local interest, Christmas in Georgia, say, are likely affected less by Internet technologies because being on a shelf in a used bookstore in Atlanta might expose that book to a large fraction of its potential market anyhow. With these data, they can analyze the particularly dramatic and interesting transition of used book sales from brick-and-mortar stores to online.

They start with the observation that Internet technologies could have two distinct effects on the market for used books: First, online price search engines, such as AbeBooks.com, should make price search and comparison much easier, perhaps intensifying competition. Second, those same online price-search engines or other tools can help consummate transactions that would never have been consummated pre-Internet because a particular buyer looking for a particular title would find such a search prohibitively costly in the world of brick-and-mortar used-book stores. (We could think of offline search costs being infinite in that case.) The first effect could result in lower prices if competition is intensified. The second effect could result in higher prices, depending
on how the surplus from those extra transactions is divided. So, the large-scale natural experiment of the market for used books moving online was conflating two effects, a competition effect from better price search and comparison, and a match-quality effect from software to help buyers find specific titles. Simply comparing online price distributions with offline price distributions tells us little about each effect separately. (In introducing the important match-quality component, this analysis may have more elements in common with job search in labor markets than the typical empirical price search paper. A paper which documents match quality improvements from technology in a product-market setting, but without the improved price search effect, is Jensen, 2007.)

Ellison and Ellison’s (2014) approach, then, is three-pronged to allow them to disentangle the two effects. First, they present a model that highlights improved match quality between buyer and title for online purchases relative to offline purchases, improved price search technology for online purchases relative to offline purchases, and heterogeneity in how well-informed online buyers are. Second, they derive a number of predictions of their model which they can test using online and offline price distributions and do so using their data. They find broad support for the theoretical model, including for the differential predictions regarding local-interest titles as opposed to other titles. Third, they develop and estimate a structural model based on the theoretical model to estimate the welfare effects of the improved match quality and enhanced price search capability when the market for used books started moving online. They find that, despite significantly higher prices online, online shopping creates substantial increases in consumer surplus in addition to profits. The reason is that improved match quality results in online transactions that never would have occurred offline, generating surplus that was divided in some fashion between sellers and buyers.

Another way to think about this improved match quality is as a manifestation of the following observation: consumers use the Internet search technologies to find products, not just compare prices. A literature is emerging that explicitly considers this fact. Using very fine-grained data on consumer browsing behavior on eBay, Blake et al. (2014) offer fascinating examples of consumer search patterns. Some consumers know precisely what product they want to buy and engage in price search and comparison. Others navigate around for quite some time before settling on a specific product. The theory and empirics of consumer search have focused the bulk of their attention on the former. The latter is an interesting area for future research.

6 CONCLUSION

Retail-price search has been a topic for economic inquiry for more than five decades. My hope is that I have provided a broad outline of the directions that the theory literature on price search has taken as well as the empirical evidence on the importance of price search, its equilibrium effects, and how technologies such as the Internet have altered it. This is not a comprehensive review of all work done on price search, but I have tried to highlight some of the areas where I think research to date has been most informative as well as a couple of areas with interesting outstanding questions.

Empirical research on obfuscation has lagged behind a significant theory literature on the topic. Although examples of obfuscation are not hard to find, formal empirical
Price search and obfuscation

303
evidence testing models of obfuscation, for instance, has been rarer. I feel that this area could be an interesting one for future researchers to pursue. In addition, there is still a lot to learn about how technologies affect price search. The few examples that explicitly compare price search in similar markets with different technologies find evidence of consumer benefit. We saw examples, though, in the models of obfuscation, where improved technologies could plausibly aid obfuscation and, as a result, harm consumers. Empirical study of the effect of technologies on obfuscation could be an interesting additional step.

NOTES

1. This chapter grew out of a short series of lectures delivered by Glenn Ellison and me at the Center for Economic Studies (CES) in Munich, January, 2007. The lectures were on the theory and empirics of price search. I would like to acknowledge Glenn Ellison’s substantial role in those lectures as well as thank him for additional comments on this chapter.
2. Chapter 14 in this volume, by Murry and Schneider, discusses the economics of automobile dealerships.
3. In Bertrand competition, undifferentiated firms compete by simultaneously posting prices. The unique Nash equilibrium of this game is for firms to post price equal to marginal cost.
4. In Cournot competition, firms choose quantities instead of prices.
5. Note also that the definition of equilibrium is problematic in such models. Our normal equilibrium notion means that in a pure-strategy equilibrium, consumers would play as if they know all prices because the prices are the firms’ strategies. In a search model with a mixed-strategy equilibrium, this issue does not arise.
6. Despite the possibility of misinterpretation, the terminology I use has become standard, so I stick with it. Both ‘shoppers’ and ‘non-shoppers’ can and do purchase. The distinction is that a ‘shopper’ is a consumer with negative search costs who, in equilibrium, shops around to find the lowest price.
7. Hickman and Mortimer document the wide variation in assortments across similar retailers in Chapter 13 of this volume.
8. Hong and Shum also mention Albrecht and Axell (1984) and Rob (1985) as relevant antecedents to their procedure.
9. The gamma distribution is a two-parameter family of continuous distributions on the positive line, of which chi-squared and exponential are special cases.
10. The marketing literature has long recognized strategies such as loss leaders, formally similar to upselling or add-on pricing. In that literature as well, empirical evidence tends to be of the documentation variety.
11. A further discussion of e-commerce and online/offline competition can be found in Chapter 18 by Smith and Zentner later in this volume. Chapter 19 by Tadelis discusses online search in two-sided markets.

REFERENCES


Honka, E. and P. Chintagunta (2015), ‘Simultaneous or sequential? Search strategies in the U.S. auto insurance industry’, working paper, University of Texas at Dallas, TX.


Jang, Y. (2014), ‘The effects of the Internet and mobile search technologies on retail markets: evidence from the Korean gasoline market’, working paper, Massachusetts Institute of Technology (MIT), Cambridge, MA.


McDonald, S. and C. Wren (2014), ‘Multiple price posting and consumer search obfuscation: evidence from an online market’, working paper, Newcastle University, Newcastle upon Tyne, UK.


Spence, F. (2014), 'Consumer experience and the value of search in the online textbook market', working paper, University of Notre Dame, South Bend, IN.


