PRICING BEHAVIOR AND EQUILIBRIUM PRICE DISPERSION IN ONLINE MARKETS

by

Christine Saliba

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

Summer 2003

© 2003 Christine Saliba
All Rights Reserved
PRICING BEHAVIOR AND EQUILIBRIUM PRICE
DISPERSION IN ONLINE MARKETS

by
Christine Saliba

Approved: ____________________________________________________________
Saul Hoffman, Ph.D.
Chair of the Department of Economics

Approved: __________________________________________________________
Michael J. Ginzberg, Ph.D.
Dean of the College of Business and Economics

Approved: __________________________________________________________
Conrado M. Gempesaw II, Ph.D.
Vice Provost for Academic and International Programs
I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _______________________________________________________
Michael A. Arnold, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _______________________________________________________
Rene Sandretto, Ph.D.
Professor in charge of dissertation

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _______________________________________________________
Evangelous Falaris, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _______________________________________________________
Yves Crozet, Ph.D.
Member of dissertation committee
I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____________________________________________
Thierry Penard, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: ____________________________________________
James Dearden, Ph.D.
Member of dissertation committee
ACKNOWLEDGEMENTS

I am deeply indebted to Dr. Arnold for his tremendous help and guidance, not only with this thesis but throughout my graduate studies. Without his continuous encouragement and support this dissertation would not have been possible. I would also like to thank the other members of my committee, Dr. Sandretto, Dr. Falaris, Dr. Crozet, Dr. Penard and Dr. Dearden, for their constructive comments and their help with last minute deadlines. I am grateful to Dr. Sandretto, who is the person responsible for my coming to the University of Delaware, for ensuring my participation in the cotutelle program and allowing me to obtain a doctorate from Université Lyon2 Lumièrè. My sincere thanks to Deborah Sharpley for her help with all the administrative matters during my four years as a graduate student. I am grateful for the many friends I have. Each one of them helped me in his or her own way. In particular I would like to thank Hayette Meddah for her support during the last couple of months. I must also thank my parents Sarah and François. I wouldn’t be who I am without them.

Finally, my deepest love and appreciation to Adel Abumohor. This dissertation is dedicated to him. Thanks for your love and support during the last five years, you made life much more enjoyable.
# TABLE OF CONTENTS

**LIST OF FIGURES** .......................................................... vii  
**LIST OF TABLES** ........................................................... viii  
**ABSTRACT** ................................................................. ix  

**Chapter**

1 INTRODUCTION ............................................................. 1  
2 LITERATURE REVIEW ....................................................... 5  
   2.1 Theoretical Literature on Price Dispersion ....................... 5  
   2.2 Empirical Literature on Price Dispersion ....................... 12  
      2.2.1 Evidence of Price Dispersion in Traditional Markets .... 12  
      2.2.2 Evidence of Price Dispersion in Online Markets .......... 16  

3 PRICE DISPERSION IN ONLINE MARKETS: THE CASE OF  
   COLLEGE TEXTBOOKS ...................................................... 19  
   3.1 Introduction ......................................................... 19  
   3.2 The Model ............................................................ 24  
   3.3 Data ................................................................. 29  
   3.4 Testing the Hypothesis ............................................. 34  
   3.5 Conclusion .......................................................... 40  

4 AN EQUILIBRIUM MODEL OF ONLINE PRICE  
   COMPARISON SITES WHEN FIRMS HAVE DIFFERENT  
   MARKET SHARES .......................................................... 48  
   4.1 Introduction ......................................................... 48
LIST OF FIGURES

3.1 Equilibrium Price Dispersion ............................................. 27

4.1 Relationship between $F_{ri}/2$ and the market share of the small firm $s_i$ when $\mu = 1$. ................................................................. 68

4.2 The advertising fee $\phi$ as a function of the market share of the small firm $s_i$. ................................................................. 69
## LIST OF TABLES

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Textbook descriptive statistics.</td>
<td>43</td>
</tr>
<tr>
<td>3.2</td>
<td>Websites descriptive statistics.</td>
<td>44</td>
</tr>
<tr>
<td>3.3</td>
<td>Reduced form regressions.</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>Two-stage least squares regression.</td>
<td>46</td>
</tr>
<tr>
<td>3.5</td>
<td>Anova test using the Gap measure.</td>
<td>47</td>
</tr>
</tbody>
</table>
ABSTRACT

Ready access to price information on the Internet has led to many claims that ignorance will not persist in online markets, and that online prices for a homogeneous good will converge to the competitive price. However, a growing body of empirical research provides strong evidence of persistent price dispersion in online markets. This dissertation addresses the question of online price dispersion empirically and theoretically. The empirical part relies on a model that illustrates how price dispersion can exist in equilibrium even when all consumers are perfectly informed about the prices charged by all firms. The equilibrium relies on the fact that although prices are common knowledge, information about the availability of the product at any particular firm can only be obtained at a positive cost. The model is tested using data from the online market for college textbooks. The empirical results demonstrate that price dispersion exists in the online market for college textbooks, and support the predictions of the theoretical model that in an equilibrium with price dispersion some firms adopt a high-price, high-availability strategy while others adopt a low-price, low-availability strategy. In a second part I develop a theoretical model in which firms serve segmented markets of different sizes. A market for information is introduced in which a monopoly gatekeeper optimally sets advertising fees. In equilibrium, price dispersion appears on the gatekeeper’s site. The small firm advertises its prices more often than the large firm, and it advertises higher prices on average. Moreover if the small firm market share becomes too small relative to the big firm, the market for information can no longer exist, in which case price dispersion disappears as all firms charge the monopoly price.
Chapter 1

INTRODUCTION

Online markets are characterized by very low price search costs for consumers. This has lead to many claims that online markets should be highly competitive. Given that price dispersion for a homogeneous good is attributed in most of the economic literature to the high cost of acquiring price information and therefore to consumers’ ignorance of prices across different retailers, one would expect that price dispersion would disappear if price search costs were close to zero. Recent empirical studies suggest that this is not the case in online markets. Price dispersion exists, and is persistent on the Internet. The goal of the dissertation is to shed some light on the forces driving price dispersion on the Internet and influencing the degree of price competition in online markets. The dissertation contains two sections, one empirical and one theoretical, as well as a literature survey of theoretical price dispersion models and empirical evidence. The first section is an empirical study of price dispersion in the online market for college textbooks. The second section investigates the role that firm heterogeneity plays in sustaining price dispersion by extending existing theoretical models to allow for varying market shares across firms in a model with a gatekeeper.

The first part of the dissertation is an empirical study of the factors driving price dispersion in the online market for college textbooks. The empirical tests rely on Arnold’s (2000) theoretical model demonstrating that even if consumers are fully informed of the price charged by each retailer, price dispersion can exist in
equilibrium if consumers incur a positive search cost in order to obtain availability information. In this model firms are capacity constrained and may run out of stock. In an equilibrium with price dispersion buyers adopt a symmetric mixed search strategy in which they are more likely to search low-price firms than high-price firms and low-price firms are less likely to have the item in stock. In equilibrium an individual buyer cannot increase his or her utility by changing his or her search strategy. Furthermore, because buyers realize that availability at a given firm is decreasing in the probability that the firm is searched and temper their response to a price reduction accordingly, firms have no incentive to compete by cutting price. Therefore in equilibrium there exists a positive relationship between price and availability. The data used for the study were collected from July 30th to October 22nd, 2001. For the thirteen-week period I gathered information on twenty-one textbooks from eight leading online college textbook retailers. The data include information on prices, availability, shipping costs and methods, and return policies. The data show evidence of the existence of price dispersion in the online college textbook market. Several econometric tests were conducted to examine the relationship between prices and availability. All of these tests give strong evidence supporting the hypothesis that firms adopt either a high-price, high-availability or a low-price, low-availability strategy. The contribution of this study is to present empirical evidence consistent with theoretical results explaining price dispersion as an equilibrium phenomenon in an environment in which price information is readily available.

Existing theoretical models of pricing in online markets generally make simplifying assumptions about firm size. Baye and Morgan’s (2001) model of firm pricing in a model with a gatekeeper assumes that all firms have identical market shares. This assumption does not match reality. Internet markets are characterized by firms of different sizes. For example, in the market for textbooks we find that there are few websites with a large market share, like Amazon.com and Barnes&Noble.com
and many smaller retailers, with smaller market shares. The second part of the dissertation explores the implication of firm heterogeneity for equilibrium pricing in online markets, by introducing firm heterogeneity into Baye and Morgan’s (2001) model of price dispersion in an online market with a gatekeeper. One way of obtaining full price information on the internet is to use shopbots or price comparison sites (gatekeepers). These websites are information intermediaries. They can charge firms an advertisement fee and consumers an access fee. The goal of the second part of the dissertation is to contribute to the literature on the effect of these gatekeepers on the price dispersion of a homogeneous good. Baye and Morgan present a theoretical model in which a gatekeeper (the information provider) allows firms, for an advertising fee, and consumers, for a subscription fee, to expand their markets which otherwise would be limited to a local market. Gatekeepers are particularly important in online markets because of the high uncertainty about sellers location on the Internet. Consumers are only familiar with few websites names in each market. These websites have large market shares. Many other websites are unknown to most consumers and therefore have only small market shares. The introduction of a gatekeeper allows firm to advertise not only their prices but also their existence to consumers. Thus the gatekeeper allows them to expend their market, but at the same time it imposes on them fiercer competitive environment. When a firm does not advertise on the gatekeeper’s site it sells only to consumers from its local market, consumers visiting it directly. A firm can charge the monopoly price to these consumers. When advertising on the gatekeeper’s site a firm charging the monopoly price with certainty will systematically be undercut and would never make any sale. In equilibrium, firms advertise on the gatekeeper’s site using mixed strategy balancing the benefit from the bigger market it can reach with the higher competition forcing it to lower its price. Firms also use mixed pricing strategies and prices are
dispersed on the gatekeeper site. Baye and Morgan assume that firms are homogeneous, with identical local market shares. The contribution of my research is to allow firms to have different market shares. The most important result of the model is that the existence of the gatekeeper depends on the relative market share of firms as well as the ratio of consumers participating in the market for information. In the case with a gatekeeper, the overall probability that firms choose to advertise the monopoly price increase as firms’ market shares diverge. In addition, large firms behave more competitively than small firms with respect to the price they charge – a firm with a large market share sets a lower price when it advertises than a firm with a small market share. On the other hand the firm with the small market share advertises with a higher probability than the firm with a large market share.

The dissertation proceeds as follows. Chapter 2 includes a literature review of theoretical and empirical price dispersion. The empirical study of price dispersion in the online textbook market is in chapter 3. Chapter 4 presents the theoretical study of the role of firm heterogeneity on their decision to advertise and price on a gatekeeper’s site. Finally chapter 5 concludes.
Chapter 2

LITERATURE REVIEW

2.1 Theoretical Literature on Price Dispersion

In his famous 1961 article, Stigler introduces the notion of search. Search is defined as the process of acquiring information about a product’s price across different sellers (or across different buyers). If consumers are not fully informed about prices charged by different retailers for a homogeneous good, and getting this information requires costly search, then different retailers may charge different prices in equilibrium. In other words price dispersion may exist and it is “the manifestation of ignorance in the market”. Stigler argues that search is more costly for sellers because of the high uncertainty about the location of buyers, and therefore buyers will be the ones to engage in search. Sellers may use devices like advertising or localized transactions to reduce buyers’ search costs and induce them to enter the market (such devices will also reduce price dispersion). Buyers conduct search and buy at the lowest price revealed by their search. In Stigler’s model, search does not reveal the same information to all buyers, and therefore some buyers become more informed than others and buy at a lower price. In equilibrium high-price sellers sell only to the buyers with “non-productive” search and have lower sales than the low-price sellers. Stigler also argues that if conditions on the market are stable (i.e. buyers and sellers are fixed in the long run) then price dispersion must disappear as consumers’ search reveals all the price distribution in the market. On the other hand, higher instability of supply and demand increases the size of price dispersion.
Stigler contends that as the market size increases, it will become profitable for a third-party to sell information to the buyers about the price distribution (in 1961 examples of these information providers were newspapers or trade journals, today we can think of shopbots or internet information gatekeepers providing this service). Stigler predicts that there will be a “monopoly in the provision of information”.

Following Stigler’s article several economists developed equilibrium models where consumers do not have full information. Diamond (1971) models a market with many identical firms and consumers. The model covers several time periods. In each period each firm sets a price. Consumers are not aware of all prices across all firms. Each consumer visits one firm and either purchases the good, according to an underlying demand curve, or concludes that the price is too high and leaves the market to reenter the following period. Consumers face uncertainty about prices in the future and must compare the expected cost of searching with the expected gain from further search. Firms have perfect information. They know the reservation price of consumers, their demand functions and their search costs. Given the market structure, the equilibrium prices at different firms converge to the monopoly price, and there is no price dispersion in the market.

Salop and Stiglitz (1977) model a market in which consumers are imperfectly informed about prices at different retailers and search is costly. Their major assumption is that consumers are heterogeneous with respect to their search costs. There exists two types of consumers, high-search cost consumers and low-search cost consumers. For both types of consumers a fixed cost search (that is high for some consumers and low for others) reveals all price information on the market. Firms are fully informed about the two search costs and the proportion of each type of consumers in the market. Consumers must decide whether to engage in search (and have a total expenditure of search fixed cost plus the minimum price of the good) or shop randomly (and have an expected total expenditure of the mean price charged
in the market). Consumers want to minimize their expected total expenditure, and therefore search only if the difference between the mean price and the minimum price charged on the market is greater than the fixed search cost. Each firm knows this information and maximizes its profits accordingly (taking other firms’ prices as given). A necessary condition for an equilibrium to exist is the model’s zero profit assumption. To respect this requirement and obtain price dispersion, Salop and Stiglitz use U-shaped average cost curves. Given these assumptions, a one-price equilibrium is not possible in this market. Two-price equilibria are possible. In an equilibrium with price dispersion the low-search cost consumers always search and buy from the low price seller and the high-search cost consumers never search and shop randomly. Firms charge one of two prices: low price (same as the competitive price) or high price (either the monopoly price or a price just high enough so that high-search cost consumers don’t search). The firms charging the low price sell to the low-search cost consumers and to some lucky high-search consumers. Firms charging the high price sell exclusively to high-search cost consumers. Another interesting characteristic of this model is that equilibria with more than two prices are not possible even with more than two consumer types.

Unlike Salop and Stiglitz, Butters (1977) assumes that consumers are ex-ante homogeneous. Consumers are looking to buy one unit of a homogeneous good sold by several retailers that strategically choose their price. In this model buyers do not engage in any search, but firms engage in costly advertising. Firms are fully informed about the number of potential buyers on the market and their purchasing strategies. Each firm chooses the number of ads to send to consumers as well as the prices to advertise (firms can choose to send two ads with two different prices) in a profit maximizing way. Ads are allocated randomly among consumers in such a way that a firm cannot predict which consumer will receive which ad or whether the same consumer will receive more than one ad. Buyers receive a number of ads, observe the
different prices and buy at the lowest advertised price (if a buyer receives zero ads, he or she makes no purchase). Butters argues that with the above assumptions no one-price equilibrium exists. He even generalizes the result to assert that “not more than one seller can be advertising at any given price”. In equilibrium with a large number of sellers (approaching infinity), an infinite number of prices are advertised. Sellers’ advertising strategies are “completely indeterminate”. In equilibrium all advertised prices lead to zero expected profits.

Varian (1980) examines temporal price dispersion instead of the traditional spatial price dispersion. In a temporal price dispersion model there are several periods, and each store varies its price over time. It is assumed that there are two, exogenously determined, types of consumers (informed and uninformed) who wish to buy one unit of a homogeneous good. The informed consumers buy at the lowest advertised price during a given period and the uninformed consumers shop at random. Each period firms set their prices according to a mixed-strategy. The firm with the lowest advertised price during a given period sells to the informed consumers plus a part of the uniformed consumers. A firm does not set the lowest price on the market sells only to a part of the uninformed consumers. In equilibrium expected profits for each period are zero for all firms. In Varian’s model there exists no symmetric equilibrium where all the stores charge the same price.

Shilony (1977) examines a segmented market (according to location, but also possibly according to brand). Each segment of the market consists of a number of consumers and one seller. Consumers can purchase costlessly from the store in their sub-market, but must engage in costly search if they want to acquire information about prices in other sub-markets. Shilony shows that no Nash equilibrium exists in pure pricing strategies. On the other hand, he establishes the existence of an equilibrium mixed strategy – that is, a strategy where firms randomize their prices. In a dynamic setting, in each period sellers decide a price to advertise and buyers
observe all prices and decide on their buying store according to their location and the distribution of prices.

In a 1982 article Salop and Stiglitz abandon the assumption of having two types of consumers. All consumers are ex-ante identical and know the price distribution in the market. Each consumer lives for two periods. She enters the market the first period and shops randomly. A consumer never undertakes a second search in the same period. She then has the choice between purchasing two units and storing one unit at a positive cost for consumption in the second period or purchasing one unit and reentering the market the second period at a positive cost. In making their decisions, consumers compare the cost of purchasing two units and storing one with the expected price to be paid in the second period plus the cost of reentering the market. Given some restrictions on the storage cost as well as the cost of reentering the market, an equilibrium with two prices exists. In equilibrium the high price is the monopoly price and the low price is the reservation price that leaves consumers indifferent between buying two units or reentering the market the second period. “Lucky” consumers are consumers who shop at the low price retailers and they buy two units in the first period. “Unlucky” consumers are consumers shop at a high price retailer, buy one unit in the first period and reenter the market in the second period. At the beginning of the second period new consumers are born. They enter the market considering this period to be their first. Both types of firms make zero profits, but the low price firms sell a larger quantity of units than the high price firms.

Burdett and Judd (1983) examine a market with identical consumers and firms. Consumers are not fully informed about prices charged by different retailers for a homogeneous good. The authors argue that price distribution may appear in equilibrium if there is positive probability that a random consumer may know only one price. In other words a necessary condition to obtain price dispersion is, ex-post,
to have consumers with heterogeneous information. Burdett and Judd examine non-sequential search strategies with noise. Non-sequential search may reveal different information to different consumers. Even though ex-ante consumers are homogeneous, ex-post consumers become heterogeneous. Employing non-sequential search strategies, they do not obtain full information about prices at different retailers, and they buy from the cheapest store in their sampled retailers. This model offers two equilibria: a one-price equilibrium with the monopoly price or a dispersed price equilibrium (a one-price equilibrium always exists with non-sequential search, whereas the dispersed price equilibrium depends on some coefficients).

In Wilde and Schwartz 1979’s model consumers are imperfectly informed of the prices sellers charge. They can receive price quotations by investing in search activity. Consumers are ignorant of the price distribution and therefore the type of search they engage in (if they decide to search) is a non-sequential search. All firms are assumed to be subject to the same capacity constraint, and each firm must charge the same price to all buyers. The Consumers that decide to search visit more than one store before making a decision on where to purchase the good. Consumers are also assumed to be heterogeneous. Some consumers always engage in search: they always visit more than one store before purchasing. These consumers are called shoppers. The other type of consumers randomly samples one store before purchasing. They are called non-shoppers. The equilibrium of this model depends on the proportion of shoppers to non-shoppers in the population. Price dispersion exists in equilibrium and is the result of having the non-shopper consumers who are only aware of one price in the market. If relatively few shoppers exist, the equilibrium price distribution is continuous over a certain interval. As the proportion of shoppers increases, more firms charge the competitive price.

Salop (1977) models a market with a monopolist and two subgroups of consumers. Efficient searchers, with low search cost and inefficient searchers, with high
search cost. Consumers use sequential search strategies. The monopolist uses price
dispersion to discriminate among the two subgroups of consumers. The efficient
consumers pay a lower price on average than the less efficient consumers. The less
efficient consumers have a less elastic demand than the efficient consumers. With the
right set of coefficients about elasticities in the two sub-markets, price discrimination
is more profitable than having a single price.

Reingamun (1979) is interested in demonstrating that price dispersion can
appear in a setting very close to a simple setting. A simple setting refers to a market
with the following assumptions: identical consumers who demand exactly one unit
of a homogeneous good, consumers search sequentially, the cost of search is fixed
per observation, and firms are identical and compete monopolistically. The two
deviations from the simple setting Reingamun’s model relies on are heterogeneous
costs among firms, and nonzero elasticity of consumers’ demand curves (consumers
do not demand a single unit of the good). Buyers know the price distribution in the
market but do not know which firm sells at which price. Buyers can sample with
recall from the known distribution, using a sequential search strategy. In equilibrium
there may exist a nondegenerate price distribution for a homogeneous good. Buyers
do not purchase the same amount of the product for all prices at or below the
reservation price. Because of buyers’ substitution effects and income effects, it is
not optimal for firms to charge the reservation price.

In Baye and Morgan’s 2001 model there are several geographically separate
towns each served by a local firm. Transaction costs preclude consumers living in
one town from shopping in another town. Each local firm acts like a monopoly. The
Internet is introduced and a gatekeeper offers to the local firms the possibility of
advertising on its site and therefore the possibility of selling to consumers outside
their geographical area. Firms observe the advertising fee, they then decide whether
to advertise on the gatekeeper site or not, and they must also decide the price to
charge (whether they advertise or not). Consumers can subscribe to the Internet and for a fee they can observe the prices advertised on the gatekeeper site and purchase at the lowest price firm (even if it is different from their local firm). Non-subscribing consumers can only visit their local firm and purchase the good only if the price is equal to or below their reservation price. Consumers observe the subscription fee and decide whether to use the gatekeeper site or not. The monopoly gatekeeper takes into account all the decisions made by firms and consumers, maximizes its profits and sets the advertising fees charged to firms and the subscription fees charged to consumers. The gatekeeper finds that its profits are maximized when there is price dispersion and all consumers subscribe to its service. The gatekeeper behavior induces price dispersion. In equilibrium, the gatekeeper charges a fee for firms that does not induce the participation of all firms, and a subscription fee that induces all consumers to subscribe. A firm that does not advertise its prices on the gatekeeper’s site charges the monopoly price. Firms advertise using mixed strategies. Firms use mixed advertising and pricing strategies when advertising on the gatekeeper’s site. In conclusion, establishing a market for information leads to more competitive pricing on the part of firms. Prices posted on the site are closer to competitive levels than the monopoly price charged by local firms, but they remain above marginal cost.

2.2 Empirical Literature on Price Dispersion

2.2.1 Evidence of Price Disperison in Traditional Markets

Pratt, Wise and Zeckhauser (1979) gather empirical evidence on the existence of price dispersion in several markets. They also show the source of such price dispersion. First they model a market where consumers do not have full price information but know the price distribution in the market. Search costs are fixed per observation, and consumers use sequential search with the possibility of recall. If the number of sellers is infinite, the monopoly price is an equilibrium. On the
other hand if the number of sellers is small, the monopoly price may no longer be an equilibrium, and price dispersion may appear. A second model is presented where consumers learn about the price distribution as they search. An equilibrium exists in this model where price do not converge to a unique price, but exhibit dispersion.

To examine the empirical price dispersion, the authors collect data on fifty products selected randomly. Sellers listed in the Yellow Pages were contacted and asked about their selling price. The reported price dispersion was substantial. For eighteen of thirty-nine products, the highest price was over twice the lowest. The authors also find a positive relationship between the standard deviation and the mean price of a good. The explanations given for such high price dispersion include positive buyers’ search costs, heterogeneity among consumers, but also quality differences (firms’ location or services), different geographic markets, or different sellers’ costs.

Dahlby and West (1986) use data on automobile insurance in Alberta to examine the relationship between price dispersion and consumers’ search costs. They choose insurance to conduct their study because it is a homogeneous good, with significant amount of price dispersion. They base their study on a model by Carlson and McAfee (1983). They collect data on premiums for the period 1974-1981 for 54 companies in 1974 and 61 in 1981 (respectively 83.2 and 93.5 percent of the industry). They found that price dispersion occurred in all driver classes, for all geographical territories in Alberta for all years (they use the ratio of maximum premium to minimum premium, the mean, the standard deviation and the coefficient of variation to document price dispersion). After showing that price dispersion exists, the authors test some predictions made in the theoretical models. First they test whether firms’ premiums are correlated across driver classes (this would indicate that some firms charge on average higher prices than others). They find that firms charging higher (or lower) premiums at the end of the sample period were also charging higher (or lower) premiums over time. Second they test whether a firm’s
relative market share varies inversely with the deviation of the firm’s premium from the average premium for a particular class. This would show that firms charging higher prices sell fewer units. They find that this relationship holds for the market for drivers over age 25 but it does not hold for drivers below age 25 (except married males). One explanation they propose for the poor results of the below age 25 market is that drivers in this category do not conduct insurance search but rather purchase insurance in the same company as their parents. They finally test whether the variance of premiums is lower in the driver classes where more search occurs. They find that this relationship holds in their data. Their overall conclusion is that there is evidence price dispersion in the automobile insurance market in Alberta can be explained by costly consumer search.

Adams (1997) uses a model based on Salop and Stiglitz (1977) to test empirically whether higher search costs per unit per customer result in having higher price dispersion (the highest price becoming much higher than the lowest price). In other words he is testing whether a reduction in search costs to consumers causes the degree of price dispersion for a given product in a particular market to fall. To conduct his test, Adams collects data from 20 convenience stores in Auburn/Opelika Alabama from November 13 through November 16, 1994 on gasoline prices and on 22 in-store items. The author uses the same locality for all the convenience stores in his data set in order to avoid locational asymmetry that could be responsible for price dispersion. The reason he collects these two sets of data is the difference in search costs between gasoline and in-store items. Gasoline price are usually displayed on an exterior sign that is easily visible to consumers from many feet away (low search costs), whereas in-store items prices are only obtained by entering the store (high search costs). First the author tests for the loss-leader hypothesis, that convenience stores charge low gas prices to attract customers and then charge them high in-store items prices. He finds that this hypothesis is not supported by the
data. Then the author tests whether the degree of price dispersion for gasoline is lower than price dispersion for in-store items. Two arguments are given in favor of this hypothesis. First search costs for gasoline are lower than search costs for in-store items. Second, gasoline represents a higher percentage in the budget of most households than in-store items, which leads one to believe that consumers will be more price sensitive to gasoline price and would search more for the lowest available price. The sample variance for each item is used as a measure of price dispersion. The author finds that 20 of the 22 in-store items exhibit a statistically significantly greater degree of price dispersion than the degree of dispersion of gasoline. This result supports the tested hypothesis.

Sorensen (2000) examines the retail market for prescription drugs to study the price dispersion that arises from imperfect information. As in Adams’ paper the main assumption in this paper is that items with frequent purchases, or high proportion in the expenditure will induce more search and therefore lower the price dispersion in that product’s market. For prescription drugs, Sorensen’s hypothesis is that drugs that need to be purchased frequently will exhibit lower price dispersion than drugs that are purchased infrequently. The author collects data from individual pharmacies in upstate New York on March 28, 1998 for the 152 top-selling prescriptions. This paper documents that price dispersion exists in the drug prescription market and that it is substantial. The author finds that unlike in the automobile insurance market examined by Dahlby and West (1986) price rankings are inconsistent across drugs in the same pharmacy. The author estimates that pharmacy effects on price dispersion must not be very significant and must only account for about one-third of the variation in drug prices about their means. The main finding of this paper is that observed price distributions are consistent with the theoretical hypothesis. Measures for both dispersion and markups are significantly lower for
drugs that are purchased repeatedly. Price ranges for one-time prescriptions are estimated to be 34 percent larger than those for prescriptions that must be purchased monthly. Absolute markups for one-time prescriptions are estimated to be 41 percent higher than those for prescriptions purchased monthly, other things equal. This result is consistent with Adams’ (1997) result about gasoline and in-store items price dispersions.

2.2.2 Evidence of Price Dispersion in Online Markets

Clay, Krishnan and Wolff (2001) examine the relationship between competitive structure, advertising, price and price dispersion on the Internet. Their data include five categories of books. Price dispersion is measured by the standard deviation of unit price, the difference between minimum and maximum price, the difference as a percentage of average price and the percent deviation as a percentage of average price. First the authors examine the effects of market structure on pricing behavior. They find that the presence of an additional big firm (Amazon, Barnes and Noble or Borders) lowered the price in the market. The presence of a fringe firm also lowered the price. More competition drives the price down. Second the authors examine price dispersion. They find that price dispersion is substantial among firms, but it is not significant for the same firm across periods (evidence that online stores do not change prices often). When they examine the effect of market structure on price dispersion they find that more competition leads to a lower standard deviation. Also more heavily advertised books have lower price dispersion. The authors explain some of the price dispersion they find as arising from differences in firms’ strategies. A substantial number of independent physical bookstores have established websites and are listed in major comparison shopping engines despite their relatively high prices. These stores seem to view the Internet as adding value primarily by advertising their physical stores and providing additional services for existing customers. On the other hand, some stores appear to offer lower levels of
branding than the big three combined with prices that are only marginally below Amazon’s prices. Two of the eight stores in this category have since been bought out and one store has failed.

Morton, Zettelmeyer and Silva-Risso (2001) examine the effect of Internet car referral services on dealer pricing of automobiles in California. They find that dealerships offer lower prices to consumers who arrive via the Internet than to other consumers. They conclude that prices that are offered to consumers are a choice variable of the dealers. The authors cannot distinguish among several potential causes of the lower prices. They suggest that they can be the result of better consumer information, bargaining by the referral service, or potentially lower costs of selling online.

Goolsbee (2001) uses a micro data set on individual computer purchases and estimates the sensitivity of venue choice (Internet versus traditional retail markets) to variations in the relative price. The data used are data on computer purchases from the December 1998 proprietary mail survey of Forrester Research. The author’s results suggest that there is significant competition in the case of computer equipment, especially for sales to experienced computer users and desktop buyers. The author concludes that the variation in retail prices has a significant impact on the likelihood of buying directly from the manufacturer. The conditional cross-price elasticity, overall, of buying remotely versus buying in a store with respect to the retail price is in excess of one and suggests that online and offline sales of computers are unlikely to be truly separate markets.

Lynch and Ariely (2000) conduct an experiment to test conditions under which lowered search costs increase or decrease price sensitivity. In their experiment they vary independently three different search costs via electronic shopping; search cost for price information, search cost for quality information within a given store
and search cost for comparing across two competing electronic stores. Seventytwo MBA and Ph.D. students and staff participated in an experiment using an electronic shopping system similar to Vineyards.com. Respondents were told that wines would be sold at significant discounts relative to prices for the same wines from area merchants. There were two competing electronic wine merchants with a total of 100 wines. The results of the experiment are as follows. First, lowering search cost for quality information reduced price sensitivity. Second, price sensitivity for wines common to both stores increased when cross-store comparison was made easy. Third, making information environments more transparent by lowering all three search costs produced welfare gains for customers.
Chapter 3

PRICE DISPERSION IN ONLINE MARKETS: THE CASE OF COLLEGE TEXTBOOKS

3.1 Introduction

The development of the Internet as an electronic marketplace provides yet another example of the relevance of Stigler’s (1961) seminal work identifying the important role that information plays in determining market prices. Stigler argued that “price dispersion is a manifestation ... of ignorance in the market (p. 214).” Several years later a large literature evolved investigating the existence of price dispersion in equilibrium. In general, this literature relies on positive search costs and some degree of consumer heterogeneity to generate equilibrium price dispersion.\(^1\) However, consumer ignorance about market prices, and therefore price dispersion,

\(^1\) For example, Salop and Stiglitz (1977) assume that the cost of gathering information differs across two or more classes of consumers. Butters (1977) assumes that advertisements are randomly allocated across consumers, so different consumers receive different price information. Salop and Stiglitz (1982) assume two types of consumers, lucky and unlucky, who expend the same search cost but receive different quantities of information. Burdett and Judd (1983) assume noisy search that reveals different information to different consumers. Baye and Morgan (2001) demonstrate that price dispersion occurs in equilibrium in an Internet market with a monopoly gatekeeper and two types of customers, those who subscribe to the gatekeeper’s services, and those who do not.
does not exist if search costs are zero. Consumers only buy at the lowest price and equilibrium price equals marginal cost.\(^2\)

Access to price information at a very low cost provided by the Internet has spurred claims that online markets should be highly competitive.\(^3\) In effect, these claims imply the consumer ignorance highlighted by Stigler should not exist in online markets. These arguments are bolstered by the fact that a few clicks of the mouse provide consumers with price information from several retailers. Moreover price comparison sites or shopbots, such as MySimon.com and Shopper.com, effectively eliminate the cost of comparing price information for a homogeneous good across several online retailers. Given the low cost of obtaining and essentially zero cost of comparing prices online, the claims that competitive pricing should prevail are consistent with the theoretical models of equilibrium price dispersion which indicate that price dispersion cannot persist in the absence of ignorance about prices.

A substantial and growing body of empirical evidence suggests that the claims

\(^2\) If consumer heterogeneity is eliminated and search costs are positive, so all consumers are equally ignorant, then all firms sell at the monopoly price, as was demonstrated by Diamond (1971).

\(^3\) For example, a November 18th, 1999 article in *The Economist* argued that “the explosive growth of the Internet promises a new age of perfectly competitive markets. With perfect information about prices and products at their fingertips, consumers can quickly and easily find the best deals. In this brave new world, retailers profit margins will be competed away, as they are all forced to price at cost.” Similarly, in their discussion of what the collapse of the Nasdaq tells us about the role of information technology, DeLong and Summers [2001] argue that “innovations in information and communications technologies are competition’s friend....In the past, you could comparison-shop only by trudging from store to store. In the present, you can use the World Wide Web. ... Thus, in the ‘new economy,’ more markets will be contestable. Competitive edges based on past reputations, or brand loyalty ... will fade away. As they do so, profit margins will fall: Competition will become swifter, stronger, more pervasive, and more nearly competitive.”
are not correct.\textsuperscript{4} Despite the low cost of search on the Internet, price dispersion exists in online markets as it does in traditional markets. Furthermore, the empirical evidence indicates that this price dispersion is persistent. It is increasingly implausible to attribute price dispersion online to the immaturity of online markets.

The goal of this paper is to offer insight into how price dispersion can persist in online markets despite the facts that price information is readily available on the Internet, and price comparisons can be easily conducted using shopbots. We present a simple theoretical model which demonstrates that even if prices are common knowledge, costly search to obtain information about other product characteristics of value to the consumer, specifically, information about product availability, can lead to price dispersion in equilibrium. The model predicts that in an equilibrium with price dispersion some firms adopt a high-price, high-availability strategy while others adopt a low-price, low-availability strategy. Data from the online market for college textbooks are used to test the model. The empirical results identify considerable price dispersion in this market and support the predictions of the theoretical model.

Although most models of equilibrium price dispersion rely on consumer ignorance about prices to support more than one price in equilibrium, recent theoretical work by Bakos (1997) suggests a potentially fruitful extension of earlier models that is particularly pertinent to online markets and may lead to insight into how equilibrium price dispersion can arise in a market with perfect information about prices. Bakos' innovation is the introduction of separate consumer search costs for identifying price information and information about other product characteristics. Harrington (2001) demonstrates that while no symmetric, pure-strategy equilibrium exists in Bakos' model as the cost of searching for price information goes to zero, Examples include Clemons, Hann and Hitt (1998), Bailey (1998), Brynjolfsson and Smith (1999), Clay, Krishnan, and Wolff (2001), and Baye, Morgan and Scholten (2001, 2002).
the issue of whether other equilibria exist remains an interesting topic for future research. This work raises the following question. If factors other than price are important to consumers, then might price dispersion arise and persist even if price information is freely available? Arnold (2000) demonstrates that under the assumption that prices are common knowledge – consumers incur no search cost to compare prices across firms – if obtaining information about the availability of a homogeneous good is costly to consumers, then an equilibrium with price dispersion can exist. In this model, firms are capacity constrained and buyers are unable to observe availability without first incurring a positive search cost. In an equilibrium with price dispersion buyers adopt a symmetric mixed search strategy in which they are more likely to search low-price firms than high-price firms, and low-price firms are less likely to have the item in-stock. In equilibrium an individual buyer cannot increase his utility by changing his search strategy. Furthermore, because buyers realize that availability at a given firm is decreasing in the probability that firm is searched and they temper their response to a price reduction accordingly, firms have no incentive to compete by cutting price.5

To test the hypothesis that unobserved availability can explain price dispersion in online markets, we investigate pricing behavior in the online market for college textbooks. For a period of thirteen weeks we gathered information on twenty-one textbooks from eight leading online college textbook retailers. The data include information on prices, availability, shipping costs and methods, and return policies. In the empirical analysis below, we first verify that price dispersion exists in the

---

5 Dana (2001) analyzes the impact of unobserved availability in a model in which firms compete in both price and availability. His restriction to the case of symmetric equilibria precludes the equilibrium price dispersion found by Arnold. Baye and Morgan (2001) demonstrate that equilibrium price dispersion can arise in a market with a profit maximizing monopolist gatekeeper (shopbot). However, this equilibrium breaks down if all consumers observe prices at zero cost.
online college textbook market, and then proceed to test the relationship between price and availability for a given textbook. We find strong evidence supporting the hypothesis that firms adopt either a high-price, high-availability or a low-price, low-availability strategy. In contrast to most previous empirical studies which document persistent price dispersion in online markets but do not explain it, we present evidence consistent with theoretical results which explain price dispersion as an equilibrium phenomenon.

Baye, Morgan, and Scholten (2002) also directly investigate the cause of persistent price dispersion in an online market. Using data for thirty-six consumer electronic products they demonstrate that characteristics which can be observed and compared directly using Shopper.com explain an average of seventeen percent of the price dispersion for the products in their sample, and the identity of the website posting the price explains an additional 55 percent of this dispersion. Their findings are consistent with Varian’s (1980) model in which price dispersion persists over time because firms adopt mixed pricing strategies in equilibrium. Varian argues that if firms adopt pure pricing strategies, price dispersion will disappear over time as buyers learn the firms’ strategies and only purchase from firms offering the best price. Our goal differs from theirs in that we investigate whether product characteristics which can be compared across websites only after incurring a positive search cost can explain the persistence of price dispersion online, after controlling for the product and for firm specific effects. Our results suggest a related but alternative explanation of persistent price dispersion to that offered by Varian. While firms in our model adopt pure pricing strategies and buyers all know these prices, price dispersion persists because buyers must incur a search cost to determine, with certainty, whether a given firm has the item in-stock, and buyers are willing to pay a premium for the increased probability that the item will be in-stock offered by firms charging above average prices.
In Section 3.2 we present a simple theoretical model. In Section 3.3 we describe the data used in the empirical study. We test the hypothesis and discuss the empirical results in Section 3.4. Section 3.5 concludes.

3.2 The Model

In this section we present a simple theoretical model to illustrate the important role that unobserved availability can play in a market in which consumers can costlessly compare prices across several firms. The model is suggestive of a more general relationship between a positive search cost to obtain information about product characteristics other than price which are important to the consumer and the existence of equilibrium price dispersion.

Consider the following simplified version of Arnold’s (2000) model. There is a continuum of consumers of measure $\lambda$, each with a common valuation $v$ for a homogeneous product offered by one of two firms. Each firm has a fixed, commonly known capacity constraint parameter $k$ which impacts the probability that the firm has the item in-stock when a customer arrives. Firms simultaneously post prices and buyers observe these prices at no cost. However, buyers incur a search cost to visit the firm and determine whether the item is in-stock. To capture the disutility of searching a firm which does not have the item in-stock, represent the buyer’s utility from searching firm $i$ by $u_i = v - p_i - c/\alpha_i$ where $\alpha_i$ is the probability that firm $i$ has the item in-stock.

Availability $\alpha_i$ at firm $i$ depends upon the market demand parameter $\lambda$, the firm’s capacity parameter $k$, and the probability $\pi_i$ that

---

6 This utility can be thought of as the result of a search problem with value $u$ in which the consumer receives a surplus of $v - p_i$ if the item is in-stock, and the problem restarts if the item is out-of-stock, so $u = \sum_{i=1}^{n} \pi_i [-c + \alpha_i (v - p_i) + (1 - \alpha_i)u]$. Noting that in any equilibrium mixed search strategy the expression in brackets must be equal for all $i$ and that $\sum_{i=1}^{n} \pi_i = 1$, the utility from searching any firm $i$ simplifies to $u_i = v - p_i - c/\alpha_i$. 24
buyers search firm $i$. To keep the model simple, assume $\alpha_i = k/(k + \lambda \pi_i)$.\(^7\)

We consider subgame perfect equilibria of the game in which firms simultaneously announce their prices and buyers then adopt a symmetric mixed search strategy $\pi \equiv \{\pi_1, \pi_2\}$. We limit our analysis to equilibria in which $\pi_1 + \pi_2 = 1$ (so choosing not to search is not an option – the buyer must buy the good), and assume $v < c + 3\lambda c/2k + \lambda^2 c/2k^2$.\(^8\) Note that in equilibrium $u_1 = u_2$. Otherwise, all buyers would search the firm offering the higher expected surplus and the other firm would reduce its price. Setting $u_i = u_j$, substituting the above expression for $\alpha_i$, and imposing the condition that $\pi_1 + \pi_2 = 1$, the buyers’ equilibrium search strategy $\pi^* \equiv \{\pi_1^*, \pi_2^*\}$ must satisfy

$$\pi_i^* = 1/2 + k(p_j - p_i)/2\lambda c.$$\(^9\)

Realizing the buyer response $\pi^*$ to the prices they establish, firms simultaneously choose prices to maximize their expected return $R_i$, where $R_i$ is simply the price $p_i$ times the expected number of units sold (the demand $\lambda \pi_i$ realized by firm $i$ multiplied by the probability $\alpha_i$ the item is in-stock when the customer arrives),

\(^7\) This assumption is not purely arbitrary. This value for $\alpha_i$ is equivalent to the steady state in-stock probability for an alternating renewal process in which buyers arrive at firm $i$ according to a Poisson process with rate $\lambda \pi_i$, out-of-stock items are restocked according to a Poisson process with rate $k$, and the process is “on” while the item is in-stock and “off” while the item is being restocked. Note that $\alpha_i$ is increasing in the capacity parameter $k$ and decreasing in both market demand $\lambda$ and the probability $\pi_i$ that buyers search firm $i$.

\(^8\) This restriction on $v$ is made to limit mathematical complexity while still illustrating the equilibrium price dispersion result. For a more general treatment see Arnold (2000).

\(^9\) More generally, in a market with $n$ firms the equilibrium must satisfy $u_1 = u_2 = \cdots = u_n$. This combined with $\sum_{i=1}^{n} \pi_i = 1$ implies that $\pi_i = 1/n + k[\sum_{j\neq i} p_j - (n - 1)p_i]/n\lambda c$. 

25
all minus a fixed operating cost $d$;\(^{10}\)

\[
R_i = p_i \lambda \pi_i \alpha_i - d
= \lambda(1/2 + k(p_j - p_i)/2\lambda c)(k/(k + \lambda(1/2 + k(p_j - p_i)/2\lambda c)))p_i - d.
\]

This maximization is done subject to the constraints that $\pi^*_i + \pi^*_j = 1$, and that $u_i \geq 0$. The first of these constraints implies $p_j + \lambda c/k \geq p_i \geq p_j - \lambda c/k$, and the second implies $p_i \leq 2(v - c) - \lambda c/k - p_j$. Imposing the restriction $v < c + 3\lambda c/2k + \lambda^2 c/2k^2$, deriving Kuhn-Tucker conditions for this maximization problem, and solving for firm $i$'s reaction function yields

\[
p_i = \begin{cases} 
  p_j + \lambda c/k + 2c - \sqrt{2c(2c + \lambda c/k + p_j)} & \text{if } p_j \leq \tilde{p} \\
  2(v - c) - \lambda c/k - p_j & \text{if } p_j \geq \tilde{p}
\end{cases}
\]

where $\tilde{p} = v - 7c/4 - \lambda c/k + \sqrt{c^2 + 8cv}/4$.\(^{11}\) Firm $j$'s reaction function is symmetric.

To characterize the set of equilibria, we need to identify any fixed points of the two reaction functions. To do this, define the price $\bar{p}$ as the solution to $p = p + \lambda c/k + 2c - \sqrt{2c(2c + \lambda c/k + p)}$, so $\bar{p} = \lambda c/k + \lambda^2 c/2k^2$. Note that $\bar{p}$ is the fixed point of the reaction functions if the constraint $p_i \leq 2(v - c) - \lambda c/k - p_j$ is not binding. Finally, note that the restriction on $v$ implies $\tilde{p} < \bar{p}$ so the constraint $p_i \leq 2(v - c) - \lambda c/k - p_j$ is binding in any equilibrium. This analysis verifies the following proposition which is illustrated in Figure 3.1.

---

\(^{10}\) Adding a marginal cost $m$ of producing each unit, so the firm receives $p_i - m$ for each unit sold, slightly complicates the analysis, but does not change the results of the model.

\(^{11}\) There are two roots to the first-order condition for the firm’s profit maximization problem. The second-order condition verifies that the smaller root maximizes profit and defines the firm’s reaction function if $p_j \leq \tilde{p}$. The value $\tilde{p}$ is the value of $p_j$ at which the constraint $p_i \leq 2(v - c) - \lambda c/k - p_j$ becomes binding for firm $i$. The restriction on $v$ along with the fact that $p_j \leq v - c$ (otherwise buyers would not search firm $j$) ensures the constraint $p_j + \lambda c/k \geq p_i \geq p_j - \lambda c/k$ is never binding.
Figure 3.1: Equilibrium Price Dispersion
Proposition 1  There exist multiple subgame perfect equilibria with price dispersion. In particular, any price vector $p^* = (p^*_i, p^*_j)$ satisfying $p^*_i = 2(v - c) - \lambda c / k - p^*_j$ is a fixed point of the two firms’ reaction functions.

Surprisingly, despite the fact that buyers have perfect information about all prices, and firms compete by choosing price, the standard Bertrand equilibrium story does not hold. This outcome arises because the capacity constraint prevents firms from engaging purely in price competition in the sense described by Bertrand. A price reduction by firm $i$ makes $i$ more attractive to buyers because they can purchase from $i$ at a lower price. However, this benefit is only realized by a buyer if the item is actually in-stock, and the buyer must incur a search cost to determine this. In effect, the capacity constraint introduces an externality to the market. The externality arises because the increase in $\pi_i$ associated with a reduction in $p_i$ causes the probability $\alpha_i$ that firm $i$ has the item in-stock to decline, and this makes firm $i$ less attractive to buyers. Similarly, because $\pi_i + \pi_j = 1$, the increase in $p_i$ causes $\pi_j$ to fall which causes $\alpha_j$ to increase. As a result, a price reduction by firm $i$ does not generate the same buyer response, as measured by either the increase in $\pi^*_i$ or the reduction in $\pi^*_j$, that would occur in the absence of the externality. Along the segment of the reaction function where the constraint $p_i \leq 2(v - c) - \lambda c / k - p_j$ is binding, a price reduction by firm $i$ actually supports a higher price for firm $j$ in equilibrium. Rather than setting a lower price, firm $j$ responds to the increase in $\alpha_j$ that occurs as buyers increase $\pi_i$ (and reduce $\pi_j$) by setting a higher price. The extent to which the externality mitigates buyer response to a price reduction by firm $i$ is sufficiently large to override the incentive to cut price that arises in the typical Bertrand price competition story.

As demonstrated above, the theoretical model predicts that price dispersion can arise in equilibrium even if market prices are common knowledge. In an equilibrium with price dispersion firms that charge an above average price will be more
likely to have the item in-stock, while those that charge a below average price will be less likely to have the item in-stock. Finally, arguments similar to those made by Varian (2001) suggest that price dispersion can persist despite the fact that buyers know which firms charge the lowest prices because buyers are unable to costlessly observe which firms have the item in-stock, and firms charging lower prices are less likely to have the item in-stock in equilibrium. Our empirical analysis of the online market for college textbooks tests these theoretical predictions.

3.3 Data
To test empirically the relationship between prices and availability, we collected data for 21 textbooks for a period of 13 weeks, from July 30th to October 22nd, 2001, on 8 different websites. College textbooks are well suited to testing the theoretical model for several reasons. First, because a college student purchasing a specific textbook assigned by her instructor will not consider alternate textbooks as a substitute for the assigned textbook, issues of obfuscation and complex pricing strategies, like those analyzed by Ellison and Ellison (2000), in which a website posts a low price for one product (which it may not have in-stock) in order to attract buyers who can be offered a menu of close substitutes are minimized. Second, because instructors typically consider the quality and relevance of the content rather than price when selecting a textbook for a course, prices posted by individual websites reflect competition between sellers for a fixed pool of student buyers rather than an attempt to influence overall market demand for a particular textbook. Third, because each textbook is uniquely identified by an ISBN, we are able to compare homogeneous goods without concern that sellers are offering slightly differentiated products in an effort to price discriminate. Fourth, availability plays an important role in this market. In most cases a student ordering online needs the textbook for the current term. If the book is unlikely to be available at the website offering the lowest price, a student may be willing to pay a premium to shop at a website.
which is likely to have the books she needs in-stock. The possibility of this behavior is central to testing the predictions of the theoretical model. Finally, when shopping for textbooks consumers have a unit demand, not a downward sloping demand function. This is also consistent with the theoretical model.

Several of the websites in our sample offer used as well as new textbooks. One might argue that there are two distinct buyer types, those who buy only new books and those who buy used books. If this is the case, then used books are not a viable substitute for new books, and the availability of used textbooks would not complicate our empirical analysis. However, we were initially concerned that the availability of used books might affect pricing decisions for new books. To test this concern, we collected data on whether a used book alternative was listed with the new book information. Used book information typically appeared below, but on the same webpage as the new book. Including a dummy variable indicating the availability of a used book in the regressions discussed in Section 4 below had no effect on the regression results; pricing and availability of new textbooks online was unaffected by availability of used books. Adverse selection problems offer one potential explanation for this result. A student searching for a new book online might not consider a used book offered online as a substitute because the condition of the used textbook cannot be easily verified. Adverse selection is exacerbated by the fact that in most cases the used textbook was being sold by an individual, not the website – the website simply provided individuals a means of advertising a used book for sale.

The twenty-one books in our sample are top selling textbooks covering several fields of study. To identify websites for inclusion in our sample we consulted several search engines, including msn.com, yahoo.com and google.com to generate a list of the top websites selling textbooks online. After excluding from this list sites such as fatbrain.com which did not sell books directly but merely provided a link to
another website (Barnes and Noble’s website bn.com in the case of fatbrain), we were left with eight leading online textbook sellers. During the 13 week period, data were collected directly from each website, not from a shopbot. Although shopbots which compare prices offered by online textbook retailers do exist, these shopbots do not consistently provide information on other factors such as availability, shipping policies and costs, product guarantees, and return policies, that are likely to influence the buyer’s decision. For example, bigwords.com provides no availability information for either Amazon.com or BooksAMillion.com and provides availability information only on used books for several other retailers. This illustrates the fact that while price comparisons for a specific textbook can be effected relatively easily using a shopbot, additional search costs must be incurred to compare information about other characteristics, such as availability and shipping charges, across websites.

For each textbook we collected information on price, shipping charges, availability (whether the textbook was in-stock, and if it was in-stock, the time until it would be shipped), delivery time, return policy, and availability of used textbooks from each website. To ensure data were collected for the same edition of the same textbook across different websites we used the ISBN to identify the textbook. The search technology and customer interface were similar across different websites used in our data set. Each allowed the customer to search by ISBN, author, textbook name or keyword. The time required to obtain price information for a specific textbook also was similar across different websites. The search cost associated with obtaining availability information varied across websites. In some cases, such as

---

12 We also excluded efollet.com because it is as an online outlet for specific college bookstores which does not compete directly with other online retailers. For example, one must search by a course number for a specific college rather than by ISBN to obtain textbook information on efollet.

13 Studentmarket.com, which was an online textbook retailer during our sample period but now operates as a price comparison site, provides no textbook availability information.
Amazon.com or bn.com, this information was provided on the first or second screen for the book, while with others such as alltextbooks4less.com, studentmarket.com and textbooksource.net, availability information was not provided until after an order was actually placed. While two of these three websites made availability information more accessible during our sample period, textbooksource.net still does not confirm availability until one to two days after an order is placed. As a result of these practices, the cost of obtaining and comparing availability information was high relative to the cost of comparing price information.\textsuperscript{14}

The average price of the textbooks in our sample is $88.70 with an average minimum price of $80.79 and an average maximum price of $95.45. The highest price of a textbook during the sample period is $132.95 and the lowest price is $45.18. Different measures of price dispersion are used in the literature. Brynjolfsson and Smith (1999) use the range between the lowest and highest price of the good. Using the range measure we find that price dispersion exists for all textbooks in our sample. The average range is 18.95\%.\textsuperscript{15} The highest price dispersion within one period is 52.59\% and the lowest is 10.72\%. Baye, Morgan and Scholten (2001) define a narrower measure of price dispersion. They argue that the lowest two prices for a given good are the most important indicator of equilibrium price dispersion. If the lowest two prices are equal (and equal to marginal cost), then the two firms offering this price earn zero profit. If all buyers go to these two firms, then all firms offering the good at higher prices also earn zero profits. No firm has an incentive to change its price, and the good is only sold at the lowest price offered, so there is no price dispersion despite the fact that some firms offer different prices. To capture this idea of price dispersion they propose using a “gap” measure which is simply the

\textsuperscript{14} Two more recent entrants to the online textbook market, bookbundle.com and classbook.com also do not provide availability information until after an order has been placed.

\textsuperscript{15} This is about 2/3 of the price dispersion found by Brynjolfsson and Smith (1999).
difference between the two lowest prices. Using the gap measure in our sample, price
dispersion occurred in 238 of the 273 observations,\textsuperscript{16} while price dispersion did not
exist in 35 of the 273 observations. The average gap including the 35 observations
that did not report price dispersion is 4.45\%. Excluding observations with no price
dispersion, the average gap is 5.1\%\textsuperscript{17}. A third measure of price dispersion used
by Sorensen (2000) is the coefficient of variation, $\frac{\sigma}{\mu}$. The average coefficient of
variation in our sample (in percentage terms) is 6.13\%.

Charges for standard shipping ranged from free shipping to more than $7 per
textbook. At one extreme, ecampus.com offered free shipping starting on Septem-
ber 9th, 2001 through the end of our sample period. At the other extreme, var-
sitybooks.com calculated shipping charges as a percentage of the textbook’s price
($3.95 per order plus three percent of the total order amount). On the most ex-
pensive book this reached $7.91. Shipping time was similar across websites because
they use comparable shipping methods (mainly UPS ground).

Availability information was reported somewhat differently across websites.
Some provided specific information about whether or not the textbook was cur-
rently in-stock. Others provided an estimated number of days until the book would
be delivered to the customer’s location without specifying whether the book was
currently in-stock. The estimated delivery time was reported as a range. The re-
ported ranges could be placed in one of two categories, either two weeks or less, or
more than three weeks. None of the observations had an estimated delivery time in
the sixteen to twenty day range. We coded a book as available if either the book
was reported as in-stock or the estimated time until delivery was 15 days or less.

\textsuperscript{16} While our entire data set consists of 2184 observations (21 books at 8 websites
for 13 periods), it is reduced to 273 observations (one gap value per book for 21
books over 13 periods) when using the gap measure.

\textsuperscript{17} This is similar to the average gap of 5\% reported by Baye, Morgan and Scholten
and we coded a book as unavailable if it was either not in-stock or the time until delivery was more than three weeks. In our data set textbooks were unavailable in thirty percent of the observations.

Return policies were similar across websites. The standard return policy states that books can be returned within thirty days from the date of purchase with some websites offering a full refund while others deduct ten percent of the price of the textbook as a restocking fee. Table 3.1 lists the twenty-one textbooks and corresponding descriptive statistics. Table 3.2 lists the eight websites, their shipping costs and return policies, and summary statistics.

3.4 Testing the Hypothesis

In this section we conduct several different tests of the hypothesis implied by the theoretical model that the price a website charges for a particular textbook and the availability of that book at that website should be positively related. In particular, the theoretical model predicts that price dispersion can occur and persist in an equilibrium in which websites that charge above average prices also have above average availability, while those that charge below average prices have below average availability. Because a website’s pricing decision impacts buyer search strategies which, in turn, affect the availability of the item at that website (as well as at other websites), both price and availability are endogenous variables. Therefore, ordinary least squares estimators produced by regressing price on availability and the exogenous variables would be inconsistent. To obtain estimates which do provide a valid test of the relationship between price and availability, we begin

\[ \text{If one is comfortable assuming that availability is not correlated with the disturbance term in the price equation, then an OLS regression of price on availability and exogenous variables is valid. This OLS regression produces a coefficient of 1.06 on the availability dummy variable (which takes a value of 1 if the book is in-stock and 0 otherwise) which is significant at the 99.9 percent level. This coefficient implies that the premium for having the book in-stock is $1.06.} \]
by estimating the implicitly defined reduced form equations expressing price and availability as a function of the exogenous variables:

\[
\begin{align*}
\text{price} & = f(\text{listprice}, \text{website}, \text{shippingcharge}, \text{demand}) \\
\text{availability} & = f(\text{listprice}, \text{website}, \text{shippingcharge}, \text{demand}).
\end{align*}
\]

The variable \text{listprice} controls for textbook specific effects. Using dummy variables instead of \text{listprice} to identify the twenty-one textbooks did not change the results. However, because the use of textbook dummy variables places greater demands on the data by requiring that more coefficients be estimated, the coefficient estimates from regressions using \text{listprice} to control for textbook specific effects should be more precise. Both \text{listprice} and textbook dummy variables cannot be included in the same regression because \text{listprice} can be expressed as a linear combination of the textbook dummy variables. The \text{website} vector of seven dummy variables controls for website specific effects. Even though shipping methods are similar across websites (mainly UPS ground), shipping costs differ across websites. The \text{shippingcharge} variable controls for the impact of these differences. A \text{demand} dummy variable is added to control for a possible change in demand during the period from mid-August to mid-September of our sample. Demand may have differed during this period because most universities begin their academic years during this time. The \text{demand} dummy is assigned a value 1 if the observation is during the high-demand period and the value 0 if not. Rosenzweig and Wolpin (1980) make a strong case for utilizing estimates of the reduced form equations specified above. They demonstrate that simultaneous equation techniques yield no more information than that provided by the reduced form equation estimates. Unless there is an obvious choice for an instrument which is correlated with only one of the endogenous variables, the information obtained from simultaneous equations is redundant because the choice of instrument is otherwise arbitrary; any one of the exogenous variables could be omitted from one of the reduced form equations and used as an instrument. While
we include results from a two-stage least squares procedure below, following the arguments of Rosenzweig and Wolpin we begin our analysis by estimating the reduced form equations. Information about the relationship between price and availability is captured by the coefficients on the website dummy variables in these equations. The theoretical model predicts that if a particular website charges a high price, it should also have above average availability. This implies that the coefficients on the dummy variable for a particular website should have the same sign in both the estimated price and availability equations. The reduced form price equation was estimated using ordinary least squares. Because availability is binary, taking the value 1 if the textbook is in-stock, and the value 0 if it is not, the reduced form availability equation was estimated using a probit regression. Table 3.3 presents the results of the two reduced-form regressions. In both Table 3.3 and Table 3.4 standard errors are reported in parentheses below the coefficients, and three asterisks denote significance at the 99 percent level, two asterisks denote significance at the 95 percent level, and one asterisk denotes significance at the 90 percent level. The first column presents the results of the price regression. As expected, listprice has a positive and highly significant impact on price, with a coefficient of 0.93. This coefficient implies that on average textbooks are sold at 93 percent of the list price online. Shipping charges also have a positive and significant effect on prices. This somewhat counter-intuitive finding reflects the fact that websites such as amazon.com and bn.com which charge above average prices also have above average shipping charges. Combined with the results of the availability regression discussed below, it suggests that firms that provide above average availability receive a premium both in price and shipping charges. It also indicates that firms do not attempt to use a low price, high shipping charge strategy to confuse buyers and appear to offer a more attractive deal. The demand variable has a positive effect on prices; prices are higher during the period of high demand.
The website coefficients suggest that three websites, varsitybooks.com, bn.com and amazon.com, charge significantly higher prices than the omitted website, alltextbooks4less.com. Studentmarket.com and textbooksource.net have positive, but insignificant website coefficients, and ecampus.com has a negative but insignificant coefficient. This implies that the pricing strategies of these three websites do not significantly differ from alltextbooks4less.com. The negative and statistically significant website coefficient of -4.74 for booksamillion.com implies that they do charge lower prices than the other websites. However, booksamillion.com also charges $5 membership fee in order to receive their discounted price. Non-members are offered a substantially higher price. This was the only evidence of attempted obfuscation we observed in the online textbook market. Adding the $5 membership fee to booksamillion.com’s prices resulted in a positive but insignificant website coefficient for booksamillion.com.

The second column of Table 3.3 reports results for the availability regression. Listprice has a positive coefficient. The coefficients on shippingcharges and demand are both negative. The website coefficients imply that two websites, bn.com and amazon.com have significantly higher availability than alltextbooks4less.com. This is consistent with the theory because, as shown in the price regression, both of these websites also charge higher prices than alltextbooks4less.com. As was the case in the price regression, in the availability regression studentmarket.com and textbooksource.net have positive, but insignificant coefficients, and ecampus.com has a negative but insignificant coefficient. In addition, as expected after accounting for the $5 membership fee, the booksamillion.com website coefficient also is insignificant in the availability regression. This implies that availability at these four websites does not differ from availability at alltextbooks4less.com. Varsitybooks.com is the only website for which the website coefficient has opposite signs in the two regressions. The coefficient on varsitybooks.com is positive and significant in the price
regression and negative and significant in the availability regression.

The evidence from the reduced-form regressions is consistent with the theory for seven of the eight websites in our sample. Bn.com and amazon.com charge above average prices and have above average availability whereas studentmarket.com, textbooksource.net, ecampus.com, alltextbooks4less.com, and booksamillion.com charge below average prices and have below average availability. Varsitybooks.com charges high prices but has low availability. While varsitybooks.com’s strategy does not correspond with the equilibrium strategies predicted by the model, this divergence may be explained by their increased focus on providing textbook outsourcing services. As stated on their website, their services “include the partnership program, a program through which educational institutions ... outsource their new textbook-selling operations to Varsitybooks.com.” This focus is further evidenced by the fact that their website directs students to click on a regional map to find books for their particular school. Following these links only leads the student to particular textbooks if the school is a participating member of the partnership program. Based on this evidence, it is apparent that Varsitybooks.com is not competing directly with the other websites in our sample to sell textbooks to the general public. Estimates of the reduced form equations with varsitybooks.com omitted from the sample and the five dollar membership fee added to booksamillion.com prices are presented in the third and fourth columns of Table 3.3. These results reconfirm that amazon.com and bn.com have above average prices and availability while studentmarket.com, textbooksource.net, ecampus.com, alltextbooks4less.com, and booksamillion.com have below average prices and availability.

While the reduced form results support the predictions of the theoretical model, they do not provide a direct estimate of the marginal effect of price on availability. Simultaneous equation procedures are needed to obtain such an estimate. We adopted the reduced form equations approach above following Rosenzweig and
Wolpin’s arguments that unless a clear choice for an instrument exists, the choice of an instrument is arbitrary and the resulting two-stage least squares estimates add no additional information. However, based on the results of the reduced form estimates, an argument can be made that listprice serves as an instrument for price. Assuming the publisher’s choice of both the list price for the textbook and the wholesale price charged to online retailers are related, the impact of wholesale book costs on the retailers pricing decision will be directly related to list price. The strong connection between list price and the price actually posted by each website is clearly evident in the reduced form estimates of the price equation in Table 3.3; listprice is highly significant, and the $R^2$ in this equation is .95. However, because list price is set by the publisher, and not by online retailers, and because students’ purchasing decisions are unaffected by the list price, the list price will impact availability only indirectly through its effect on price. Therefore, an argument can be made for using listprice as an instrument for price. Results of the two-stage least squares estimation using listprice as an instrument are presented in Table 3.4. Not surprisingly, the two-stage least squares results confirm the information found in the reduced form estimates; the coefficient on fitted price is positive and significant, and the sign and significance of each website dummy variable corresponds to the reduced form results. The two-stage least squares estimates also enable us to compute the marginal effect of price on availability. The estimated marginal effect is a function of the textbook price and the values of the exogenous variables. Using the values from our data in this function, the marginal effect of price on availability ranged from .0076 to .0108 with an average of .009. This suggests that on average a one dollar increase in the price of a given textbook will increase the equilibrium probability the website has that book in-stock by nearly one percentage point.

Our final test of the hypothesis that websites which charge a higher price have higher availability in equilibrium is based on Baye, Morgan and Scholten’s
argument that the gap between the two lowest prices is the most appropriate indicator of price dispersion. To incorporate this concept of price dispersion we use an Anova test to compare availability at the website posting the lowest price for a particular book in a particular period with availability at the website posting the second-lowest price for that book in that period. The Anova test results are presented in Table 3.5. The average availability for the lowest priced textbook is 58.82%, and the average availability for the second lowest priced book is 78.57%. The Anova test rejects the hypothesis that availability for the lowest and second lowest priced textbooks is the same. Using only the two lowest prices to quantify price dispersion indicates that the website offering the lowest price has lower availability than the website offering the second lowest price. The results using the gap measure are of additional interest because, unlike the reduced form and two-stage least squares estimates above which are based on the entire sample, estimates based on the gap measure do not impose the assumption that an individual website applies the same pricing and availability strategy across all textbooks it sells.

3.5 Conclusion

Ready access to price information on the Internet has led to many claims that prices for a homogeneous good offered online will converge to the competitive price. However, a growing body of empirical research provides strong evidence of persistent price dispersion in online markets. This paper sheds some light on the discrepancy between the claims and the empirical evidence. Our results imply that observations made by Stigler (1961) in his seminal work on the economics of information still apply in the information age. Price dispersion continues to be an indication of ignorance in the market. However, because consumers shopping online have ready access to a great deal of price information provided by shopbots and price comparison websites, it is a stretch to argue that price dispersion in online markets results from ignorance about market prices. Rather, we argue that price
dispersion online is an indication of consumer ignorance about non-price parameters which are valued by the consumer. In particular, our results support the hypothesis that unobserved availability is sufficient to generate price dispersion in equilibrium even if prices are common knowledge.

The simple theoretical model presented in Section 3.2 illustrates how price dispersion can persist in equilibrium even when all prices are known. Because both price and availability affect consumer utility (searching a website at which the book is unavailable imposes additional search costs to find a website which has the book in-stock), unobserved availability can lead to equilibrium price dispersion in which some firms charge below average prices and have below average availability while other firms charge a premium for offering above average availability. Although consumers know which firms charge the lowest prices, they also know that, on average, the low price firms are less likely to have the item in-stock. Therefore, they continue to shop (according to a mixed strategy) at high price firms, and equilibrium price dispersion persists.

Data from the online market for college textbooks are used to test this hypothesis. Several different empirical tests conducted in Section 3.4 all support the predictions of the theoretical model. These tests indicate that amazon.com and bn.com charge above average prices and have above average availability, and alltextbooks4less.com, booksamillion.com, ecampus.com, textbooksource.net, and studentmarket.com charge below average prices and have below average availability. Our estimates imply that the marginal effect of a one dollar increase in the price of a particular textbook at a particular website is approximately a one percentage point increase in the availability of that textbook at that website in equilibrium. An Anova test based on the gap measure defined by Baye, Morgan and Scholten (2001) using only the two lowest prices for each book in each period to quantify price dispersion also supports the theoretical predictions; the website offering the
lowest price has lower availability than the website offering the second lowest price.

The claims that online markets should be perfectly competitive (or at least nearly so) depend crucially on the assumption that information is readily available online. While this is true for price information (because shopbots enable consumers to quickly compare prices posted by several retailers), it is not true in general for all parameters of interest to consumers. Because a buyer cannot walk into an online retailer and physically determine an item’s availability and quality, and then walk out with the item in hand, consumer information when purchasing online is incomplete in many dimensions. This paper presents evidence regarding the important role information about product availability and its effect on prices in online markets. The fact that inadequate information regarding other aspects of the transaction also inhibits perfectly competitive pricing in online markets is suggested by the significant coefficients on the website dummy variables. Smith and Brynjolfsson (2001) illustrate the strong influence of brand on consumer purchasing decisions and the prices firms charge online in much greater detail. Collectively these results suggest that further insights into why the predictions of perfectly competitive pricing online have not been realized may be found in the information captured by website specific effects.
### Table 3.1: Textbook descriptive statistics.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Calculus</td>
<td>1234567890</td>
<td>$59.95</td>
<td>$40.00</td>
<td>$39.99</td>
<td>$59.99</td>
<td>$3.56</td>
<td>0.12</td>
<td>10, 90</td>
<td>50-100</td>
<td>$59.95</td>
</tr>
<tr>
<td>Physics</td>
<td>9876543210</td>
<td>$49.99</td>
<td>$39.99</td>
<td>$49.99</td>
<td>$49.99</td>
<td>$1.75</td>
<td>0.21</td>
<td>50, 100</td>
<td>50-100</td>
<td>$49.99</td>
</tr>
<tr>
<td>Chemistry</td>
<td>1122334455</td>
<td>$59.99</td>
<td>$50.00</td>
<td>$69.99</td>
<td>$59.99</td>
<td>$3.45</td>
<td>0.08</td>
<td>70, 100</td>
<td>50-100</td>
<td>$59.99</td>
</tr>
<tr>
<td>Biology</td>
<td>5566778899</td>
<td>$59.99</td>
<td>$50.00</td>
<td>$60.00</td>
<td>$59.99</td>
<td>$2.56</td>
<td>0.12</td>
<td>60, 100</td>
<td>50-100</td>
<td>$59.99</td>
</tr>
<tr>
<td>English</td>
<td>9988776655</td>
<td>$69.99</td>
<td>$60.00</td>
<td>$70.00</td>
<td>$69.99</td>
<td>$3.45</td>
<td>0.12</td>
<td>70, 100</td>
<td>50-100</td>
<td>$69.99</td>
</tr>
<tr>
<td>History</td>
<td>5544332211</td>
<td>$59.99</td>
<td>$50.00</td>
<td>$60.00</td>
<td>$59.99</td>
<td>$2.56</td>
<td>0.12</td>
<td>60, 100</td>
<td>50-100</td>
<td>$59.99</td>
</tr>
<tr>
<td>Economics</td>
<td>1122334455</td>
<td>$59.99</td>
<td>$50.00</td>
<td>$60.00</td>
<td>$59.99</td>
<td>$2.56</td>
<td>0.12</td>
<td>60, 100</td>
<td>50-100</td>
<td>$59.99</td>
</tr>
</tbody>
</table>

*Note: Price Percentiles Available range from 10 to 100, with 50 representing the median.*
### Table 3.2: Websites descriptive statistics.

<table>
<thead>
<tr>
<th>Website</th>
<th>Average Percent Availability</th>
<th>Average Price</th>
<th>Shipping Costs</th>
<th>Return Policies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecampus.com</td>
<td>71%</td>
<td>$85.97</td>
<td>UPS Standard (1-5 Business Days): $2.99 Flat Rate + $0.99 Each Item</td>
<td>Items purchased from ecampus.com may be returned (excluding shipping charges) for a refund up to 30 days.</td>
</tr>
<tr>
<td>Alttextbooks4less.com</td>
<td>69%</td>
<td>$86.98</td>
<td>Standard Shipping: $3.49 per order + $0.99 per item</td>
<td>Within 30 days of confirmation of your order, you may return any item, for any reason, and receive a refund in the amount of the price paid for the merchandise less a 10% restocking fee.</td>
</tr>
<tr>
<td>Varsitybooks.com</td>
<td>48%</td>
<td>$94.97</td>
<td>UPS Ground (2-3 business days): $3.95 + 3% of your total order amount</td>
<td>Books can be returned up to 45 days from initial purchase date. If the book is in its new condition than there will be a full refund.</td>
</tr>
<tr>
<td>Bn.com</td>
<td>92%</td>
<td>$91.97</td>
<td>Standard Ground (3 to 6 business days): $2.49 per order + $0.99 per item</td>
<td>If you are not completely satisfied with your purchase, simply return it to us within 30 days for a full refund.</td>
</tr>
<tr>
<td>Amazon.com</td>
<td>81%</td>
<td>$92.87</td>
<td>Standard Shipping (3-7 Business Days): $3.00 per shipment + $0.99 per item</td>
<td>Within 30 days of your order, you may return any item, for any reason, and receive a full refund.</td>
</tr>
<tr>
<td>Textbooksource.com</td>
<td>69%</td>
<td>$97.48</td>
<td>Standard Shipping: $3.49 per order + $0.99 per item</td>
<td>Within 30 days of confirmation of your order, you may return any item, for any reason, and receive a refund in the amount of the price paid for the merchandise less a 10% restocking fee.</td>
</tr>
<tr>
<td>Studentmarket.net</td>
<td>72%</td>
<td>$97.36</td>
<td>Standard Shipping: $3.49 per order + $0.99 per item</td>
<td>Within 30 days of confirmation of your order, you may return any item, for any reason, and receive a refund in the amount of the price paid for the merchandise less a 10% restocking fee.</td>
</tr>
<tr>
<td>Booksamillion.com</td>
<td>69%</td>
<td>$92.70</td>
<td>Standard Ground (3 to 7 business days): $2.49 per order + $0.99 per item</td>
<td>Return unwanted merchandise within 30 days after you receive your order and you will receive a refund. Shipping and handling charges are not refundable.</td>
</tr>
</tbody>
</table>
Table 3.3: Reduced form regressions.

<table>
<thead>
<tr>
<th></th>
<th>Price</th>
<th>Availability</th>
<th>Price</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>-2.234</td>
<td>0.504</td>
<td>-1.669</td>
<td>0.199</td>
</tr>
<tr>
<td></td>
<td>(0.616)**</td>
<td>(0.180)**</td>
<td>(0.694)**</td>
<td>(0.194)**</td>
</tr>
<tr>
<td>List Price</td>
<td>0.934</td>
<td>0.002</td>
<td>0.930</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.005)**</td>
<td>(0.0014)*</td>
<td>(0.0053)**</td>
<td>(0.0015)**</td>
</tr>
<tr>
<td>Ecampus</td>
<td>-0.612</td>
<td>-0.01</td>
<td>-0.7</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.41)</td>
<td>(0.118)</td>
<td>(0.434)</td>
<td>(0.119)</td>
</tr>
<tr>
<td>Varsitybooks</td>
<td>7.453</td>
<td>-0.411</td>
<td>5.170</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td>(0.427)**</td>
<td>(0.121)**</td>
<td>(0.418)**</td>
<td>(0.14)****</td>
</tr>
<tr>
<td>Bn</td>
<td>5.221</td>
<td>0.884</td>
<td>5.978</td>
<td>0.378</td>
</tr>
<tr>
<td></td>
<td>(0.396)**</td>
<td>(0.139)**</td>
<td>(0.412)**</td>
<td>(0.119)****</td>
</tr>
<tr>
<td>Amazon</td>
<td>6.003</td>
<td>0.371</td>
<td>5.978</td>
<td>0.378</td>
</tr>
<tr>
<td></td>
<td>(0.39)**</td>
<td>(0.119)**</td>
<td>(0.412)**</td>
<td>(0.119)****</td>
</tr>
<tr>
<td>Textbooksource</td>
<td>0.501</td>
<td>0.010</td>
<td>0.501</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>(0.388)</td>
<td>(0.1124)</td>
<td>(0.410)</td>
<td>(0.1126)</td>
</tr>
<tr>
<td>Studentmarket</td>
<td>0.384</td>
<td>0.084</td>
<td>0.384</td>
<td>0.083</td>
</tr>
<tr>
<td></td>
<td>(0.3881)</td>
<td>(0.113)</td>
<td>(0.41)</td>
<td>(0.113)</td>
</tr>
<tr>
<td>Booksamillion</td>
<td>-4.738</td>
<td>-0.045</td>
<td>0.211</td>
<td>-0.030</td>
</tr>
<tr>
<td></td>
<td>(0.396)**</td>
<td>(0.1146)</td>
<td>(0.418)</td>
<td>(0.115)</td>
</tr>
<tr>
<td>Shippingcharge</td>
<td>0.231</td>
<td>-0.054</td>
<td>0.180</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>(0.077)**</td>
<td>(0.022)**</td>
<td>(0.082)**</td>
<td>(0.022)**</td>
</tr>
<tr>
<td>Demand</td>
<td>0.399</td>
<td>-0.002</td>
<td>0.461</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.210)*</td>
<td>(0.06)</td>
<td>(0.238)*</td>
<td>(0.069)</td>
</tr>
<tr>
<td>R^2</td>
<td>0.95</td>
<td>0.063</td>
<td>0.94</td>
<td>0.04</td>
</tr>
</tbody>
</table>
Table 3.4: Two-stage least squares regression.

<table>
<thead>
<tr>
<th></th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.676</td>
</tr>
<tr>
<td></td>
<td>(0.129)***</td>
</tr>
<tr>
<td>Estimated Prices</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.006)***</td>
</tr>
<tr>
<td>Ecampus</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
</tr>
<tr>
<td>Bn</td>
<td>0.902</td>
</tr>
<tr>
<td></td>
<td>(0.139)***</td>
</tr>
<tr>
<td>Amazon</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>(0.119)***</td>
</tr>
<tr>
<td>Tetxbooksource</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>(0.113)</td>
</tr>
<tr>
<td>Studentmarket</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>(0.114)</td>
</tr>
<tr>
<td>Booksamillion</td>
<td>-0.024</td>
</tr>
<tr>
<td></td>
<td>(0.115)</td>
</tr>
<tr>
<td>Shippingcharges</td>
<td>-0.037</td>
</tr>
<tr>
<td></td>
<td>(0.022)*</td>
</tr>
<tr>
<td>Demand</td>
<td>-0.032</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.049</td>
</tr>
</tbody>
</table>
Table 3.5: Anova test using the Gap measure.

<table>
<thead>
<tr>
<th></th>
<th>Lowest Price</th>
<th>Second-lowest Price</th>
<th>All</th>
<th>Anova F-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>58.82%</td>
<td>78.57%</td>
<td>68.70%</td>
<td>22.51077</td>
</tr>
<tr>
<td>Availability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

47
Chapter 4

AN EQUILIBRIUM MODEL OF ONLINE PRICE COMPARISON SITES WHEN FIRMS HAVE DIFFERENT MARKET SHARES

4.1 Introduction

In the late nineties, with the rapid expansion of the Internet, many economists predicted that prices in online markets would not be dispersed. Homogenous goods would be sold at a common price. The motivation behind this prediction is that search costs on the Internet are negligible, and consumers can easily obtain full price information. Since most price dispersion models rely on consumers’ imperfect price information to obtain a distribution of prices, bringing search costs to nearly zero was supposed to converge prices to a unique price.

Furthermore, the increased popularity of the Internet and the decrease in the cost of technology necessary to gather price information opened the way to a profitable new line of business: Information provision by price comparison sites. Price comparison sites, such as Shopper.com and Mysimon.com acting like gatekeepers, are information intermediaries. Having developed software that can search for the price of a product across several e-retailers and present these prices to online consumers almost instantaneously, these gatekeepers enable consumers to compare quickly prices charged by several online retailers before choosing from whom to purchase the good.
Baye and Morgan (2001) present a theoretical model in which a gatekeeper allows firms, for an advertising fee, and consumers, for a subscription fee, to expand their markets which otherwise would be limited to a local market. This model by Baye and Morgan is the first, and to the best of my knowledge, the only model of an information provider decision making. The most important finding of their model is that in equilibrium firms use mixed pricing strategies and prices are dispersed on the gatekeeper site.

In Baye and Morgan’s model firms are homogeneous, with identical local market shares. After observing online markets one can clearly see that not all websites have the same market share. Electronic markets are usually characterized by few firms with familiar names which capture most of the customers in the market and many smaller websites unknown to most consumers. Given that consumers are more likely to shop from few firms in the market, it is important to understand the impact of heterogeneous market shares on retailer advertising and pricing behavior in equilibrium in a market with an information gatekeeper. This paper extends the Baye and Morgan model to analyze the equilibrium in a market with a gatekeeper in which firms have different market shares. In the Baye and Morgan model, even though the gatekeeper sets its subscription fees at a level that induces full consumer participation in the market for information, it sets advertising fees in such a way to limit firm participation on its website, and to avoid bringing prices to marginal cost. Introducing the gatekeeper in the market lowers prices, but not down to the social optimum. An important question is whether the gatekeeper maintains such a level of power when the market share of some firms becomes very high. It is also important to know the effect of different market shares on a firm participation in the market for information, and on its pricing decision. Moreover, examining the effect of the gatekeeper on price dispersion under realistic assumptions, particularly when firms are assumed to have different market shares, is very important since the
role price comparison sites play in online markets is growing.

The analysis in this chapter demonstrates that an equilibrium exists in which prices are dispersed on the gatekeeper’s site. Firms with different market shares advertise on the gatekeeper’s site with different probabilities. The firm with the big market share advertises with a lower probability, but prices more competitively when it advertises on the gatekeeper’s site. The firm with a small market share advertises more frequently, but advertises the monopoly price with a positive probability. These results can help explain the absence of several well known retailers, such as Circuit City from the data set of more than 4 million observations of electronic products collected by Baye, Morgan and Scholten (2001) from Shopper.com.

The attractiveness of the gatekeeper’s service to consumers is that it dramatically decreases search costs. Consumers need only visit one website in order to obtain extensive price information from competing retailers. Moreover firms also benefit from the gatekeeper’s services. When shopping on the Internet for a specific good, most consumers are only familiar with the name of a few websites selling that good, if any. Most websites are relatively unknown to consumers. They may get discovered only by chance. Introducing a gatekeeper in the market allows online firms to increase their demand. Since both firms and consumers potentially can benefit from the gatekeeper’s services, both would be willing to pay a fee to the gatekeeper. Consumers could pay a subscription fee to have access to the information gathered by the gatekeeper, and firms could pay an advertising fee to have their prices appear on the gatekeeper’s site. The gatekeeper makes profits from the difference between the fees collected and the cost of the search engine software. Firms would be willing to pay an advertising fee to the gatekeeper as long as there are consumers who subscribe to the gatekeeper’s site. Consumers would be willing to pay a subscription fee as long as they can save on search cost. On the other hand if a market has no price dispersion, consumers will be unwilling to pay a fee to have access to a list where
prices are the same at every retailer. Paying the subscription fee would not reduce the expected price of the good they want to purchase. Knowing this, consumers do not subscribe to the gatekeeper’s site if prices are not dispersed. In turn, firms would be unwilling to pay an advertising fee bringing them no potential gain in demand. In this case the gatekeeper cannot remain in the market with no source of revenue. Thus it is necessary to have price dispersion in order for the gatekeeper to exist the market. This can be hard to achieve if all consumers have full information. By providing price information to consumers, the gatekeeper creates a perfect competition environment for firms which could force them to charge a unique price (possibly the competitive price). The gatekeeper needs to insure that this does not happen, possibly by excluding some firms from advertising on its site, or similarly by excluding some consumers from subscribing to its site.

The idea of information providers is not unique to the Internet or new to the economic literature. Stigler (1961) argues that if the size of a market where there is price dispersion increases, it becomes profitable for a third-party to collect information about the distribution of prices and sell it to the buyers. In the sixties, seventies and eighties these information providers took the form of newspapers, trade journal or brokers. Moreover Stigler predicted that there will be a "tendency toward monopoly in the provision of information". Later these information providers appeared in the background of several price dispersion articles. Salop and Stiglitz (1977) assume that for a fee consumers can obtain full price information, suggesting that there exists an information provider in the market. In their model, consumers are heterogeneous with respect to the fee (determined exogenously) they must pay in order to obtain full price information. If a consumer decides to pay the fee, then she buys the good at the lowest price, otherwise she shops randomly. Varian (1980) assumes that some consumers have access to a list of prices across different retailers. These informed consumers buy at the lowest advertised price. Other consumers
are uninformed and must shop randomly. The type of a consumer is exogenously determined in Varian’s model as well. One can think of the informed consumers in this model as consumers with access to a price comparison site giving them access to the list of prices.

In Section 4.2 we present the model. In Section 4.3 we discuss the comparative statics. We analyse the case with full consumer participation in section 4.4. Section 4.5 concludes.

4.2 The Model

The model in this paper builds on Baye and Morgan (2001). In their model there is a continuum of consumers, each of whom has a demand function \( q(p) \) that is continuous and non-increasing in price. Consumer surplus at price \( p \) is denoted \( S(p) \). The measure of consumers is normalized to be unity, it is assumed that they are entirely divided among \( n \) local markets, and that each local market is served by a single local firm. Local markets are completely segmented. Furthermore it is assumed that the \( n \) local markets are homogeneous in size. Firms set linear prices and sell identical products at constant marginal cost \( (c \geq 0) \). Firms have a monopoly power in their local market, and their monopoly price is denoted by \( r \). It costs a consumer \( \rho \) to visit a local store, and \( \rho \) is sufficiently small that \( S(r) > \rho \).

A monopoly information gatekeeper that controls a central clearinghouse for information charges a fee \( \phi \) to firms for advertising their prices and \( \kappa \) to consumers for subscribing to its price list. A fraction \( \mu \in (0, 1] \) of the consumers subscribes to the price comparison site and a fraction \( (1 - \mu) \) does not. By advertising on the gatekeeper website a firm has access to subscribers from other firms’ local markets. Similarly by subscribing to the gatekeeper price list, consumers can buy from any advertising firm (not only their local firm). Firms cannot discriminate between subscribers and non-subscribers, therefore they charge the same price to all consumers. After observing the advertising fee \( \phi \) and the fraction of consumers subscribing to
the gatekeeper’s site $\mu$, firms make advertising and pricing decisions. Consumers decide whether to subscribe or not, and then shop.

The main results of their model are as follows. First, if the gatekeeper charges a strictly positive advertising fee that is not too excessive, it can preclude full firm participation on its website. Firms use mixed advertising strategies, and only a fraction of firms advertise their prices on the gatekeeper’s site. Similarly, advertising firms use mixed pricing strategies that result in price dispersion on the gatekeeper’s site. A non-advertising firm charges the monopoly price. Therefore advertised prices are lower than non-advertised prices and consumers have an incentive to subscribe to the gatekeeper site. The subscription fee $\kappa$ set by the gatekeeper determines the fraction $\mu$ of subscribers. The gatekeeper sets $\phi$ and $\kappa$ in such a way to insure that price dispersion exists in the market. Thus it may be optimal for the gatekeeper to induce full consumer participation on its site by setting low subscription fees but to preclude full firm participation by setting a relatively high advertising fee. In equilibrium prices are higher than the socially optimal price – price exceeds marginal cost.

In our model the setting is similar to Baye and Morgan’s with the exception that firms have different market shares. Consumers are not evenly divided among the two firms. Firm $i$’s share of consumers is denoted by $s_i$. We also suppose that $\mu$ and $\kappa$ are exogenous. Assuming $\mu$ and $\kappa$ to be exogenous is without any loss of reality as access to most price comparison sites is free to consumers and nonetheless not all consumers subscribe to such services. The gatekeeper sets the advertising fee $\phi$, to maximize its profits, firms observe the advertising fee $\phi$ and the fraction of subscribers $\mu$ and then select a price $p_i \in [c, r]$ to charge and make an advertising decision $a_i \in \{A, N\}$, where $A$ indicates that the firm advertises and $N$ indicates that the firm does not advertise, given their market share $s_i$. Then consumers shop and (possibly) purchase the product. To allow for the possibility of
mixed strategies let $\alpha_i$ denote the probability that firm $i$ chooses to advertise, and let $F_i(p) = \Pr\{p_i \leq p\}$.

We consider the case when all consumers are divided among two firms, firm $i$ and firm $j$. Firm $i$ has a market share $s_i$ and firm $j$ has a market share $s_j$.

In any shopping subgame, non-subscribing consumers visit and purchase from their local firm at the local price given that their surplus $S(r)$ at the monopoly price is higher than the cost of visiting their local firm $\varepsilon$. The marginal cost for a subscriber from visiting the gatekeeper’s site is zero, whereas the cost of visiting her local firm is $\varepsilon > 0$, therefore a subscriber always visits the gatekeeper’s site first. Knowing that the firms that don’t advertise charge higher prices than the firms that advertise a subscriber buys at the lowest advertised price on the gatekeeper’s site if there are advertised prices on the site. If no firm is advertising, a subscriber visits and purchases from her local firm since her surplus from purchasing the good even at the monopoly price is higher than the cost $\varepsilon$ of visiting the local firm.

4.2.1 The Firms’ Problem

A seller $i$ that does not advertise on the gatekeeper site attracts only customers from its local market. These customers consist entirely of the $(1 - \mu)s_i$ non-subscribers and the subscribers who did not find prices at the gatekeeper site. Because firm $j$ advertises with probability $\alpha_j$, the expected number of subscribers who purchase from firm $i$ when it does not advertise is $(1 - \alpha_j)\mu s_i$. Both types of buyers will pay up to the monopoly price. Therefore it is optimal for a non-advertising firm to charge the monopoly price $r$, and the expected profits $E\pi_i(N, r)$ of a non-advertising firm are:

$$E\pi_i(N, r) = (1 - \alpha_j)\mu s_i \pi_r + (1 - \mu)s_i \pi_r,$$

(4.1)

where $\pi_r = (r - c)q(r)$ is the firm’s monopoly profit.
If firm \( i \) advertises on the gatekeeper’s site, sells to its \((1-\mu)s_i\) non-subscribers and to all subscribers, regardless of their local market in the case it advertises the lowest price on the gatekeeper’s site, otherwise it makes no sales on the gatekeeper’s site. Assuming firm \( i \) advertises a price \( p \) the probability that firm \( j \) advertises a lower price than firm \( i \) is \( \alpha_j F_j(p) \). Therefore the probability that all subscribers purchase from firm \( i \) is \( 1-\alpha_j F_j(p) \). We can now express the expected profit \( E\pi_i(A, p) \) of an advertising firm as:

\[
E\pi_i(A, p) = \mu \pi_p (1 - \alpha_j F_j(p)) + (1 - \mu) s_i \pi_p - \phi,
\]

where \( \pi_p = (p - c)q(p) \) is the firm’s profit from making a sale at the price \( p \).

Let \( \bar{p} \) be the upper support of \( F_i(p) \). Given that firm \( j \) charges a price lower than \( \bar{p} \) (when it advertises) with probability one, it follows that firm \( i \) makes a sale when advertising a price \( \bar{p} \) only if firm \( j \) is not advertising. In the case firm \( j \) does not advertise, it is optimal for firm \( i \) to charge the monopoly price. Therefore the upper support of \( F_i(p) \) is the monopoly price \( r \).

Since \( \pi_p \) is increasing up to the monopoly price, it follows that the upper support of \( F_i(p) \) is the monopoly price. The only time firm \( i \) can sell on the gatekeeper site when it charges the upper support is when the other firm is not advertising. If the upper support is below the monopoly price and firm \( j \) is not advertising, then firm \( i \) can increase its profits if it increases its price to the monopoly price \( r \). Because firm \( j \) does not advertise with probability \( \alpha_j = 1 \) as shown in lemma 1 below, it follows that it is optimal for a firm to set the upper support of its distribution at the monopoly price. Furthermore in order for \( F \) to be part of a Nash equilibrium, firm \( i \)’s expected profits must be constant for all prices in the support of \( F \).

**Lemma 1** For an advertising fee \( 0 < \phi < \mu \pi_r (1 - s_j) \), there exists no pure-strategy advertising equilibrium in which firms advertise their prices with probability 0 or 1.
Proof. Suppose firm $j$ advertises it prices with probability $\alpha_j = 1$ and its price is chosen from the cumulative distribution function $F(p)$ with support $[p, r]$ with $p \leq r$. Firm $i$’s expected profits from not advertising are: $E\pi_i(N, r) = (1 - \mu) s_i \pi_r$. Firm $i$’s expected profits from advertising the monopoly price $r$ (which are equal to profits from advertising any price $p$ in the support of $F$) are: $E\pi_i(A, r) = (1 - \mu) s_i \pi_r - \phi$. Given $\phi > 0$, firm $i$’s best response to firm $j$ advertising with probability 1 is not to advertise. Now given that firm $i$ advertises with probability 0, firm $j$’s expected profits from not-advertising are: $E\pi_j(N, r) = s_j \pi_r$. Firm $j$’s expected profits from advertising a price $p$ are: $E\pi_j(A, p) = \mu \pi_p + (1 - \mu) s_j \pi_p - \phi$, and from advertising the monopoly price $r$ are: $E\pi_j(A, r) = \mu \pi_r + (1 - \mu) s_j \pi_r - \phi$. Given that $\pi_p \leq \pi_r$ for any $p \leq r$, it follows that $E\pi_j(A, p) \leq E\pi_j(A, r)$, and it is optimal for firm $j$ to advertise the monopoly price $r$. Furthermore for $\phi < \mu \pi_r (1 - s_j)$ it is optimal for firm $j$ to advertise. Therefore it is optimal for firm $j$ to always advertise the monopoly price if firm $i$ does not advertise. But if firm $j$ advertises the monopoly price with probability 1, then there exists an $\epsilon > 0$ such that firm $i$’s expected profits from advertising and charging a price $r - \epsilon$ is $E\pi_i(A, r - \epsilon) = \mu \pi_{r - \epsilon} + (1 - \mu) s_i \pi_{r - \epsilon} - \phi > (1 - \mu) s_i \pi_r = E\pi_i(N, r)$, where the inequality follows from the fact that $\pi_p$ is continuous in $p$. Therefore it becomes optimal for firm $i$ to advertise.

In addition to the fact that firms cannot adopt pure advertising strategies in equilibrium, it is obvious that a firm $i$ cannot follow a pure pricing strategy when advertising because the competing firm $j$ would simply undercut firm $i$’s price. Therefore we now consider the case of mixed-strategy equilibrium. First we solve for the probability firm $j$ advertises. In equilibrium the expected profits from advertising any price in the support of $F_i$, including the monopoly price $r$ (which is the upper support of $F_i$) must be equal to the expected profits from not-advertising and charging the monopoly price. Equating the expressions for the expected profits...
in equations (4.1) and (4.2) and setting the advertised price equal to the monopoly price \( r \), we get the probability \( \alpha_j \) that firm \( j \) advertises (note that \( F_j(r) = 1 \)):

\[
\alpha_j = 1 - \frac{\phi}{\mu \pi_r s_j} \tag{4.3}
\]

As expected the probability a firm advertises increases in the fraction of subscribers \( \mu \) when the advertising fee \( \phi \) is fixed. A higher \( \mu \) increases the expected profits from advertising, and therefore it increases the probability a firm advertises. Similarly higher profits increase the probability firms advertise, this can be explained as the result of lower relative advertising fees \( \phi \). Advertising is important for a firm not only to increase its demand but also to protect its share of subscribers. This can be seen in the relationship between \( \alpha_j \) and the market share \( s_j \). A higher market share increases the probability a firm advertises because the higher the market share the bigger the loss the firm incurs when it loses the demand from its subscribers.

We now need to solve for the cumulative distribution function of advertised prices. Equating (4.2) and (4.2) at the monopoly price allows us to solve for \( F_j \):

\[
F_j(p) = \frac{1}{\alpha_j} \left( 1 - \frac{(1 - \alpha_j)\mu \pi_r + (1 - \mu)s_i(\pi_r - \pi_p)}{\mu \pi_p} \right) \tag{4.4}
\]

Substituting for \( \alpha_j \) from (4.3) yields:

\[
F_j(p) = -\pi_r s_i (s_i + \mu - s_i \mu - 1) + \frac{\pi_p (\mu - 2s_i \mu + s_i - s_i^2 + s_i^2 \mu - \phi)}{\pi_p (-\mu \pi_r + \mu s_i \pi_r + \phi)} \tag{4.5}
\]

Equation (4.5) defines a continuous mixed strategy pricing distribution. It is necessary to calculate the lower support of \( F_j \), \( p_j \). We set \( F_j(p_j) = 0 \) and we get:

\[
p_j = \pi^{-1} \left( \frac{\phi + s_i(1 - \mu)}{\mu(1 - s_i)(\mu + s_i(1 - \mu))} \right) \tag{4.6}
\]

Equation (4.6) shows that the firm with the higher market share has a lower lower support. However, in equilibrium the lower support adopted by the two firms cannot differ because the firm with the lowest lower support can increase its expected profits by increasing its lower support to its competitor’s lower support minus \( \epsilon \) (where \( \epsilon > 0 \).
is an arbitrary small number). Because equations (4.5) and (4.6) implicitly assumed that neither firm has a mass point in its pricing distribution we have verified the following lemma.

**Lemma 2** There is no equilibrium in which both firms $i$ and $j$ price according to atomless distributions $F_i$ and $F_j$.

Since $F$ cannot be atomless, we now consider distributions with mass points.

**Lemma 3** In equilibrium the only price at which either firms’ mixed strategy price distribution $F(p)$ can have a mass point is the monopoly price $r$.

**Proof.** Suppose firm $j$ has a mass point at price $p = p^m$. Consider the expected profits of firm $i$ when advertising a price $p = p^m + \epsilon$, $E\pi_i(A, p^m + \epsilon) = \mu \pi_{p^m + \epsilon}(1 - \alpha_j F_j(p^m + \epsilon)) + (1 - \mu) s_i \pi_{p^m + \epsilon} - \phi$. Similarly consider the expected profits of firm $i$ when advertising a price $p = p^m - \epsilon$, $E\pi_i(A, p^m - \epsilon) = \mu \pi_{p^m - \epsilon}(1 - \alpha_j F_j(p^m - \epsilon)) + (1 - \mu) s_i \pi_{p^m - \epsilon} - \phi$. The expected profits of firm $i$ must be the same at every price in the support of the c.d.f., which requires that $E\pi_i(A, p^m + \epsilon) = E\pi_i(A, p^m - \epsilon)$. Taking $\lim_{\epsilon \to 0}\{E\pi_i(A, p^m + \epsilon) - E\pi_i(A, p^m - \epsilon)\}$ we get $\mu \pi_{p^m} \alpha_j (F_j(p^m - \epsilon) - F_j(p^m + \epsilon))$. Given that firm $j$ has a mass point at $p^m$, $\lim_{\epsilon \to 0}\{F_j(p^m - \epsilon) - F_j(p^m + \epsilon)\} < 0$, and there exists an $\epsilon > 0$, such that $E\pi_i(A, p^m + \epsilon) < E\pi_i(A, p^m - \epsilon)$, which implies that either $p^m - \epsilon$ or $p^m + \epsilon$ cannot be part of the support of the c.d.f. It follows that $p^m$ must be at the limits of the distribution. Charging the marginal cost is not profitable given the fixed cost $\phi$ imposed by the gatekeeper. Therefore the only price at which a firm can have a mass point is the monopoly price $r$. ■

**Lemma 4** In equilibrium, it is not possible for both firms to have mass points.

**Proof.** Given that the only price at which a firm can have a mass point is the monopoly price, suppose both firms have mass points at $p = r$. Given firm $i$'s
behavior, there exists an $\epsilon > 0$ such that firm $j$ can increase its profits by moving its mass point from $r$ to $r - \epsilon$. By putting the mass probability at $r - \epsilon$ firm $j$ increases its chances of capturing all the subscribers, while keeping its profits per unit virtually the same as at the monopoly price. Since as shown in lemma 3 the only price at which a firm can have a mass point is $r$, it follows that only one firm can have a mass point in its distribution.

We now proceed to find an equilibrium in which firm $i$ has a mass point at the monopoly price $r$ and firm $j$ prices according to an atomless distribution $F_j$. Firm $i$’s expected profits are given by equations (4.1) and (4.2). Equations (4.3), (4.5) and (4.6) respectively give the probability firm $j$ advertises, its distribution function $F_j$ and the lower support of its distribution. Firm $j$’s expected profit from not advertising is:

$$E\pi_j(N, r) = (1 - \alpha_i)\mu s_j \pi_r + (1 - \mu) s_j \pi_r$$  \hfill (4.7)

Firm $j$’s expected profit from advertising a price $p \neq r$ is:

$$E\pi_j(A, p) = \mu \pi_p (1 - \alpha_i F_i(p)) + (1 - \mu) s_j \pi_p - \phi$$  \hfill (4.8)

Firm $j$’s expected profit from advertising the monopoly price $r$, given that it gets half of the consumers if firm $i$ is also advertising the monopoly price, is:

$$E\pi_j(A, r) = \mu \pi_r (1 - \alpha_i (1 - 0.5 F_{ri})) + (1 - \mu) s_j \pi_r - \phi$$  \hfill (4.9)

where $F_{ri}$ is the value of the mass point.

We now solve for the probability firm $i$ advertises ($\alpha_i$), its distribution function ($F_i(p)$) and the mass point ($F_{ri}$). First we solve for $\alpha_i$ by setting (4.7) = (4.9):

$$\alpha_i = \frac{(1 - s_j)\mu \pi_r - \phi}{\mu \pi_r (1 - s_j - F_{ri}/2)}$$  \hfill (4.10)

We then solve for the distribution function $F_i$ by setting (4.8) = (4.9):

$$F_i(p) = \frac{1}{\alpha_i} \left(1 - \frac{\pi_r(\mu(1 - \alpha_i + \alpha_i F_{ri}/2) + (1 - \mu) s_j - (1 - \mu) s_j \pi_p)}{\mu \pi_p}\right)$$  \hfill (4.11)
Setting \( F_i(p) = 0 \), we get the lower support of firm \( i \)'s distribution of prices, that we denote \( p_i \):

\[
p_i = \pi^{-1}(\pi r (1 - \alpha_i + \alpha_i F_{ri}/2) \mu + s_j(1 - \mu)) / (\mu + s_j(1 - \mu)) \tag{4.12}
\]

Finally we solve for the mass point by setting the lower bounds of both distributions \( p_i \) and \( p_j \) equal:

\[
F_{ri} = 2(1 - \alpha_j \mu + s_j(1 - \mu)) / (\alpha_i \mu + s_i(1 - \mu)) \tag{4.13}
\]

**Proposition 2** In a market with two firms, only the firm with the small market share can have the mass point in its distribution.

**Proof.** In equilibrium, the mass point must be positive and smaller than 1, \( 0 < F_{ri} < 1 \). Substituting for \( \alpha_i \) and \( \alpha_j \) in equation (4.13) we get:

\[
F_{ri}/2 = (\mu \pi r (1 - s_j) - \phi)(\mu + s_j(1 - \mu)) s_j - (\mu \pi r (1 - s_i) - \phi)(\mu + s_i(1 - \mu))(1 - s_j) / (\mu \pi r (1 - s_j) - \phi)(\mu + s_j(1 - \mu)) s_j - (\mu \pi r (1 - s_i) - \phi)(\mu + s_i(1 - \mu)) (4.14)
\]

Which simplifies to:

\[
F_{ri}/2 = 1 + s_j(\mu \pi r (1 - s_i) - \phi)(\mu + s_j(1 - \mu)) / (\mu \pi r (1 - s_j) - \phi)(\mu + s_j(1 - \mu)) s_j - (\mu \pi r (1 - s_i) - \phi)(\mu + s_i(1 - \mu)) (4.15)
\]

The numerator \( s_j(\mu \pi r (1 - s_i) - \phi)(\mu + s_j(1 - \mu)) \) is always positive if \( \phi < \mu \pi r (1 - s_i) \), which is a necessary condition for \( \alpha_j > 0 \), and lemma 1 implies \( \alpha_j > 0 \) is necessary for an equilibrium. Given that \( (\mu \pi r (1 - s_i) - \phi)(\mu + s_j(1 - \mu)) s_j > 0 \), it follows that the denominator of (4.15) \( (\mu \pi r (1 - s_j) - \phi)(\mu + s_i(1 - \mu))(1 - s_i) - (\mu \pi r (1 - s_i) - \phi)(\mu + s_j(1 - \mu)) \) must be negative in order for \( F_{ri}/2 \) to be less than 1. Now, note that equation (4.14) also can be stated as:

\[
F_{ri}/2 = (1 - \mu) \pi r (1 - s_j) + \phi \mu)(s_i - s_j) / (\mu \pi r (1 - s_j) - \phi)(\mu + s_i(1 - \mu)) s_j - (\mu \pi r (1 - s_i) - \phi)(\mu + s_j(1 - \mu)) (4.16)
\]

As was just shown, the denominator of equation (4.15) must be negative. Because the denominator is the same in both equations (4.16)) and (4.15), in order for
\( F_{ri}/2 \) to be positive, the numerator in equation (4.16) must also be negative. Since 
\((1 - \mu)\mu\pi_r (1 - s_i)(1 - s_j) + \phi \mu \dot{\mu} > 0 \) for any \( s_i \) and \( s_j \), the numerator in (4.16) is negative if and only if \((s_i - s_j) < 0\), or \( s_i < s_j \); the firm \( i \) with the mass point has a lower market share than the firm \( j \) with no mass point.

The result that the firm with the small market share has a mass point shows that the firm with the big market share will adopt a more competitive mixed strategy price distribution than the small firm. The introduction of the gatekeeper in the market is more beneficial to the firm with the small market share than it is to the firm with the big market share. The firm with the big market share is threatened by the introduction of an information market. It has more customers to protect than the firm with the small market share and therefore it behaves more competitively when it advertises in order to keep its customers.

**Proposition 3** In an equilibrium with two firms, the firm with the small market share advertises with a higher probability than the firm with the large market share.

**Proof.** Proposition 2 shows that only the firm with the small market share can have a mass point, or \( s_i < s_j \), which implies that \( \frac{\mu + s_j (1 - \mu)}{\mu + s_i (1 - \mu)} > 1 \). It follows from equation (4.13) that in order for the mass point \( F_{ri} \) to be positive \( \frac{\alpha_i}{\alpha_i} \frac{\mu + s_j (1 - \mu)}{\alpha_j} \frac{\mu + s_i (1 - \mu)}{\mu + s_i (1 - \mu)} \) must be less than 1, which requires \( \frac{\alpha_j}{\alpha_i} < 1 \), or \( \alpha_j < \alpha_i \).

### 4.2.2 The Gatekeeper’s Problem

The gatekeeper wants to maximize its profits by choice of \( \phi \):

\[
\max_{\phi} \pi_G = (\alpha_i + \alpha_j)\phi + \mu \kappa - F
\]

\[s.t. \quad \alpha_i < 1\]

\[\alpha_i > 0\]

\[\alpha_j > 0, \text{ and}\]

\[F_{ri}/2 \leq 1/2.\]
Replacing $F_{ri}/2$, $\alpha_i$ and $\alpha_j$ by their expressions we get the following maximization problem:

$$\max_{\phi} \pi G = \frac{[\mu + s_j(1-\mu)] + (s_j-s_i)(\mu + s_i(1-\mu))}{s_j(\mu + s_i(1-\mu))}$$  \hspace{1cm} (4.19)

$$-\frac{\phi}{\mu, r_i, s_j(\mu + s_i(1-\mu))} [\phi + \mu \kappa - F]$$

s.t. $\phi > \mu \pi_r(1-s_i) \frac{(1-\mu)(s_i - s_j)}{(\mu + s_i(1-\mu))(1-s_i) - (\mu + s_j(1-\mu))}$

$\phi < \mu \pi_r(1-s_i) \frac{\mu s_j + (1-\mu)(s_j - s_i)}{\mu s_i + (1-\mu)(s_j - s_i)}$

$\phi < \mu \pi_r(1-s_i)$

and $\phi < \mu \pi_r(1-s_i) \frac{(\mu + s_j(1-\mu))(2s_j - 1) + (1-s_j)(\mu + s_i(1-\mu))}{(\mu + s_j(1-\mu))(2s_j - 1) + (1-s_i)(\mu + s_i(1-\mu))}$

The first inequality imposes a lower bound on $\phi$ and the last three inequalities impose an upper bound on $\phi$. Since $s_j > s_i$ it follows that $\mu s_j + (1-\mu)(s_j - s_i) > \mu s_i + (1-\mu)(s_j - s_i)$ and $\mu \pi_r(1-s_i) \frac{\mu s_j + (1-\mu)(s_j - s_i)}{\mu s_i + (1-\mu)(s_j - s_i)} > \mu \pi_r(1-s_i)$. Furthermore $(\mu + s_j(1-\mu))(2s_j - 1) + (1-s_j)(\mu + s_i(1-\mu)) < (\mu + s_j(1-\mu))(2s_j - 1) + (1-s_i)(\mu + s_i(1-\mu))$ and $\mu \pi_r(1-s_i) \frac{(\mu + s_j(1-\mu))(2s_j - 1) + (1-s_j)(\mu + s_i(1-\mu))}{(\mu + s_j(1-\mu))(2s_j - 1) + (1-s_i)(\mu + s_i(1-\mu))} < \mu \pi_r(1-s_i)$, which implies that if $F_{ri}/2 < 1/2$ then $\alpha_j > 0$ and $\alpha_i > 0$. The four conditions on $\phi$ can be summarized by:

$$\phi > \mu \pi_r(1-s_i) \frac{(1-\mu)(s_i - s_j)}{(\mu + s_i(1-\mu))(1-s_i) - (\mu + s_j(1-\mu))} = C_1$$  \hspace{1cm} (4.20)

$$\phi < \mu \pi_r(1-s_i) \frac{(\mu + s_j(1-\mu))(2s_j - 1) + (1-s_j)(\mu + s_i(1-\mu))}{(\mu + s_j(1-\mu))(2s_j - 1) + (1-s_i)(\mu + s_i(1-\mu))} = C_2$$  \hspace{1cm} (4.21)

**Proposition 4** Let $\phi' = \frac{\mu \pi_r s_i}{2}(1 + (s_j - s_i) \frac{\mu + s_i(1-\mu)}{\mu + s_j(1-\mu)})$. In equilibrium the gatekeeper charges $\phi^*$ such that

$$\phi^* = \phi' \text{ if } C_1 < \phi' < C_2$$  \hspace{1cm} (4.22)
\[ \phi^* = C_1 \text{ if } C_1 > \phi' \]
\[ \phi^* = C_2 - \varepsilon \text{ if } C_2 < \phi'. \]

4.2.3 Market Equilibrium

Proposition 5 In a market with two firms, firm \( i \) and firm \( j \) with market shares \( s_i \leq s_j \) and a consumer subscription ratio \( \mu \), an equilibrium with a market for information exists if and only if \( \frac{1}{2(3\mu - 3)}(3\mu - 4 + \sqrt{-3\mu^2 + 4}) \leq s_i \leq 1/2 \). In this equilibrium:

(i) The gatekeeper sets an advertising fee \( \phi^* \),

(ii) firm \( j \) advertises on the gatekeeper’s site with probability \( \alpha_j^* = 1 - \frac{\phi^*}{\mu s_i s_j} \), charges the monopoly price \( r \) when it does not advertise, and when it advertises, its distribution of advertised prices is given by the cumulative distribution function \( F_j^*(p) = \frac{1}{\alpha_j^*}(1 - \frac{(1-\alpha_j^*)\mu \pi_r + (1-\alpha_j^*)s_j(\pi_r - \mu)}{\mu \pi_p}) \), which has an upper support \( p = r \), and a lower support \( p^* = \pi^{-1}(\frac{\phi^* + s_j(1-\mu)}{\mu(1-s_i)/(\mu + s_i(1-\mu))}) \).

(iii) firm \( i \) advertises on the gatekeeper’s site with probability \( \alpha_i^* = \frac{(1-s_j)\mu \pi_r - \phi^*}{\mu s_i (1-s_j-F_i^*/2)} \), charges the monopoly price when it does not advertise, and when it advertises, its distribution of advertised prices is given by the cumulative distribution function \( F_i^*(p) \) which has a mass point \( F_{ri}^* \) at the monopoly price. \( F_{ri}^* = 2(1 - \frac{\alpha_j^* \mu + s_j(1-\mu)}{\alpha_i^* \mu + s_i(1-\mu)}) \), and \( F_i^*(p \neq r) = \frac{1}{\alpha_i^*}(1 - \frac{\pi_r(\mu(1-\alpha_i^* + \alpha_j^* s_j^2/2) + (1-\mu)s_j - (1-\mu)s_j \pi_p)}{\mu \pi_p}) \). The upper support of the distribution is the monopoly price \( r \) and the lower support of the distribution is \( p^* = \pi^{-1}(\frac{\phi^* + s_j(1-\mu)}{\mu(1-s_i)/(\mu + s_i(1-\mu))}) \).

If \( \frac{1}{2(3\mu - 3)}(3\mu - 4 + \sqrt{-3\mu^2 + 4}) > s_i \), a market for information cannot exist in equilibrium.

Proof. A necessary condition for the existence of a market for information is that \( C_1 < \phi < C_2 \). In order for this condition to be satisfied the interval between \( C_1 \) and \( C_2 \) cannot be an empty set. This requires that \( C_1 < C_2 \). The polynomial \( \frac{(1-\mu)(s_i-s_j)}{\mu + s_i(1-\mu)(s_i-s_j) - (\mu + s_j(1-\mu))} \frac{(\mu + s_j(1-\mu))(2s_i-1) + (1-s_j)(\mu + s_i(1-\mu))}{(\mu + s_i(1-\mu))(2s_j-1) + (1-s_i)(\mu + s_j(1-\mu))} \) has four roots: \( s_i = \)
\{1, \frac{\mu}{\mu-1}, \frac{1}{6(\mu-1)}(3\mu - 4 \pm \sqrt{4 - 3\mu^2})\}. Since \(s_i\) is the small firm’s market share, \(s_i\) must be between 0 and 1/2. Furthermore it can be shown that for \(0 \leq \mu \leq 1\), 
\[\frac{\mu}{\mu-1} < 0 \quad \text{and} \quad \frac{1}{6(\mu-1)}(3\mu - 4 - \sqrt{4 - 3\mu^2}) > 1.\] It follows that the condition necessary for the existence of an equilibrium is

\[\frac{1}{6(\mu-1)}(3\mu - 4 + \sqrt{4 - 3\mu^2}) < s_i < 1/2 \quad (4.23)\]

4.3 Comparative Statics

The model presented above allows us to understand the effect of firm heterogeneity with respect to their market shares on the dynamics of market for information on the Internet. As was shown in proposition 9, a market for information does not always exist. Two important factors determine whether a market for information could exist: the ratio of subscribers \(\mu\), and the market shares of firms \(s_i\) and \(s_j\). As the ratio of subscribers \(\mu\) increases, smaller and smaller \(s_i\) can insure the existence of a market for information. As we show in the numerical example below, for the case of full consumer participation in the market for information (i.e., \(\mu = 1\)), a market for information exists for any value of \(s_i\), \(0 < s_i < 1/2\). On the other hand as \(\mu\) decreases, the lower bound on \(s_i\) necessary for a market to exist is increasing. Even relatively high \(s_i\) can not ensure the existence of a market for information.

The limit necessary for the existence of a gatekeeper of the lower support of \(s_i\) as \(\mu\) approaches zero is \(\frac{1}{3}\), furthermore, for \(s_i > 1/3\), as \(\mu\) goes to zero, \(\phi\) goes to zero. It follows that for firms having equal market shares as in Baye and Morgan (2001) a market for information always exists. However, this does not generalize to the case in which firms have different market shares. The market for information will not exist if firm market shares are too different. This suggests that the gatekeeper will have an incentive to advertise its services, to attract more consumers to subscribe to its website and thus insure the existence of a market for information.
Equation (4.22) shows that the advertising fee $\phi^*$ is proportional to the firm’s monopoly profit $\pi_r$. Higher monopoly profits allow the gatekeeper to charge higher fees. Replacing $\phi^*$ in equations (4.3), (4.14) and (4.10) shows that the advertising probability $\alpha^*_j$ of firm $j$ as well as the mass point $F^*_{ri}$ do not change if the monopoly profit $\pi_r$ changes. Similarly replacing $\phi^*$ and $F^*_{ri}$ in equation (4.10) shows that the probability $\alpha_i$ that firm $i$ advertises does not change in the monopoly profit $\pi_r$. If the gatekeeper charges a higher advertising fee, and firms do not change their advertising behavior the gatekeeper’s profit increases. In the numerical examples that follow, we set the monopoly profit to one, which allows us to express the advertising fee $\phi$ as a percentage of the monopoly profit. This implies that the advertising fee would change from market to market according to the monopoly profit in each market, but the advertising behavior of firms would be the same similar market shares $s_i$ and $s_j$ and consumers subscription rates $\mu$.

From equations (4.3) and (4.22) it can also be shown that the advertising probability of the large firm $\alpha^*_j$ is always smaller than 1/2.

$$\alpha^*_j = \frac{1}{2} - \frac{s_j - s_i \mu + s_i(1 - \mu)}{2 \mu + s_j(1 - \mu)}$$  \hspace{1cm} (4.24)

Equation (4.24) also shows that $\alpha^*_j$ is increasing in $s_i$. As the two firms market shares converge the large firm behavior becomes more competitive, it advertises with higher probability, and thus charges the monopoly price with lower probability. This can be explained by the fact that as the two firms’ market shares converge, the size of the non-local market the large firm has the potential to capture by advertising is increasing. On the other hand when the market share of the large firm is much higher than the small firm’s market share, the probability the large firm advertises falls. Moreover, a higher consumer subscription ratio $\mu$ increases the benefit firms get from advertising, and thus it leads to a higher probability $\alpha^*_j$ firm $j$ advertises, even though the advertising fee $\phi^*$ is also increasing in the subscribers ratio $\mu$. 

65
4.4 Numerical Example, Case when \( \mu = 1 \)

Let \( \mu = 1 \) and \( \pi_r = 1 \). As shown above, the monopoly profit \( \pi_r \) does not change the probabilities firm \( i \) and firm \( j \) advertise, nor does it change the the mass point. \( \pi_r \) is proportional to the advertising fee and to the profit at the lower support \( p \). Thus setting the monopoly profit \( \pi_r = 1 \) allows us to express the advertising fee as a percentage of the monopoly profit. When \( \mu = 1 \) there is full consumer participation in the market for information. Since the gatekeeper’s profit is increasing in the subscription rate \( \mu \), one can expect the gatekeeper to try to attract full consumer participation in the market for information. In this case, consumers buy from their local firm only if no prices are advertised on the gatekeeper’s site.

First we examine the range for \( s_i \) for which the market for information exists. The conditions on \( \phi \) are:

\[
\begin{align*}
\phi & > 0 \quad (4.25) \\
\phi & < \frac{s_j^2}{2s_j - s_i} \quad (4.26)
\end{align*}
\]

It follows from these equations that a market for information exists for any \( 0 < s_i < 1/2 \). The advertising fee \( \phi \) that maximizes the gatekeeper’s profits is \( s_j^2 \). For \( s_i \leq 1/3 \), the higher bound condition on \( \phi \) will be binding, meaning \( \phi^* = \frac{s_j^2}{2s_j - s_i} \). For \( 1/3 < s_i < 1/2 \), the condition on the advertising fee \( \phi \) is not binding and \( \phi^* = s_j^2 \).

For \( 0 < s_i < 1/3 \), we get:

\[
\begin{align*}
\phi^* &= \frac{s_j^2}{2s_j - s_i} - \epsilon \\
\alpha_j^* &= \frac{1 - 2s_i}{2 - 3s_i} \\
F_{ri}^*/2 &= 1/2 - \epsilon \\
\alpha_i^* &= \frac{2 - 4s_i}{2 - 3s_i}
\end{align*}
\]

For \( 1/3 \leq s_i < 1/2 \), we get:

\[
\phi^* = s_j^2
\]

(4.28)
\[
\alpha_j^* = s_i \\
\overline{F}_{ri}^* / 2 = 1 - \frac{s_i}{s_j} \\
\alpha_i^* = s_j
\]

For \( \mu = 1 \), the mass point \( \overline{F}_{ri}^* \) is decreasing in \( s_i \), and reaches 0 for \( s_i = 1/2 \). Figure 4.1 shows the relationship between \( \overline{F}_{ri}^* / 2 \) and the market share of the small firm for \( s_i > 1/3 \). The probability \( \alpha_i^* \) that firm \( i \) advertises is decreasing in \( s_i \), whereas the probability \( \alpha_j^* \) that firm \( j \) advertises is increasing in \( s_i \). As the market shares of the two firms converge, so does the probability they advertise.

Figure 4.2 shows the relationship between the advertising fee \( \phi \) and the market share of the small firm for \( s_i > 1/3 \). The advertising fee \( \phi^* \) is decreasing in the market share of the small firm \( s_i \), as the market shares of both firms converge, the gatekeeper can extract lower fees from the firms. Given that the sum of the probabilities the two firms advertise \( \alpha_i^* + \alpha_j^* = 1 \), it means that the gatekeeper’s profit is decreasing in the small firm’s market share \( s_i \).

### 4.5 Conclusion

In this paper we investigate information provision in a market where firms have different market shares. Baye and Morgan (2001) provide analysis of information provision in a market where firms are homogeneous. They find that when all firms have the same market share, an equilibrium in which a gatekeeper exists is always possible. In this equilibrium firms advertise with a positive probability that is strictly smaller than one, and use mixed pricing strategies when advertising. They also find that it is optimal for the gatekeeper to induce full consumer participation in the market for information. When generalizing their model to include the case when firms have different market shares, the central finding is that a market for information is not always possible. We identify two cases, one with a gatekeeper, and one without. A subgame perfect Nash equilibrium with a market for information
Figure 4.1: Relationship between $F_{ri}/2$ and the market share of the small firm $s_i$ when $\mu = 1$. 
Figure 4.2: The advertising fee $\phi$ as a function of the market share of the small firm $s_i$. 
exists conditional on the market shares of the two firms not being too different given
the rate of consumer subscription to the gatekeeper’s site. The larger the big firm
becomes, the harder it is for the gatekeeper to exist. The bigger the market share of
the large firm, the higher is the ratio of consumers subscribing to the gatekeeper’s
site needed to insure the gatekeeper’s existence. When there is full consumer par-
ticipation in the market for information, the gatekeeper exists at any firms’ market
shares. In the absence of a market for information price dispersion disappears and
both firms charge the monopoly price.

Another important result is that the small firm advertises with a higher
probability than the large firm, but it advertises the monopoly price with a positive
probability despite the added competition faced by the small firm on the gatekeeper’s
site. The big firm advertising probability is decreasing in its market share. This result
is consistent with the fact that CircuitCity.com, a firm with a large market share
in the electronics market, is absent from the list of prices provided in the Baye et
al. (2001) research of more than 4 million observations of computer and electronic
products collected from Shopper.com.

Finally we note that the results of our model apply to any market regard-
less of the good sold. As we have seen, the probabilities firms advertise and the
mass point placed by the small firm at the monopoly price do not change with the
monopoly profit. The advertising fee charged by the gatekeeper is proportional to
the monopoly profit. This implies that the advertising fee will differ from market
to market, which leads us to think that, at least in markets where monopoly profits
are high, several firms will be attracted to the market of information provision will
attract several gatekeepers. An interesting generalization of the model would be to
allow for competition among gatekeepers.

Even though our analysis is complete only for the case of two firms, it is
consistent with at least some casual observation of information provision in online
markets. Some of the most popular online price comparison sites Shopper.com and Mysimon.com provide price information for computer and electronic products. The market for computer and electronic products is one of the most mature and active online markets. It is also characterized by high consumer participation in the market for information. So even though firms may have dramatically different market shares, gatekeepers are still able to exist.

Examples of further possible research include the endogenization of the consumer subscription ratio and the subscription rate charged by the gatekeeper to consumers. As this model shows the consumer participation in the market for information is a key determinant of the existence of the gatekeeper. Therefore we would expect the gatekeeper to take measures insuring a high consumer subscription rate. Another important follow-up to this research would be to conduct an empirical test of the model. One could study the relationship between the market share of the firms that advertise on a price comparison site and the frequency of their advertisement.
E-commerce is rapidly becoming an important and major part of the economy. Understanding online markets is thus a crucial question for economists. As several empirical studies have recently shown, it is not accurate to think of online markets as a realization of the perfectly competitive markets where prices converge to marginal cost. It is true that price information is available at very low cost on the Internet, which in theory should eliminate price dispersion, but at the same time Internet has changed the competitive environment that faces firms. Consumers are faced with uncertainties when shopping online that didn’t exist in traditional markets. Moreover the low cost of running and operating a website has incited many sellers to enter online markets. This has resulted in a big variety of firms of different sizes providing services with different qualities. Firm heterogeneity is therefore very common on the Internet.

This dissertation’s main goal is to contribute to the understanding of online markets. First using existing theoretical models I test the hypothesis that unobserved product characteristics in particular availability is sufficient to generate price dispersion in equilibrium even if prices are common knowledge. Data from the online market for college textbooks are used to test this hypothesis. All of the empirical tests conducted in section (3.4) support the predictions of the theoretical model. These tests indicate that some online firms such as amazon.com and bn.com charge above average prices and have above average availability, while others, such as,
alltextbooks4less.com, booksamillion.com, ecampus.com, textbooksource.net, and studentmarket.com charge below average prices and have below average availability. Furthermore, the website offering the lowest price has lower availability than the website offering the second lowest price. The claims that online markets should be perfectly competitive (or at least nearly so) depend crucially on the assumption that information is readily available online. While this is true for price information, it is not true in general for all parameters of interest to consumers. Because a buyer cannot walk into an online retailer and physically determine an item’s availability and quality, and then walk out with the item in hand, consumer information when purchasing online is incomplete in many dimensions. The first part of the dissertation presents evidence regarding the important role of information about product availability and its effect on prices in online markets. The fact that inadequate information regarding other aspects of the transaction also inhibits perfectly competitive pricing in online markets is an important area for future research. Many predictions of perfectly competitive pricing online have not been realized and further study may provide an explanation. Online prices are far from the perfectly competitive markets predicted from full price information, and it may be the case that firms use prices to signal some important quality to consumers.

The second part of the dissertation studies the effect different market shares have on online information provision. One important motivation for the predictions that prices for a homogeneous good offered online would converge to the competitive price is the existence of price comparison sites. Price comparison sites are gatekeepers that allow a flow of information between firms and consumers for some fee. A theoretical model by Baye and Morgan (2001) shows that if firms are identical, the gatekeeper is always able to establish an information market where prices are dispersed. The model presented in chapter 4 makes the more realistic assumption that which firms have different market shares. The major finding of this model is that
consumer participation in the market for information is a major factor determining whether a gatekeeper can exist or not. The other important factor is the relative market shares of firms. If some firms control a big part of the market, they may be able to preclude the existence of the gatekeeper. In the case of existence of a market for information, the small firms tend to advertise their prices with a higher probability than the big firm, but they advertise higher prices on average.

The Internet is still relatively young and most online markets are not mature. As time passes, more studies and research become necessary to assess better the difference between regular and online markets. Examples of further research include several empirical studies of the relationship between the unobserved product characteristics and price on the Internet. The test conducted in chapter 3 can be generalized to other markets over a longer period of time. Empirical evidence is also necessary whether firms use prices to signal product quality which otherwise would be unobserved by consumers. Another important empirical study is to examine the relationship between a firm’s market share, the probability it advertises on a gatekeeper’s site, and the distribution of prices it advertises. Note that the model predicts that if the firm with the big market share advertises, it advertises lower prices than firms with low market share, but if it does not advertise it charges the monopoly price. Another interesting generalization would be to allow firms to have different capacity constraints in Arnold’s (2000) model. This would be a closer match to online reality. Internet markets are characterized by firms with different capacities. For example, in the market for textbooks there are a few websites with a large capacity, such as amazon.com and barnes&noble.com, which warehouse books, and many smaller retailers, with more limited capacities, that do not carry most of the textbooks. Finally, as information provision becomes a profitable business, it is important to study the effect competition between gatekeepers has on prices.
REFERENCES


