ESSAYS ON TWO SOURCES OF MARKET IMPERFECTIONS:
CONSUMER SEARCH COSTS AND CONSUMER MYOPIA

A Dissertation in
Economics
by
Marlon L. Williams

© 2013 Marlon L. Williams

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

December 2013
The dissertation of Marlon L. Williams was reviewed and approved* by the following:

Mark J. Roberts  
Professor of Economics and NBER Research Associate  
Dissertation Co-Adviser and Co-Chair of Committee

Edward Green  
Professor of Economics and Mathematics  
Dissertation Co-Adviser and Co-Chair of Committee

Edward Coulson  
Professor of Economics and Professor of Real Estate Economics  
Jeffrey and Cindy King Fellow In Real Estate

Lisa Posey  
Associate Professor of Business Administration  
The Department of Risk Management in Smeal School of Business

Vijay Krishna  
Distinguished Professor of Economics  
Director of Graduate Studies, Economics

*Signatures are on file in the Graduate School.
Abstract

The dissertation comprises three self-contained essays in the general field of industrial organization, and focuses on two sources of market imperfections: Consumer search costs in the automobile insurance industry and consumer myopia in the banking industry. The first two essays analyze the interaction between consumers’ search behavior and firms’ advertising decisions. Both papers are motivated by the observation that advertising spending has exploded in the auto insurance industry over the last two decades, while simultaneously consumer search costs have arguably fallen equally dramatically. The overarching goal of these two papers is to provide a cogent, coherent link between increased advertising and decreased consumer search costs in a market where firms produce a homogeneous product. Establishing such a connection is an important step in extending our understanding of the role of consumer search costs in shaping the competitive landscape of various industries.

To fully appreciate the intricate interplay between consumers’ search behavior and firms’ advertising decisions, we must have an acute understanding of how the pricing decisions of firms are influenced by changes in consumer search costs. Chapter 2 provides a thorough discussion of the pricing implications of such changes. Chapter 3 analyzes firms’ pricing decisions when consumers are faced with a different type of limitation – consumer myopia. I investigate the factors that influence how banks price overdraft services, an add-on good, in the presence of myopic consumers.

In Chapter 1, “Estimating the Effects of Non-informative Advertising when Consumers Search Sequentially,” I use the auto insurance industry as a case study to establish that consumer search costs have fallen significantly, at least in some industries. The advent of the Internet has allowed consumers in many industries to now access product information much easier and much more efficiently than in previous decades. Specifically, in the auto insurance industry, as
more firms increase their online presence and with increased Internet penetration, consumers increasingly are able to obtain multiple quotes for insurance products in a much shorter time. Though this assertion is non-contentious, it does not necessarily imply that consumer search costs have fallen, because other changes, such as increased value for time, could outweigh the reduction in search time.

In chapter 1, I rigorously establish that search costs have been falling over the last two decades in the private passenger auto insurance industry of California, while I simultaneously estimate the sensitivity of the probability that a firm is visited to the advertising levels of that firm and of its competitors. This requires that I estimate the consumer search costs distribution function. The challenge is to undertake an estimation procedure that generally requires detailed consumer-level data, when we only have highly aggregated firm-level data. I make use of Hortaçsu and Syverson’s (2004) model, which addresses this problem. Consistent with a priori expectations, I find that overall consumer search costs have been falling over time, especially between 2005 and 2010. The results also confirm that consumers are more likely to visit a firm that advertised in the past, and that this advertising effect is increasing over time.

Chapter 2, “Dynamic Choice of Non-informative Advertising when Consumers Search Sequentially: A Numerical Approach,” builds on the findings of chapter 1 by providing a theoretical model that shows how decreasing consumer search costs can lead to increased advertising activity, when firms produce a homogeneous product. This result is an exception to the results of traditional advertising models, which directly or indirectly posit that consumer search and advertising are substitutes. In these models consumers’ incentives to search fall as firms provide more and better advertising messages. Similarly, firms’ incentives to advertise decrease as consumers search more intensely. This implies that advertising should fall as consumer search costs fall. In chapter 2, I establish reasonably weak conditions under which firms actually choose optimally to increase advertising in response to a decrease in search costs.

I do so by showing that the relationship between advertising and search costs is not monotone because the relationship between search costs and the firms’ profitability is first not monotone. I use a dynamic model, where firms choose advertising to affect their market exposure, the state variable; and choose prices in a product market, where consumers with heterogeneous search costs search sequentially. In short, I embed a version of Carlson and McAfee’s (1983) demand model into Doraszelski and Markovich’s (2007) dynamic advertising framework. The results of chapter 2 have significant policy implications, as they suggest that though markets may ultimately become very competitive at very low levels of search costs, markets may first have to pass through stages of being highly concentrated as search costs fall.
Chapter 3, “Overdraft Pricing and Myopic Consumers,” investigates whether the presence of myopic consumers impacts how banks choose the prices of bank overdraft services. The pricing of bank products has long been a focus of empirical researchers. But in recent years overdraft services have received additional attention because banks and credit unions routinely earn an annual percentage rate (APR) of over 1,000 percent on these products. Such abnormally high returns in an arguably competitive industry require explanation. Gabaix and Laibson (2006) propose a model where banks optimally choose to suppress information on the price of add-on services, and charge prices that are substantially higher than marginal cost. They argue that firms are able to do so because some consumers are myopic, in the sense that they do not consider some add-on services when choosing the base product. The authors also show that no firm has an incentive to “educate” the myopic consumers if the fraction of myopic consumers is sufficiently large.

In chapter 3, I test the applicability of Gabaix and Laibson’s (2006) model to the pricing of overdraft services using a repeated cross-section of U.S. banks and credit unions. I use three demographic characteristics to serve as proxies for the fraction of myopic consumers in each market, and find that two of the three proxies are significant and have the predicted signs. These results, though not conclusive, suggest that structural estimation of Gabaix and Laibson’s model may add substantially to our understanding of pricing decisions in aftermarkets.
Table of Contents

List of Figures ix

List of Tables x

Acknowledgments xi

Chapter 1

Estimating the Effects of Non-informative Advertising when Consumers Search Sequentially 1

1.1 Introduction .................................................. 1

1.2 Theoretical Model ........................................... 5
   1.2.1 Theory ................................................ 5
   1.2.2 Identification and Estimation Strategy .................. 10
      1.2.2.1 Heterogeneous Firms with Equal Sampling .......... 10
      1.2.2.2 Unequal Sampling with Homogeneous Firms ........ 11

1.3 Data ......................................................... 13
   1.3.1 Data description ...................................... 13
   1.3.2 Defining the Unit of Sale .............................. 15
   1.3.3 Defining the Market .................................. 17

1.4 Estimation Results .......................................... 18
   1.4.1 Version 1: Homogeneous Firms with Unequal Sampling ... 18
      1.4.1.1 Consumer Search Cost Distribution ................. 18
      1.4.1.2 Sampling Probability ............................. 21
   1.4.2 Version 2: Heterogeneous Firms with Equal Sampling ..... 26

1.5 Concluding Remarks ......................................... 29

Bibliography .................................................... 32
# Chapter 2

**Dynamic Choice of Non-informative Advertising when Consumers Search Sequentially: A Numerical Approach**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 Introduction</td>
<td>33</td>
</tr>
<tr>
<td>2.2 Model Overview</td>
<td>36</td>
</tr>
<tr>
<td>2.2.1 Environment</td>
<td>37</td>
</tr>
<tr>
<td>2.2.2 Timing</td>
<td>37</td>
</tr>
<tr>
<td>2.3 The Pricing Game</td>
<td>38</td>
</tr>
<tr>
<td>2.3.1 Set-up</td>
<td>38</td>
</tr>
<tr>
<td>2.3.2 Consumer Search and Firm Demand</td>
<td>38</td>
</tr>
<tr>
<td>2.3.3 Profit-maximization and Equilibrium</td>
<td>42</td>
</tr>
<tr>
<td>2.4 Stage I: The Dynamic Advertising Choice</td>
<td>47</td>
</tr>
<tr>
<td>2.4.1 State-to-State Transitions</td>
<td>47</td>
</tr>
<tr>
<td>2.4.2 Bellman Equation and Policy Function</td>
<td>48</td>
</tr>
<tr>
<td>2.4.3 Equilibrium and Parameterization</td>
<td>49</td>
</tr>
<tr>
<td>2.5 Results</td>
<td>51</td>
</tr>
<tr>
<td>2.5.1 Product Market Competition</td>
<td>51</td>
</tr>
<tr>
<td>2.5.2 Value Function and Policy Function</td>
<td>55</td>
</tr>
<tr>
<td>2.5.3 Limiting Distribution</td>
<td>62</td>
</tr>
<tr>
<td>2.5.4 The Big Question Answered</td>
<td>65</td>
</tr>
<tr>
<td>2.5.5 Welfare Discussion</td>
<td>66</td>
</tr>
<tr>
<td>2.6 Concluding Remarks</td>
<td>70</td>
</tr>
</tbody>
</table>

**Bibliography**

## Chapter 2 Appendix

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.A Supplemental Figures and Tables</td>
<td>74</td>
</tr>
<tr>
<td>2.A.1 Pricing and Advertising when ( b = 50 )</td>
<td>74</td>
</tr>
<tr>
<td>2.A.2 Simulation Results when ( b = 40 ) and ( 30 )</td>
<td>75</td>
</tr>
<tr>
<td>2.A.3 Summary Results Tables: Value and Policy Functions for ( b = 40, 30, ) and ( 20 )</td>
<td>77</td>
</tr>
<tr>
<td>2.B Existence of Pure Strategy Nash</td>
<td>80</td>
</tr>
</tbody>
</table>

# Chapter 3

**Bank Overdraft Pricing and Myopic Consumers**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Introduction</td>
<td>85</td>
</tr>
<tr>
<td>3.2 A Model of Add-on Pricing: Gabaix and Laibson (2006)</td>
<td>87</td>
</tr>
<tr>
<td>3.3 The regression model</td>
<td>92</td>
</tr>
<tr>
<td>3.4 Data description</td>
<td>93</td>
</tr>
<tr>
<td>3.5 Results</td>
<td>95</td>
</tr>
<tr>
<td>3.5.1 Control variables</td>
<td>96</td>
</tr>
</tbody>
</table>
3.5.2 Variables of interest: Proxies for the fraction of myopic consumers ........................................... 101
3.6 Concluding remarks .................................................................................................................... 103
List of Figures

1.1 CDFs with Unequal Sampling (1996-2001) .............................................. 21
1.2 CDFs with Unequal Sampling (2005-2010) .............................................. 22
1.3 Sampling Probability Distribution .......................................................... 25
1.4 CDFs with Equal Sampling ................................................................. 27
1.5 Indirect Utility Distribution ................................................................. 29

2.1 Demand Functions in a Bertrand Model and McAfee Model ...................... 41
2.2 Demand and Profit Functions .............................................................. 43
2.3 Best Response Functions ................................................................. 46
2.4 Pricing Function for Different Search Costs Levels ................................ 52
2.5 Profit Function for Different Search Costs Levels ................................ 54
2.6 Value Function and Policy Function for $b = 40$ and $b = 30$ ............... 55
2.7 Value Function and Policy Function for $b = 20$ ................................... 57
2.8 Value Functions: $b = 20, 30, 40$ .................................................... 57
2.9 Simulations when $b = 20$ and Initial State is $(2, 2)$ .............................. 61
2.10 Limiting Distribution for Different Levels Consumer Search Costs .......... 63
2.11 Consumer Search Costs Incurred ....................................................... 68
2.12 Consumer Surplus and Producer Surplus ......................................... 69
2.A.1 Comparison Figures for $b = 50$ ..................................................... 75
2.A.2 Simulations with $b = 40$ and Initial State, $(15, 1)$ .......................... 76
2.A.3 Simulations with $b = 30$ and Initial State $(15, 1)$ ............................. 77
2.B.1 CDF and PDF when $c \sim \ln N(1, 0.5)$ .................................................. 81
2.B.2 Demand and Profit Functions when $c \sim \ln N(1, 0.5)$, $\rho_1 = 0.5$ and $p_2 = $18.88 .............................................................. 82
2.B.3 Best-response Function when $c \sim \ln N(1, 0.5)$ and $\rho_1 = 0.5$ ........ 83
## List of Tables

1.1 Estimates for Consumer Search Costs Distributions (Unequal Sampling) ........................................... 19
1.2 Estimates for Sampling Probability Functions, $\rho$ ................................................................. 23
1.3 Implied Sampling Probability Distributions (1996-2010) .......................................................... 24
1.4 Estimated Search Cost Distributions and Indirect Utility Distributions ................................................. 28

2.A.1 Summary Results for Firm 1’s Value Function ........................................................................... 78
2.A.2 Summary Results for Firm 1’s Policy Function ........................................................................... 79
2.A.3 Limiting Distribution Market Structure ..................................................................................... 80

3.1 Summary Statistics .................................................................................................................... 96
3.1 Results of OLS regression .......................................................................................................... 98
Acknowledgments

I would not have been able to successfully complete this dissertation without the help of countless individuals. And at the risk of missing many persons who played a part, directly or indirectly, I will say thanks to a few persons here. Obviously, this list is by no means exhaustive.

I would like to extend my sincere appreciation to Professors Ed Coulson and Lisa Posey for serving on my thesis committee and to Professors Mark Roberts and Ed Green, my co-advisers, who provided invaluable support, direction, and advise. Thanks to Professors Damien King and Michael Witter of the University of the West Indies, along with Professors John Riew, Robert Marshall, Vijay Krishna, and Neil Wallace, who were instrumental in me getting the opportunity to study at the Pennsylvania State University.

Thanks to the Department of Economics, PSU, for providing the vast majority of the funds needed for my degree. Thanks to Bates White LLC for a very worthwhile internship experience, during which the idea for chapter 3 was born; and for funding my third year by providing a fellowship. Thanks to the College of Liberal Arts, PSU, for a dissertation support grant, which allowed me to purchase the necessary data for chapter 1 from A. M. Best Ltd. Thanks to Luciano Gobbo and other staff members at the California Department of Insurance, who graciously provided much of the data for chapter 2, along with important insights into the auto insurance industry. Cindy Fillman of the Pennsylvania Insurance Department also helped me to gain a better understanding of the industry, especially with regards to the policy issues; for this assistance I am very grateful. Thanks to Ulrich Doraszelski and Sarit Markovich, who willingly made their programming codes available. These codes greatly reduced my programming burden.

Thanks to friends, such as Dr. Gary Lyn, Dr. Yanping Liu, Dr. Van Anh Vuong, Winston Robotham, and Orion Knepp, for their support and assistance of all types. Thanks to Pastor Leroy Thompson, my church families, my life group family, and to my other relatives and friends for their encouragement and for providing some semblance of balance in my life during this period. Thanks to
administrative staff in the Economics Department (past and present), including Lynn Sebulsky and Krista Winkelblech for going beyond their job descriptions to help students. Thanks to members of the Directorate of International Student Advising (DISA), PSU, such as Jessica Ouedraogo and Julie Mortimore, who provide excellent support for international students. Thanks also to many members of the PSU library staff, especially Diane Zabel, Kevin Harwell, and Debora Cheney, who were instrumental in helping me to access needed advertising data.

Special thanks to my mom, Annette Williams, who through-out my life has done everything in her power to give her kids the best opportunities to succeed. Thanks mom for your boundless love. Thanks to my two kids, Breanne and Zachary, who perhaps without knowing, provided daily motivation. And a very special thanks to my wife, Cameisha, who provided constant support during a difficult but exciting period in our lives. Thanks Mishka for loving me sacrificially.

And ultimately, thanks to my Father, without who this would not be possible. Actually without Him, nothing is possible.
Dedication

I dedicate this dissertation to my ever-supportive wife, Cameisha; my (almost) three-year-old daughter, Breanne; my four-month-old son, Zachary; and my ever-dependable mom, Annette; who all endured less than their rightful share of my time, to ensure that I could successfully complete this degree. I pray that this sacrifice will allow us to have a better future, and that I will use the knowledge and experience gained during this period to improve not only our lives, but also the lives of our neighbors, near and far.
Chapter 1

Estimating the Effects of Non-informative Advertising when Consumers Search Sequentially

1.1 Introduction

In this paper I estimate the consumer search costs distribution function in the auto insurance industry, while simultaneously recovering the sensitivity of the probability that a firm is visited to the advertising levels of that firm and of its competitors. Consumer search cost has repeatedly been stressed by microeconomic theorists as being an important feature in shaping the competitive landscape of many industries. Stahl (1989), for example, shows that in the presence of even small positive search costs, an increase in the number of firms can actually lead to more monopolistic pricing. It is easy to show that non-informative advertising has significant potential to exacerbate this problem.

In spite of the overwhelming theoretical support for the importance of consumer search costs, there is only a handful of empirical research that quantify the magnitude and distribution of search costs.

\[\text{The goal of this paper is two-fold:} \]

---

\[\text{1 One recent example of this type of work is Santos, Hortaçsu, and Wildenbeest (2012) that estimate the search-costs distribution for consumers that purchased books online from Amazon, Barnes and Noble, and Books Clubs. These authors are able to do so because they were able to obtain detailed web browsing data.} \]
(1) I quantify the effects of advertising in an industry where advertising is be-
coming increasingly important, (2) I provide additional empirical evidence of the
significance of consumer search costs even in a mature industry, and (3) I confirm
that consumer search cost has been falling in the auto insurance industry in recent
years.

One of the main factors that have stifled empirical work in this area is the
difficulty of obtaining suitable data. The data requirement is extremely high for
estimating consumer search costs. The traditional estimation procedure uses the
fact that the optimal decision rule in standard sequential search theory requires
that the expected benefits from an additional search must be equal to the expected
cost of that search. As such, we are able to estimate the minimum bound on
a consumer’s search cost if we know the consumer’s available options and have
an estimate of the indirect utility derived from each option. Thus, to estimate
consumer search costs in this way requires not only knowledge of the consumers’
actual choices, but also knowledge of each consumer’s choice set, which is likely to
vary significantly across consumers. In most industries, including auto insurance,
this type of detailed consumer-level data is extremely difficult to obtain, so the
traditional econometric approach cannot be used in many applications.

In this paper I am able to overcome this significant data requirement hurdle
by applying an alternative estimation strategy that was proposed by Hortaçsu and
Syverson (2004) (H&S). By expressing market shares as functions of cutoffs in
the cumulative distribution function and probability density function, H&S reduce
our data needs to little more than market shares and prices. However, Hortaçsu
and Syverson’s method of estimation lightens the burdensome data requirement
at the cost of imposing an extremely stringent set of institutional requirements. I
can confidently use the H&S approach to estimate consumer search costs in the
private passenger auto insurance industry of California for the following reasons:

1. Estimating search-costs distributions, using any method, presupposes that
costly search is important. There is abundant anecdotal evidence in the con-
tents of auto insurance advertisement campaigns that suggest that companies
are attempting to reduce the perceived search costs of consumers in the auto
insurance industry. Specific examples include one of Geico’s tag lines, “Fif-
teen minutes could save you 15% or more on car insurance.” Progressive uses
a less direct approach by giving consumers the option to receive quotes for the same product of three of its competitors. Also, in one of Progressive’s ads a couple complains about being tired of shopping around. Allstate tells consumers that “Breaking up is easy to do” because an Allstate agent will do it for them, “saving [them] that uncomfortable moment.” All of these at least hint to the fact that search costs matter a lot in the auto insurance industry.

2. H&S is applicable only to industries where there is a great degree of dispersion in sampling probability, but very limited product differentiation. The explosion in advertising spending in the auto insurance industry over the last twenty years points to the importance of advertising (and thus, sampling probability) in this industry. National advertising expenditure increased dramatically and consistently from about $25 million in 1995 to $1.7 billion in 2010. Additionally, we find significant support in the contents of advertisement campaigns for the assertion that auto insurance products are homogeneous: 21st Century’s punch line is “Same great coverage for less.” The fact that firms continue to spend liberally on advertisements that provide virtually no information, other than the name of the company, suggests that increasing a firm’s likelihood of being visited is very important.

Also, as we discuss below, insurers operate in a heavily regulated industry, where they cannot change their prices immediately at will. In such an environment where prices are not allowed to move freely, we would expect non-price competition to play a greater role than it otherwise would. Based on the product characteristics of auto insurance and the fact that retention rates are extremely high for insurance companies, I expect advertising to be the dominant mode of non-price competition in the auto insurance industry.

2 This advertisement speaks more to switching costs than to search costs. However, as will become apparent later, the data does not allow me to distinguish between search costs and switching costs. As such, the reader should interpret the term “search costs” as the sum of search costs and switching cost throughout this paper.

3 The model can also be applied in settings where product differentiation is important but all firms have an equal likelihood of been visited. Given the significant differences in market exposure between auto insurers, this stance is far less tenable here. However, I also estimate the model under this less plausible set of assumptions for comparison purposes.
3. The institutional features of the California auto insurance industry create an almost ideal setting for conducting this econometric analysis. In 1984, the California legislature passed the Robbins-McAlister Financial Responsibility Act. The Act made driving without written proof of insurance, by itself, illegal and punishable by fine and possibly suspension of driving privileges. This type of regulation makes demand for auto insurance inelastic at the extensive margin. As such, a firm can only increase its market share at the expense of another firm. This effectively eliminates the need to try to disentangle market-expanding advertising effects from that of combative advertising effects.

Furthermore, in an attempt to regulate supply in a market where demand had suddenly jumped dramatically, Proposition 103 was passed by California voters in 1988. Beginning in November of 1989, insurance companies were required to obtain prior approval for any premium rate changes before implementing them—a requirement that arguably constitutes the sort of “menu cost” that macroeconomists cite to justify assumptions of price rigidity. The Proposition also restricted the factors that could be used in determining a consumer’s rate; chief of which should be the driver’s driving safety record, number of miles driven annually, and number of years of driving experience.

This type of regulation restricts the ability of firms to segment the market according to consumer characteristics that are not observable to the researcher. Prices, thus depend only on the drivers’ risk, which I control for by using an industry-standard definition for output and prices. This adds validity to the assumption that drivers differ only based on their search costs, which I must assume in estimating the search-costs distribution using the H&S approach.

The importance of search costs, the spike in advertising spending, the price rigidity, and the overall regulatory framework in the auto insurance industry of California all combine to make this market ideally suited for estimating the consumer search costs distribution using H&S (2004). I add to Hortacsu and Syverson’s model by incorporating advertising spending into the sampling probability

\footnote{Here I ignore the fact that some drivers will respond to an increase in prices by choosing to drive without insurance. Though I do not have direct data on the extent of this problem, I do not believe it is pervasive enough to distort our results.}
function, which allows me to speak directly to the effectiveness of advertising in increasing the likelihood that a firm is visited.

The estimated model provides strong support for our major hypotheses. I confirm that consumers are more likely to visit a firm that advertised in the past and that this advertising effect is increasing over time. For example, the results suggest that an average-aged firm in 1997 was 3.5-times more likely to be visited if it spent $10,000 in 1996 rather than not advertising at all. Consistent with a priori expectations, I also find that overall search costs have been falling over time, especially between 2005 and 2010.

The next section (1.2) summarizes Hortaçu and Syverson’s model to allow for ease of exposition and comparison. I then discuss important elements of the data in section 1.3 and present the results in section 1.4.

### 1.2 A Theoretical Model of Sequential Search and Empirical Identification

In this section I present a model of sequential search, as put forward by Hortaçu and Syverson (2004). Note that the general model as discussed here is not unique to Hortaçu and Syverson. However, they make a significant contribution by expressing market shares and the price elasticities of demand in terms of the search costs distribution function and density function at the search cost cutoffs. In so doing, they provide a novel approach to recovering the distribution of consumer search costs. Since the main features of the model is taken directly from Hortaçu and Syverson (2004), I present the model here as succinctly as possible. To enable easy referencing and comparison I also maintain the same notations, wherever possible. My value-added is the application of their approach to the auto insurance industry, and the inclusion of advertising in the sampling probability. I make a special effort to highlight departures from the original model.

#### 1.2.1 Theory

$N$ firms each offer a single, possibly vertically differentiated, auto insurance product, in that all consumers have a common ranking of the products. A continuum
of drivers must all purchase auto insurance from one of these $N$ firms. Consumers have heterogeneous search costs, drawn from the cumulative distribution, $G(\cdot)$, with support $[0, \bar{c}]$. Consumers are homogeneous on all other dimensions, and thus have identical utility functions, ignoring search cost. Assuming linear utility functions, a consumer that purchases product $j$ derives utility,

$$u_j = \beta W_j - p_j + \xi_j,$$

where $W_j$ is the vector of non-price product attributes, including quality of claims servicing and the type of interaction between the firm and consumers. $\beta$ is the associated vector of coefficients on these attributes, $p_j$ is the price of product $j$, and $\xi_j$ is an unobserved component, common across all consumers. Note that the coefficient on price is normalized to $-1$, which results in utility being expressed in dollars per unit of sales. I use a dollar’s worth of expected loss in claims as our unit of sale. This will be discussed at length in section 1.3. Consumers are aware of the distribution of indirect utilities provided in the market. However, a consumer must engage in costly search to ascertain the exact benefit to be derived from a particular firm.

I maintain three important assumptions about the search process that were made by Hortaçsu and Syverson (2004). First, a consumer’s initial search is costless. This ensures that all drivers, regardless of their search costs, will visit at least one firm, and hence will purchase insurance. Second, to allow for computational tractability, consumers are assumed to search with replacement. The distorting effect of this assumption is reduced as the number of firms increase. Third, consumers can costlessly return to a previously visited firm. This allows consumers to compare the expected benefit from continued search against the highest utility observed on all previous visits rather than against that of the last visit.

Consumers continue to search as long as the expected benefit from visiting another firm is greater than the cost of search. Let $H(\cdot)$ be the cumulative distribution of indirect utilities, with support $[u, \overline{u}]$. A consumer with search cost $c$ will

---

5 Here auto insurance is viewed as an essential good, and so I do not include an outside good. I justify this assumption by restricting our attention to drivers, for whom auto insurance is a legal requirement. I ignore those who drive illegally without such insurance.
optimally continue to search until

\[ c = \int_{u^*}^{\bar{u}} (u - u^*) dH(u), \quad (1.2.2) \]

where \( u^* \) is the highest indirect utility observed on the previous searches. Note that indirect utilities that are less than \( u^* \) are ignored in this decision rule since the consumer can costlessly revisit the firm that offers \( u^* \). This decision rule is standard for sequential search models with the three assumptions outlined above.\(^6\)

In carrying out the remainder of their analysis, Hortaçsu and Syverson make use of the assumption that consumers know the empirical distribution of \( u \) with \( N \) firms. Consumers know the array of available indirect utilities but do not know which is associated with a given firm. Firms are indexed in ascending order based on the utility they offer to consumers, so that \( k > j \) iff \( u_k > u_j \). As such, firm 1 can be viewed as offering the worst deal, while firm \( N \) offers the best deal. Let \( \rho_j \) be the probability that firm \( j \) is sampled on each search. The optimal search rule above generates critical cutoff in search cost for each firm \( (j = 1, ..., N) \) of the form

\[ c_j = \sum_{k=j+1}^{N} \rho_k (u_k - u_j). \quad (1.2.3) \]

The cutoffs above can be interpreted in the following way: A driver who purchases auto insurance from firm \( j \) must have a search cost of at least \( c_j \). Consumers with search cost of exactly \( c_j \) would be indifferent between continuing to search and purchasing from firm \( j \). However, consumers with search costs less than \( c_j \) would continue to search until they find a firm that offers sufficiently higher utility than firm \( j \). It is obvious that these cutoffs are monotonically decreasing in the index of the firm. That is, \( c_1 > c_2 > ... > c_N \), while \( u_N > u_{N-1} > ... > u_1 \). Furthermore, \( c_N \) is necessarily zero, since regardless of the consumer’s search cost, there is no incentive to continue searching once he or she visits the firm offering the best deal, firm \( N \).

The novelty of Hortaçsu and Syverson’s approach is that they show that market shares can be written almost exclusively in terms of the cumulative distribution function at these cutoffs. Consider the firm offering the lowest utility, firm 1.

\(^6\) See for example, Weitzman (1997).
Consumers that purchase from firm 1 must have very high search costs, \( c_1 \) or higher. Also, firm 1 must be the first firm they visit, which occurs with probability \( \rho_1 \). Therefore, the fraction of consumers that purchase product 1 is

\[
s_1 = \rho_1(1 - G(c_1)) \quad (1.2.4)
\]

In calculating the market share for firm 2 we recognize that two sets of consumers purchase from it. Consumers with \( c \geq c_1 \geq c_2 \) will purchase from firm 2 if that is the first firm they visit. This occurs with probability \( \rho_2(1 - G(c_1)) \). Secondly, a fraction of consumers with \( c_2 \leq c \leq c_1 \) will purchase from firm 2. These consumers will continue to search until they find a firm that offers at least \( u_2 \). Such consumers will find firm 2 with probability \( \frac{\rho_2}{1 - \rho_1} \). Note that this probability is calculated under the assumption that consumers search with replacement. Consequently, the market share for firm 2 is

\[
s_2 = \rho_2(1 - G(c_1)) + \frac{\rho_2}{1 - \rho_1} (G(c_1) - G(c_2))
\]

\[
= \rho_2 \left[ 1 + \frac{\rho_1 G(c_1)}{1 - \rho_1} - \frac{G(c_2)}{1 - \rho_1} \right] \quad (1.2.5)
\]

Hortaçsu and Syverson show that the general market share for firms 3 through \( N \) can be analogously derived:

\[
s_j = \rho_j \left[ 1 + \frac{\rho_1 G(c_1)}{1 - \rho_1} - \frac{\rho_2 G(c_2)}{1 - \rho_1} - \frac{\rho_2 G(c_2)}{(1 - \rho_1)(1 - \rho_1 - \rho_2)} \right] \\
+ \rho_j \sum_{k=3}^{j-1} \frac{\rho_k G(c_k)}{(1 - \rho_1 - ... - \rho_{k-1})(1 - \rho_1 - ... - \rho_k)} \\
- \rho_j \frac{G(c_j)}{1 - \rho_1 - ... - \rho_{j-1}} \quad (1.2.6)
\]

The system of equations above concisely and comprehensively summarize demand.
On the supply side, firms choose prices to maximize their current profits

\[ \Pi_j = M s_j(p, W)(p_j - mc_j) \]  

(1.2.7)

\( M \) is the total market size, \( mc_j \) is the marginal cost of production, while \( p \) and \( W \) are respectively, the vectors of prices and firm attributes of other firms. At an interior solution, optimizing behavior requires that

\[ s_j(p, W) + (p_j - mc_j) \frac{\partial s_j(p, W)}{\partial p_j} = 0, \]  

(1.2.8)

where the derivatives of the market shares (with respect to their corresponding prices) are given below:

\[
\frac{\partial s_1}{\partial p_1} = -\rho_1 g(c_1) \sum_{k=2}^{N} \rho_k,
\frac{\partial s_2}{\partial p_2} = -\frac{\rho_1 \rho_2^2 g(c_1)}{1 - \rho_1} - \frac{\rho_2 g(c_2) \sum_{k=3}^{N} \rho_k}{1 - \rho_1},
\frac{\partial s_3}{\partial p_3} = -\frac{\rho_1 \rho_3^2 g(c_1)}{1 - \rho_1} - \frac{\rho_2 \rho_3^2 g(c_2)}{(1 - \rho_1)(1 - \rho_1 - \rho_2)} - \frac{\rho_3 g(c_3) \sum_{k=4}^{N} \rho_k}{1 - \rho_1 - \rho_2},
\]

and for \( j = 4, \ldots, N, \)

\[
\frac{\partial s_j}{\partial p_j} = -\frac{\rho_1 \rho_j^2 g(c_1)}{1 - \rho_1} - \frac{\rho_2 \rho_j^2 g(c_2)}{(1 - \rho_1)(1 - \rho_1 - \rho_2)} - \sum_{k=3}^{j-1} \frac{\rho_k \rho_j^2 g(c_k)}{(1 - \rho_1 - \ldots - \rho_{k-1})(1 - \rho_1 - \ldots - \rho_k)} - \frac{\rho_j (\sum_{k=j+1}^{N} \rho_k g(c_j))}{1 - \rho_1 - \ldots - \rho_{j-1}}.
\]

(1.2.9)

The form of the derivatives are intuitive. Increasing the price of product \( j \) reduces firm \( j \)’s market share through its effect on the search cost cutoffs, which in turn change because a higher \( p_j \) results in a lower \( u_j \). Consider the impact of

7 In Hortaçoşu and Syverson’s (2004) model, firms only choose current prices to maximize current profits. Ideally, with the appropriate data set, I would prefer to explore a dynamic model, where firms also choose advertising levels to influence their sampling probability. However, this is not feasible given my current data set. In chapter 2 I introduce this dynamic choice in the duopoly-equivalent model, which I solve numerically, rather than estimate.
increasing $p_j$. With a lower $u_j$, consumers that find lower-valued firms have less incentive to continue searching. From equation 1.2.3 we know that $c_i$ would fall for $i < j$. Thus, the market share of firm $j$ would fall because the probability that a consumer purchases a lower-valued product increases. This impact is accounted for by the first $j - 1$ terms above.

Secondly, from equation 1.2.3 we also know that increasing $p_j$ would cause $c_j$ to increase. This occurs because higher-valued products $k > j$, are now even more preferred to $j$ than before. Having found firm $j$, consumers are now more likely to continue their search. This effect where firm $j$ loses market share to higher-valued firms is embodied in the $j^{th}$ term above.

1.2.2 Identification and Estimation Strategy

Following H&S, I estimate the model under two extreme sets of assumptions. These assumptions determine the exact form of the sampling probability function, cutoff equations, and market share equations. They also pin down the ranking of firms, which is critical to the results. In making these assumptions I allow firms to vary in either non-price attributes or sampling probabilities, but not both. I concede that neither of these extreme sets of assumptions is likely to actually hold. However, given the available data, I must shut down one possible type of heterogeneity, in order to identify the other’s effect.

1.2.2.1 Heterogeneous Firms with Equal Sampling

In one version of the model I assume that all firms are equally likely to be sampled, while allowing them to be heterogeneous in non-price attributes. Here variation in market shares are explained only by variation in the quality of the products offered. These assumptions dictate that firms be ranked based on market shares.

The system of market share equations (1.2.6) forms the basis for identifying the search cost distribution. Recall that we have a system of $N$ linear equations in $G(c_1), ..., G(c_N)$. As Hortaçsu and Syverson point out, if the sampling probabilities are known or assumed, we can solve for $G(c_1), ..., G(c_{N-1})$, with $G(c_N) = 0$. Market

---

8 Market shares also depend on prices, but these are in turn functions of the non-price attributes of the firms. Under the second set of assumptions I make the statement that market share are determined only by sampling probabilities based on similar reasoning.
shares identify the cumulative distribution function (CDF) at $N - 1$ cutoffs, and in so doing, identify the relative positions of these cutoffs. However in the case of heterogeneous firms, to recover the absolute levels of the cutoffs we also need information on the probability density function (PDF). H&S show that we can obtain this information by utilizing the optimal pricing behavior of firms, which is encapsulated in the first order conditions. I fit equations 1.2.6 and 1.2.8 using non-linear least squares for each year.

Having estimated the distribution and density function at the cutoffs, $G(c_1), \ldots, G(c_N)$, $g(c_1), \ldots, g(c_N)$, we are able to recover the $N - 1$ cutoffs. Following H&S, I use the trapezoidal method to approximate the probability that search costs fall between two adjacent cutoffs. Thus,

$$G(c_{j-1}) - G(c_j) = 0.5[g(c_{j-1} + c_j)](c_{j-1} - c_j).$$

(1.2.10)

This approximation allows us to non-parametrically recover the distribution function at the cutoffs, $c_1, \ldots, c_{N-1}$, and using any non-decreasing function we can trace out the entire distribution. Note that $g(c_N = 0)$ is not identified but its value does not significantly influence the estimated search costs in any way.

### 1.2.2.2 Unequal Sampling with Homogeneous Firms

In the second version of the model I reverse the assumptions; firms are now assumed to be homogeneous in non-price attributes, but differ in their likelihood of being visited. In this set-up, differences in market shares are explained solely by differences in sampling probabilities, which I specify to be a function of the age of the firm and its past advertising expenditures. Given that all firms offer an identical product, firms are ranked based on prices. I parameterize both the sampling probabilities and the search costs distribution function to estimate the distribution at its cutoffs. I estimate the market share equations using the following parametric form for the sampling probabilities:

$$\rho_j = \beta \frac{x_j^\alpha}{\sum_{k=1}^{N} x_k^\alpha} + (1 - \beta) \frac{z_j^\gamma}{\sum_{k=1}^{N} z_k^\gamma}.$$  

(1.2.11)

$x_j$ is the total one-period lagged advertising expenditure by firm $j$ and $z_j$ is
the firm’s age. $\alpha$, $\beta$, and $\gamma$ are parameters to be estimated. In the auto insurance industry I expect advertising, ‘informative’ or ‘persuasive’, to increase a firm’s visibility and hence its sampling probability. Advertising is especially important for firms that use the direct writing channel, since they must attract consumers directly, rather than through independent agents or brokers. I use lagged advertising expenditure in an attempt to minimize possible endogeneity problems that may arise from using contemporaneous advertising, the choice of which may be influenced by current demand shocks. The use of national advertising expenditure and not California-specific advertising should also mitigate the endogeneity problem. I recognize however, that this does not completely solve the problem and that the coefficient on advertising may still be biased upwards.

Firms that rely heavily on independent agents optimally do not advertise much, but may still have large market shares. I do not have data on the share of customers that use direct channels versus agency, and so must complement advertising expenditure with some other proxy of market exposure. I use the age of the firm, which also serves to capture some of the effects of the notoriously high retention rates of auto insurance companies. Ideally, I would like to account directly for different marketing channels. However, attempts to acquire reliable historical data on marketing channel proved futile. $\alpha$ and $\gamma$ enter as exponents to allow for non-linearity in the respective relationships. This is especially important for the advertising portion of the function given that many previous studies have found evidence of diminishing marginal returns in the advertising effects. $\beta$ provides additional flexibility by controlling for the fraction of the sampling probability that is influenced by age relative to advertising.

In this version of the model I must also specify the functional form of the search cost distribution function. I estimated the model and compared the results with many possible distributions, including log-normal, gamma, exponential, and weibull. I settled on using the Rayleigh distribution for a few reasons. As a one-parameter distribution, the Rayleigh distribution will ensure that the estimated search costs distributions do not overlap. This will give us a better picture of the movement of the overall distribution over time, which is one of my main interests. Use of the Rayleigh distribution does not sacrifice much in terms of goodness of fit. Also, the log-normal distribution, our a priori choice, would allow us to speak to
changes in the tails of the distribution, but it produces search costs distributions with unrealistically tight support.

Note that in this version of the model the cutoffs can be identified without the information from the firms’ first order conditions for prices. This is the case because firms are assumed to be homogeneous in all non-price attributes, and as such, the cutoff equations (1.2.3) reduce to

\[ c_j = \sum_{k=j+1}^{N} \rho_k (p_j - p_k). \]  

(1.2.12)

1.3 Data

1.3.1 Data description

An observation is a group in a year. I choose group-level over company-level data because consumers are unlikely to be able to differentiate sufficiently between similarly named companies that are members of the same group. I recognize that this choice requires that I aggregate up to the group-level, which may distort the differences in prices offered by companies. Having estimated both versions of the model using both company-level and group-level data, I confirm that group-level estimates are much more compelling.

I have data on written and earned premium, incurred losses, defense and cost containment expenses, and commission and brokerage costs for private passenger auto liability and physical damage insurance. These data were obtained from A. M. Best Ltd. Incurred losses will be our measure of quantity, whereas price will be the earned premium per dollar of incurred loss for a given firm. Earned premium is used in the definition of price rather than written premium because the timing of earned premium is more consistent with incurred losses from a practical and accounting standpoint. I discuss the choice of quantity and unit price in more depth below. I use the sum of defense and cost containment expenses and commission and brokerage costs as a ratio of incurred losses to proxy for the constant marginal cost for each firm.

---

9 I use the terms firm and group interchangeably below.

10 In short, auto liability insurance covers the losses inflicted by an at-fault insured driver on other drivers, whereas physical damage insurance covers losses to the insured’s vehicle.
Advertising expenditure data is obtained from Kantar Media of the Kantar Group. I have advertising data at the media-market level broken down by media type. However, in all the analyses, I make use only of national advertising expenditure data. I use national advertising, rather than California-specific advertising, for two reasons. First, national advertising, especially by the top-four advertisers, dwarfs state-specific advertising, which would mean I would miss much of the sampling probability effects if I were to focus exclusively on California-specific advertising. Second, in this paper I treat advertising as exogenous, which more closely matches national advertising than California-specific advertising. Our age variable represents how long a company has been licensed to operate in California. I obtained this information using an online search tool on the California Department of Insurance’s website. In cases where I was unable to find an exact match, I proxy the age with how long the company has been incorporated in any state or the age of the company’s parent.

I have a total of 1,209 observations covering the sixteen-year period, 1995–2010. I estimate the search cost distribution for each year, independently. By doing so, I implicitly assume that all consumers engage in search on a yearly basis. This assumption is not unreasonable given that the maximum length of an auto insurance contract is twelve months, with many companies offering only six-month contracts. In fact, consumers are usually able to switch insurance providers at any point during the life of the contract without facing a financial penalty.

Having said that, I am fully aware that some consumers’ decision to re-sign with their present provider is so automatic that for them, no real search occurs on a yearly basis. This is not overly worrisome because our goal here is to estimate the distribution of perceived, not actual search cost. A consumer who believes that the cost of switching between firms is sufficiently high should optimally re-sign with their current provider, which in the context of the model simply means that the consumer purchases after the first search.\footnote{As noted in the introductory section (1.1), anecdotal evidence suggest that a large fraction of advertising in this industry attempt to convince consumers that switching or searching is much easier than consumers believe.} Also, I remind the reader that our data set does not allow us to distinguish between search costs and switching cost, and therefore I am comfortable estimating the model for each year separately.
1.3.2 Defining the Unit of Sale

Market share is a critically important variable in our analysis. As such, it is paramount that it is defined precisely, and measured accurately. I could model demand as the aggregate result of a discrete choice, where drivers simply choose whether or not to insure with a given firm. This would lead us to measure market shares as the ratio of the number of drivers or vehicles that are insured by a firm relative to the total number insured in the market. This definition is feasible for us since I have data on the number of vehicles insured at the firm level. However, using this definition requires that I ignore the large degree of heterogeneity that exists in the amount and type of risk that consumers insure against. To adequately control for this heterogeneity I would require substantially more data than is available to me; at a minimum, I would need the distribution of drivers across firms based on the major insurance risk classification characteristics, such as age and gender.

Extensive search has revealed that it is unlikely that I will be able to access even this base-level data, and so I must settle for far more aggregated data. It is important to note however, that in the context of the H&S model, having highly aggregated data is not overly problematic. In fact, the type of data that is generally available in the auto insurance industry, such as firm-level premium, is perfectly suited for an application of Hortacsu and Syverson’s modeling strategy. But to successfully use such aggregated data still necessitates that I define the unit of sale in such a way that risk is standardized across firms. To achieve this, I utilize a definition that is somewhat commonplace in the insurance literature, but may be nonstandard in the wider industrial organization field.

I consider the substantive output of the auto insurance firms to be the services of pooling and bearing risks, and making claim payouts in the event of losses. As such, I define the total output for a firm as its total expected losses that may result from claims. This appears to be an appropriate and noncontentious measure.

\[\text{In the insurance industry this is termed as the number of exposures. Note that using this definition would significantly reduce the size of our data set as our sample period would shrink to five years.}\]

\[\text{See, for example, Tennyson and Posey (1998) and Berger, Cummins, and Weiss (1997).}\]

\[\text{Other services provided by the industry include, but are not limited to, intermediation and risk reduction. See Berger et al (1997) for a more comprehensive discussion of these and other industry services.}\]
since it reflects the amount of risk that the firm expects to cover. What is however more contentious, is using total incurred losses as a proxy for total expected losses. I follow this practice by defining our basic unit of output as a dollar of incurred loss and the unit price as the earned premium per dollar of incurred loss. Note that given these definitions, our price can be interpreted as a mark-up over the actuarially-fair price.

Berger et al (1997) cite two limitations of using actual incurred losses as the measure of output. First, this measure does not account for differences in the type and quality of services provided by insurers. Consequently, the output of firms that offer more or superior services are likely to be understated. I am not unduly perturbed by this limitation since I attribute variations in premium per dollar of expected losses to differences in service levels, at least partially in one version of our estimates.

Second, they argue that using incurred losses to measure insurance output ignores the output qualities of risk assessment, loss control, and risk management. Hence, a firm that is more successful in reducing its losses by reducing moral hazard effects or a firm that is more accurate in assessing risk, is more likely to have a lower output. This does not strike me as an overly damning critique either, because optimizing behavior on the part of both consumers and firms dictates that such mismeasurement should be small. Consider the problem of erroneous risk assessment. A firm that overestimates risk will charge a relatively high price, and in equilibrium such a firm should have a relatively small market share. Conversely, a firm that consistently underestimates risk should earn sustained losses, which will force the firm to improve its underwriting technology or exit the market. Additionally, firms are able to duplicate other firms’ underwriting practices, and so any actuarial advantage that does exist should dissipate quickly. In either case, an actuarially flawed firm should have a small market share, correctly reflecting its “low-quality” product.

A concern, however, still exists that actual losses may be significantly different from expected losses, even for a firms that correctly assess risk. This concern is likely to be greater for firms with very small market shares, some of which may actually be exiting the market. To reduce the likelihood of this occurring, I remove observations with extremely low market share and unusually high (or low) prices.
Finally, defining the unit of sale as a dollar’s worth of expected loss appears natural from the perspective of the firms, but perhaps less so for consumers. Under the following timing assumption, I suggest that our “unit of sale” definition is also appropriate for consumers. Consumers choose their desired auto insurance package, including deductible and bodily liability limit, before they engage in search. The consumer then visits firms sequentially, observe the firm’s quoted premium for his or her desired package along with other features of the firm’s product offering. Having done so, the consumer decides to purchase insurance with that firm or to visit another. Under this timing assumption consumers compare quoted premium of different firms for an identical insurance product, and hence the same total expected loss. It is thus reasonable to model consumers as making decisions based on prices that are quoted on a per dollar of expected losses basis.

1.3.3 Defining the Market

A more standard industrial organization question relates to market definition. Specifically, do firms view a dollar’s worth of expected loss differently if originating from a high risk driver as opposed to a low risk driver, and do firms specialize in different risk types? If they do, then I would err in treating them jointly in a single market, and our estimate of search cost may also include a risk component, which is absent in the application of H&S.

Adding some amount of credence to the argument above, is our observation that there are substantial differences between the prices of companies within the same group, which arguably allocate drivers to their subsidiary companies based on risk type. This may suggest the market for low-risk drivers is sufficiently different from that of high-risk drivers, such that there is likely to be significant differences in price per dollar of incurred loss between the two categories. My answer to these concerns is two-fold. First, our definitions of quantity as a dollar’s worth of expected loss and our price as the total premium as a ratio of total incurred losses implicitly assumes that firms seek to maximize expected profits, and that variance of losses are unimportant. Therefore, a consumer who has higher expected losses would simply be required to pay a scaled up price of what a low-risk consumer would pay. Second, the differences we observe in price per unit of expected loss
may simply be a reflection of differences in marginal costs, which I assume to be constant over output.

While a few firms choose to focus only on auto insurance, most insurance companies offer a wide portfolio of insurance products; including auto, renter, homeowner, and life. There is the potential for us to use this additional dimension of optimizing behavior to better identify our model. However, when information on these additional products is unavailable, as it is here, the interaction among their markets may confound our estimation strategy. This confounding element is likely to arise because consumers usually receive significant discounts by purchasing multiple insurance products from a single firm, in which case consumers are buying insurance as a composite product. Therefore, I may wrongly interpret an individual buying high priced auto insurance bundled with low priced homeowners insurance either as the consumer having high search cost or the auto insurance firm having above-average non-price attributes. The available data does not allow me to control for this possibility, and so the results may therefore overstate the absolute levels of consumer search costs. However, the estimated search costs levels suggest that if this type of error exists, it is small.

1.4 Estimation Results

1.4.1 Version 1: Homogeneous Firms with Unequal Sampling

1.4.1.1 Consumer Search Cost Distribution

Based on the institutional features of the auto insurance industry, along with other anecdotal evidence, I believe that unequal sampling is more important in explaining variations in market shares than does heterogeneity in non-price product attributes. Therefore, I first present the results of the model under the assumption that firms differ based on the likelihood of being visited by a randomly chosen driver. It is useful to bear in mind that in this version firms have the same non-price attributes, and thus are ranked from best to worst in ascending order of prices.

The main output from the estimation process is a series of consumer search cost distributions (CDFs) from 1996 to 2010. I report key statistics for each year’s
Table 1.1: Estimates for Consumer Search Costs Distributions (Unequal Sampling)

<table>
<thead>
<tr>
<th>Year</th>
<th>$R^2$</th>
<th>Coefficient</th>
<th>(Std. Error)</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Deviation</th>
<th>IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>0.345</td>
<td>0.134</td>
<td>(0.070)</td>
<td>0.168</td>
<td>0.158</td>
<td>0.088</td>
<td>0.121</td>
</tr>
<tr>
<td>1997</td>
<td>0.452</td>
<td>0.207</td>
<td>(0.039)</td>
<td>0.259</td>
<td>0.243</td>
<td>0.135</td>
<td>0.187</td>
</tr>
<tr>
<td>1998</td>
<td>0.435</td>
<td>0.416</td>
<td>(0.125)</td>
<td>0.521</td>
<td>0.489</td>
<td>0.272</td>
<td>0.377</td>
</tr>
<tr>
<td>1999</td>
<td>0.451</td>
<td>0.186</td>
<td>(0.063)</td>
<td>0.233</td>
<td>0.219</td>
<td>0.122</td>
<td>0.168</td>
</tr>
<tr>
<td>2000</td>
<td>0.442</td>
<td>0.143</td>
<td>(0.063)</td>
<td>0.179</td>
<td>0.168</td>
<td>0.094</td>
<td>0.129</td>
</tr>
<tr>
<td>2001</td>
<td>0.460</td>
<td>0.313</td>
<td>(0.145)</td>
<td>0.392</td>
<td>0.368</td>
<td>0.205</td>
<td>0.284</td>
</tr>
<tr>
<td>2002</td>
<td>0.439</td>
<td>0.939</td>
<td>(1.335)</td>
<td>1.177</td>
<td>1.105</td>
<td>0.615</td>
<td>0.851</td>
</tr>
<tr>
<td>2003</td>
<td>0.386</td>
<td>1.481</td>
<td>(4.840)</td>
<td>1.856</td>
<td>1.744</td>
<td>0.970</td>
<td>1.343</td>
</tr>
<tr>
<td>2004</td>
<td>0.396</td>
<td>4.013</td>
<td>(70.708)</td>
<td>5.029</td>
<td>4.724</td>
<td>2.629</td>
<td>3.638</td>
</tr>
<tr>
<td>2005</td>
<td>0.492</td>
<td>0.493</td>
<td>(0.295)</td>
<td>0.618</td>
<td>0.580</td>
<td>0.323</td>
<td>0.447</td>
</tr>
<tr>
<td>2006</td>
<td>0.581</td>
<td>0.434</td>
<td>(0.238)</td>
<td>0.544</td>
<td>0.511</td>
<td>0.284</td>
<td>0.393</td>
</tr>
<tr>
<td>2007</td>
<td>0.533</td>
<td>0.300</td>
<td>(0.188)</td>
<td>0.376</td>
<td>0.353</td>
<td>0.196</td>
<td>0.272</td>
</tr>
<tr>
<td>2008</td>
<td>0.568</td>
<td>0.264</td>
<td>(0.140)</td>
<td>0.331</td>
<td>0.311</td>
<td>0.173</td>
<td>0.240</td>
</tr>
<tr>
<td>2009</td>
<td>0.491</td>
<td>0.219</td>
<td>(0.075)</td>
<td>0.275</td>
<td>0.258</td>
<td>0.144</td>
<td>0.199</td>
</tr>
<tr>
<td>2010</td>
<td>0.513</td>
<td>0.201</td>
<td>(0.067)</td>
<td>0.252</td>
<td>0.236</td>
<td>0.132</td>
<td>0.182</td>
</tr>
<tr>
<td>Average</td>
<td>0.493</td>
<td>0.289</td>
<td>–</td>
<td>0.362</td>
<td>0.340</td>
<td>0.189</td>
<td>0.262</td>
</tr>
</tbody>
</table>

*(a) Scale parameter for the Rayleigh distribution
(b) Excludes values from years 2002-2004*

estimated distribution in table 1.1. Also, to avoid cluster I show some of the graphs of the CDFs in two separate figures. The model does an adequate job in explaining the variation in market shares in spite of being so parsimoniously parameterized. Although our average $R^2$ of 0.49 pales in comparison to $R^2$ values above 0.9, as reported for the mutual funds application in H&S, the model does sufficiently well to justify being applied to the auto insurance industry. I also recognize that the stark assumption of a single differentiating variable (sampling probability) necessarily limits the explanatory power of the model.

Except for years 2002-2004, the scale parameter in the Rayleigh distribution is statistically significant at roughly the 5% level. I am not able to precisely estimate this parameter for years 2002-2004 even under different distributional assumptions. This may be explained by a regime change that started in 2002 and reversed in
2005. I exclude the estimated distributions for these years for the remainder of the discussion.

Before I discuss the differences between the estimated distributions I note that the absolute values of the estimated search costs are quite reasonable. For example, the median search costs in 2010 is 0.24; that is, 24% of the expected loss that the consumer is insuring. Take for instance an individual who is offered a premium of $1,000 for 2010. Based on the mean 2010 price of $1.69, this consumer would have an expected loss of $592. Such a consumer would have a search cost of $142 since the estimated search cost is expressed as a ratio of expected losses.

Though search costs of $142 may appear high, this value is consistent with findings in previous work. Honka (2010), for example, using a different estimation procedure, finds the sum of search costs and switching costs ranges from $130 to $195 in the US auto insurance industry in 2006 and 2007. Again, recall that the available data and our chosen estimation procedure does not allow us to distinguish between search costs and switching costs.

Figures 1.1 and 1.2 show the estimated search cost distributions for the 1996-2001 period and the 2005-2010 period, respectively. Though I do not find a clear trend over the entire sample period, there is an unmistakable downward trend in search costs for the later period. Here I am able to roughly compare the distributions based on stochastic dominance mainly due to our choice of a one-parameter distribution. However, the results from using a lognormal distributional assumption does not provide any strong arguments for differential changes in the tails as oppose to the center of the distributions over time.

Our a priori expectation is that search costs have been falling over the last decade as more consumers and firms use the Internet to carry out transactions, which is usually much quicker and less stressful, at least for the younger generation. Additionally, we would expect that this change would be stronger in the latter half of the sample period as e-commerce in auto insurance matures and become more pervasive. This expectation is supported by the strong downward trend in the later years, coupled with the lack of a clear trend for the early portion of the sample period.

This initial period of little or no change in consumer search costs (or perhaps even a slight increase) followed by a significant reduction is consistent with the
findings of Brown and Goolsbee (2002). They found that price dispersion actually increased before decreasing in the term life insurance industry of U.S. after the introduction of Internet comparison shopping sites. All the statistics for the distribution are calculated based on the estimated $\sigma$, and as such does not provide any additional information about the the evolution of search costs. I report additional statistics of the estimated distribution, such as inter-quartile range, for the curious reader.

1.4.1.2 Sampling Probability

One of the primary goals of this research is to identify the impact of advertising on sampling probabilities. This in turn will give us a sense of how well differences in sampling probabilities are able to explain the large differences that we observe in market shares between the top firms and the rest. In estimating the search cost distributions under the assumption of homogeneous firms, I obtain estimates of the
coefficients on age and advertising in the sampling probability function, equation 1.2.11. Table 1.2 provides summary results for the estimates of the sampling probability function. Age is statistically significant at the 1% level in all years, with an average coefficient of 2.85. This average is slightly higher than coefficient on age of 2.6 in the mutual funds application of H&S. I use a different sampling probability function and so will not try to infer anything more about the importance of age in drawing potential consumers in the mutual funds industry relative to that of the auto insurance industry.

Given that the coefficient on age is greater than one, it is very likely that without an additional means to attract consumers to obtain a quote that the industry would become increasingly concentrated. This hints at a possible explanation for why the market is becoming less concentrated: Firms with low market shares can now use advertising to attract additional consumers from even older, bigger firms. With this possibility, advertising takes on a procompetitive look in the auto
Table 1.2: Estimates for Sampling Probability Functions, $\rho$

<table>
<thead>
<tr>
<th>Year</th>
<th>$\alpha$ (Age) Coefficient</th>
<th>(Std. Error)</th>
<th>$\gamma$ (Advertising) Coefficient</th>
<th>(Std. Error)</th>
<th>$\beta$ Coefficient</th>
<th>(Std. Error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>1.904</td>
<td>(0.450)</td>
<td>43.802</td>
<td>(0.000)</td>
<td>0.900</td>
<td>(0.167)</td>
</tr>
<tr>
<td>1997</td>
<td>2.940</td>
<td>(0.522)</td>
<td>0.143</td>
<td>(0.188)</td>
<td>0.776</td>
<td>(0.065)</td>
</tr>
<tr>
<td>1998</td>
<td>2.474</td>
<td>(0.577)</td>
<td>0.433</td>
<td>(0.250)</td>
<td>0.672</td>
<td>(0.084)</td>
</tr>
<tr>
<td>1999</td>
<td>2.857</td>
<td>(0.701)</td>
<td>0.324</td>
<td>(0.157)</td>
<td>0.671</td>
<td>(0.097)</td>
</tr>
<tr>
<td>2000</td>
<td>2.401</td>
<td>(0.674)</td>
<td>0.182</td>
<td>(0.132)</td>
<td>0.703</td>
<td>(0.138)</td>
</tr>
<tr>
<td>2001</td>
<td>2.558</td>
<td>(0.899)</td>
<td>0.197</td>
<td>(0.105)</td>
<td>0.576</td>
<td>(0.138)</td>
</tr>
<tr>
<td>2002</td>
<td>2.343</td>
<td>(0.715)</td>
<td>0.246</td>
<td>(0.108)</td>
<td>0.593</td>
<td>(0.113)</td>
</tr>
<tr>
<td>2003</td>
<td>2.519</td>
<td>(0.886)</td>
<td>0.215</td>
<td>(0.107)</td>
<td>0.545</td>
<td>(0.107)</td>
</tr>
<tr>
<td>2004</td>
<td>2.394</td>
<td>(0.884)</td>
<td>0.170</td>
<td>(0.082)</td>
<td>0.537</td>
<td>(0.111)</td>
</tr>
<tr>
<td>2005</td>
<td>2.610</td>
<td>(1.002)</td>
<td>0.235</td>
<td>(0.077)</td>
<td>0.475</td>
<td>(0.114)</td>
</tr>
<tr>
<td>2006</td>
<td>3.439</td>
<td>(0.928)</td>
<td>0.344</td>
<td>(0.101)</td>
<td>0.506</td>
<td>(0.087)</td>
</tr>
<tr>
<td>2007</td>
<td>2.696</td>
<td>(0.812)</td>
<td>0.293</td>
<td>(0.110)</td>
<td>0.518</td>
<td>(0.110)</td>
</tr>
<tr>
<td>2008</td>
<td>3.164</td>
<td>(1.086)</td>
<td>0.282</td>
<td>(0.092)</td>
<td>0.484</td>
<td>(0.112)</td>
</tr>
<tr>
<td>2009</td>
<td>3.248</td>
<td>(1.108)</td>
<td>0.294</td>
<td>(0.132)</td>
<td>0.576</td>
<td>(0.123)</td>
</tr>
<tr>
<td>2010</td>
<td>2.942</td>
<td>(0.959)</td>
<td>0.222</td>
<td>(0.108)</td>
<td>0.501</td>
<td>(0.112)</td>
</tr>
<tr>
<td>Averagea</td>
<td>2.848</td>
<td>–</td>
<td>0.268</td>
<td>–</td>
<td>0.587</td>
<td>–</td>
</tr>
</tbody>
</table>

(a) Excludes values from years 2002-2004
(b) $\rho_j = \beta \frac{x^\alpha_j}{\sum x^\alpha_k} + (1 - \beta) \frac{z^\gamma_j}{\sum z^\gamma_k}$, $x$ - age, $z$ - lagged advertising spending

The parameter $\gamma$ partly captures the effects of advertising on sampling probability. $\gamma$ is statistically significant at the 5% level for all years except 1997. The coefficients are found to be positive and less than one for all years, except 1996. Given our sampling probability function (equation 1.2.11), such coefficients indicate that firms with higher advertising levels are more likely to be sampled, but this positive effect decreases with advertising. This is consistent with the traditional assumption that advertising exhibits diminishing marginal returns, at least above some threshold. Given this finding, it is likely that the escalation in advertising

---

15 Such a large value for $\gamma$ is understandable in the first year given that very few firms advertised at this point, and so a small amount of advertise is likely to be very effective. This is easily explained in the context of advertising congestion. Note however that given that $\beta$ is 0.9, advertising only accounted for 10% of the sampling probability, regardless of how much a firm advertised.
spending over the sample period is a feature of a standard prisoners’ dilemma, where firms undertake costly advertising campaigns simply in an attempt to maintain their level of market exposure, rather than an attempt to increase it. Also, the coefficients vary considerably over the years, which is in keeping with the stochastic nature of advertising effects.

Table 1.3: Implied Sampling Probability Distributions (1996-2010)

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean</th>
<th>Median</th>
<th>Maximum</th>
<th>Std. Deviation</th>
<th>IQR</th>
<th>75th/25th</th>
<th>Max/Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>0.012</td>
<td>0.005</td>
<td>0.131</td>
<td>0.018</td>
<td>0.012</td>
<td>7.5</td>
<td>24.3</td>
</tr>
<tr>
<td>1997</td>
<td>0.011</td>
<td>0.002</td>
<td>0.076</td>
<td>0.019</td>
<td>0.010</td>
<td>19.9</td>
<td>31.2</td>
</tr>
<tr>
<td>1998</td>
<td>0.012</td>
<td>0.003</td>
<td>0.096</td>
<td>0.020</td>
<td>0.011</td>
<td>18.6</td>
<td>27.7</td>
</tr>
<tr>
<td>1999</td>
<td>0.015</td>
<td>0.003</td>
<td>0.075</td>
<td>0.019</td>
<td>0.013</td>
<td>34.3</td>
<td>26.0</td>
</tr>
<tr>
<td>2000</td>
<td>0.017</td>
<td>0.009</td>
<td>0.072</td>
<td>0.019</td>
<td>0.014</td>
<td>6.5</td>
<td>8.2</td>
</tr>
<tr>
<td>2001</td>
<td>0.014</td>
<td>0.006</td>
<td>0.077</td>
<td>0.021</td>
<td>0.023</td>
<td>14.0</td>
<td>13.0</td>
</tr>
<tr>
<td>2002</td>
<td>0.013</td>
<td>0.005</td>
<td>0.074</td>
<td>0.020</td>
<td>0.012</td>
<td>8.3</td>
<td>15.4</td>
</tr>
<tr>
<td>2003</td>
<td>0.013</td>
<td>0.003</td>
<td>0.063</td>
<td>0.018</td>
<td>0.014</td>
<td>15.4</td>
<td>19.0</td>
</tr>
<tr>
<td>2004</td>
<td>0.013</td>
<td>0.005</td>
<td>0.070</td>
<td>0.018</td>
<td>0.011</td>
<td>9.6</td>
<td>15.2</td>
</tr>
<tr>
<td>2005</td>
<td>0.014</td>
<td>0.004</td>
<td>0.074</td>
<td>0.019</td>
<td>0.017</td>
<td>16.8</td>
<td>18.3</td>
</tr>
<tr>
<td>2006</td>
<td>0.014</td>
<td>0.003</td>
<td>0.086</td>
<td>0.022</td>
<td>0.016</td>
<td>36.3</td>
<td>28.8</td>
</tr>
<tr>
<td>2007</td>
<td>0.014</td>
<td>0.003</td>
<td>0.080</td>
<td>0.019</td>
<td>0.016</td>
<td>14.5</td>
<td>23.3</td>
</tr>
<tr>
<td>2008</td>
<td>0.014</td>
<td>0.004</td>
<td>0.075</td>
<td>0.020</td>
<td>0.017</td>
<td>41.6</td>
<td>20.4</td>
</tr>
<tr>
<td>2009</td>
<td>0.015</td>
<td>0.005</td>
<td>0.070</td>
<td>0.020</td>
<td>0.018</td>
<td>41.3</td>
<td>12.7</td>
</tr>
<tr>
<td>2010</td>
<td>0.014</td>
<td>0.003</td>
<td>0.068</td>
<td>0.020</td>
<td>0.023</td>
<td>39.4</td>
<td>19.5</td>
</tr>
<tr>
<td>Average a</td>
<td>0.014</td>
<td>0.004</td>
<td>0.077</td>
<td>0.020</td>
<td>0.016</td>
<td>25.7</td>
<td>20.8</td>
</tr>
</tbody>
</table>

(a) Excludes values from years 2002-2004

β is included in the sampling probability function to directly account for the importance of age relative to advertising. It can be interpreted as the cap on a firm’s sampling probability if it chooses not to advertise. β is statistically significant at the 1% level and falls over the entire sample period, roughly. It ranges from a high of 0.9 in 1996 to a low of 0.48 in 2005, and end slightly higher in 2010 at 0.5. This clearly suggests that during the sample period, relative to age, advertising has become significantly more important in determining which firms are visited by consumers.

To get a clearer picture of how the sampling probability function is changing
over time, we need to analyze how the changes in $\alpha$, $\gamma$, and $\beta$ jointly affect sampling. Consider firms that do not advertise. A consumer is 39% more likely to visit a firm that is five years older than the average-aged firm in 2010 than an average-aged firm. This is substantial given that the oldest firm is over 50 years older than the average-aged firm in 2010, and it highlights the importance of the age of a firm in attracting consumers. However, note that the same firm would have been 51% more likely to be visited than the average-aged firm in 1997, highlighting the rise of advertising as a drawing force.

To get a better idea of the importance of advertising in shifting sampling probabilities, let us consider an average-aged firm (33) in 1997 that did not advertise in 1996. Advertising expenditure of $10,000 in 1996 would have caused this firm’s sampling probability to be 3.5-times higher. In 2010 the corresponding effect would have been 1.6 times higher. This reflects the fact that though advertising is more important in the latter half of the sample period; given that total advertising is higher, it now requires higher levels of advertising to be effective than in the past.

![Figure 1.3: Sampling Probability Distribution](image)
Table 1.3 reports some summary statistics for the implied sampling probabilities based on our estimates above. The means are simply a function of the number of firms that are active in a year. Note that the mean is significantly higher than the median in all years. This typically occurs in distributions that are positively skewed. This is evident in figure 1.3 which illustrates that the vast majority of firms are sampled less than 1% of the time both at the beginning and end of the sample period. Alongside these rarely sampled firms there is a select group of firms that are sampled close to 10% of the time. These differences in sampling probability are substantial and go a long way in explaining the differences in market shares across firms. Over the period the most popular firm is 20 times more likely to be sampled by a given consumer than the median firm.

Our expectation was that the dispersion in sampling probabilities would fall over time given that advertising can now be effectively used to attract consumers and that smaller firms have a greater incentive to advertise than do larger firms. However, there is no clear trend in the changes in sampling dispersion. This may suggest that the dramatic increases in advertising in the industry as a whole has created a minimum advertising requirement that is simply too high for the majority of firms to afford. This thesis warrants further exploration, but it is beyond the scope of this paper.

1.4.2 Version 2: Heterogeneous Firms with Equal Sampling

Recall that I estimate two separate versions of the model: The first allows firms to differ in sampling probability, while the second allows firms to differ in non-price attributes. In each case I allow for only one source of heterogeneity across firms. Though my focus is on the first version, which speaks to the importance of advertising, I am nonetheless interested in the degree of product differentiation that would be required to rationalize the observed data in the context of H&S. I am also interesting in the differences that show up in the estimated search costs distributions between versions.

Figure 1.4 shows the estimated search cost CDFs for five select years. There are two important differences that show up in the estimated search costs across
models. Firstly, the more striking difference between the versions is the differences in the absolute values of the estimated search costs. Median search costs estimates are about 5 to 10 times lower in the equal sampling version of the model than the unequal sampling version. This is not surprising, however. In a model where I impose equal sampling, the best firm necessarily has the highest market share. Therefore, the majority of variation in market shares is explained simply by ranking firms based on market share. A simple explanation thus presents itself: Consumers have very low search costs and therefore will continue to search until they find the better firms, allowing ‘better’ firms to have larger market shares.

The second difference in the estimates is that the non-parametrically identified CDFs in the equal-sampling version suggest consistent small incremental increases in search costs. This conflicts with estimated changes in the unequal-sampling version, where search costs were falling, at least for the latter half of the sample. Note that over the sample period the market was becoming less concentrated,
Table 1.4: Estimated Search Cost Distributions and Indirect Utility Distributions

<table>
<thead>
<tr>
<th>Year</th>
<th>Search Cost Distribution</th>
<th>Indirect Utility Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Median 25th 75th pct</td>
<td>Max  Mean  Median Best/Mean 75th/25th</td>
</tr>
<tr>
<td>1996</td>
<td>0.038 0.011 0.128</td>
<td>6.80  4.58  4.86  1.48  1.26</td>
</tr>
<tr>
<td>1997</td>
<td>0.049 0.014 0.150</td>
<td>8.39  5.99  5.98  1.40  1.19</td>
</tr>
<tr>
<td>1998</td>
<td>0.060 0.013 0.188</td>
<td>6.48  3.97  4.05  1.63  1.45</td>
</tr>
<tr>
<td>1999</td>
<td>0.039 0.012 0.115</td>
<td>5.49  3.70  3.66  1.49  1.33</td>
</tr>
<tr>
<td>2000</td>
<td>0.028 0.011 0.067</td>
<td>3.91  2.87  2.86  1.36  1.24</td>
</tr>
<tr>
<td>2001</td>
<td>0.041 0.013 0.087</td>
<td>4.54  3.58  3.69  1.27  1.22</td>
</tr>
<tr>
<td>2002</td>
<td>0.037 0.011 0.084</td>
<td>5.75  4.71  4.88  1.22  1.10</td>
</tr>
<tr>
<td>2003</td>
<td>0.054 0.018 0.141</td>
<td>5.40  3.88  3.88  1.39  1.16</td>
</tr>
<tr>
<td>2004</td>
<td>0.066 0.024 0.161</td>
<td>7.10  5.03  5.22  1.41  1.37</td>
</tr>
<tr>
<td>2005</td>
<td>0.069 0.025 0.173</td>
<td>7.49  5.29  5.40  1.42  1.32</td>
</tr>
<tr>
<td>2006</td>
<td>0.075 0.023 0.223</td>
<td>6.74  4.63  4.64  1.46  1.34</td>
</tr>
<tr>
<td>2007</td>
<td>0.065 0.022 0.152</td>
<td>5.91  4.00  4.20  1.48  1.45</td>
</tr>
<tr>
<td>2008</td>
<td>0.073 0.025 0.148</td>
<td>6.26  4.49  4.58  1.39  1.31</td>
</tr>
<tr>
<td>2009</td>
<td>0.072 0.022 0.167</td>
<td>7.96  5.98  5.98  1.33  1.22</td>
</tr>
<tr>
<td>2010</td>
<td>0.070 0.022 0.142</td>
<td>5.90  4.29  4.18  1.38  1.22</td>
</tr>
<tr>
<td>Average</td>
<td>0.055 0.017 0.141</td>
<td>6.18  4.35  4.42  1.43  1.29</td>
</tr>
</tbody>
</table>

(a) Percentile

suggesting that better firms were losing market share to smaller, inferior firms; at least that is the interpretation in the equal-sampling version of the model. One possible explanation for reduced concentration in this version of the model is that there is an increase in consumer search costs. An increase in search costs would force some consumers to settle with a firm of low quality rather than incur the costs of an additional search with the hope of finding a sufficiently better firm. If sampling probability is more important than product differentiation, which is likely the case, then the most likely way for smaller firms to gain market share is to increase their sampling probabilities. In this version of the model, changes in sampling probability would then be incorrectly attributed to increases in consumer search costs.

Table 1.4 also shows summary statistics of the implied indirect utility distribution. The dispersion in indirect utility appears reasonable, with the best firm offering a product that is about 50% better than that of the average firm. As
Figure 1.5: Indirect Utility Distribution

Figure 1.5 illustrates, the majority of firms offer a product that closely resembles the average product. This result along with the unrealistically low search costs that I generate in the model with equal sampling gives me confidence that the more appropriate model for this application is a model with unequal sampling but homogeneous firms. Again, ideally I would want a model that simultaneously incorporates both dimensions of heterogeneity, but our data set does not allow us that luxury.

1.5 Concluding Remarks

This paper estimates the search costs distributions for consumers in the California auto insurance industry over the 1996-2010 period by making use of the novel methodology proposed by Hortaçsu and Syverson (2004). By expressing market shares as functions of the consumer search cost distribution parameters, we are able to recover the distribution function, at least up to the firm-specific search costs
cutoffs. The main goal of this paper is to assess the importance of sampling probabilities in explaining differences in market shares. In heavily regulated industries, such as auto insurance, it is likely that non-price competition will take on added significance. If large firms are able to gain and maintain a competitive advantage by outspending competitors in marketing expenditures, the industry is likely to become more concentrated, and perhaps less competitive over time. However, on the other hand, there are many industries where advertising is used to maintain competitive balance by allowing young small firms to gain market exposure and effectively compete with bigger and older firms.

In this paper I do not attempt to answer the question of whether or not advertising is used in a procompetitive or anticompetitive way. Instead I identify the effectiveness of advertising in shifting sampling probability, which is a first step in answering the bigger question of interest. In this paper I model the probability that a firm is sampled as a function of exogenous advertising expenditure and age. I find sampling probability to be positively related to both past advertising expenditure and age. The function exhibits diminishing marginal returns for advertising, but not for age. Also, I find that advertising has become more important over time, with advertising being almost as important as age at the end of the sample period.

A secondary aim of this paper is to rigorously test the hypothesis that consumer search costs have declined in the auto insurance industry over the last fifteen years. There is no clear picture for the period as a whole. However, there is an unmistakeable reduction in search costs over the last six years of the sample period. This change is consistent with our hypothesis that increased use of the Internet has lowered search costs over time.

Having established that advertising is becoming significantly more effective in attracting consumers and that consumer search costs are falling over time in the auto insurance industry, there are still a number of interesting questions that remain unanswered. Does advertising have procompetitive or anticompetitive effects in the auto insurance industry? If advertising is so effective, why do so few auto insurers advertise? How can we explain the sudden and dramatic takeoff in advertising spending in the industry over the last decade? These are all intriguing questions that I look forward to tackling in future work.
References


Dynamic Choice of Non-informative Advertising when Consumers Search Sequentially: A Numerical Approach

2.1 Introduction

The goal of this paper is straightforward. I attempt to answer a single question: Can falling consumer search costs create the incentives for firms to optimally increase advertising spending in a market where firms produce a homogeneous product? Many economists have argued compellingly that consumer search costs have fallen dramatically in the past two or three decades. See for example Brown and Goolsbee (2002) and Hortaçsu and Syverson (2004). Arguably, this result has been produced by a combination of many factors, but perhaps mainly by the recent advent of the Internet, coupled with the increasingly high market penetration of personal computers. This has allowed consumers and firms to more easily and efficiently interact. The majority of consumers are now able to quickly and conveniently compare prices and other product attributes, as more firms embrace the idea of having a virtual presence.

Consumer protection groups welcome these changes, as they anticipate that falling search costs will necessarily lead to more competitive markets, and ultimately, to increased consumer welfare. On the other hand, firms view these
changes ominously, as they fear that this anticipated increased price competition might drastically drive down profit margins. Whether either of these outcomes comes to fruition is still to be seen, but undoubtedly, they point to the fact that falling consumer search costs has a tremendous potential to radically reshape the competitive landscape of our modern-day markets.

Answering the question of “Can falling search costs lead to increased advertising when firms produce a homogeneous product?” would add significantly to our understanding of these potential implications. This question is not trivial. Actually, if the answer is yes, this result would conflict with our conventional understanding of the link between advertising and consumer search. The traditional argument can be summarized as follows: Advertising, informative or persuasive, is most effective in the presence of high search costs. Consider informative advertising. The lower the search costs, the more consumers search, regardless of the amount of advertising. Search and advertising are thus substitutes. This is easy to see when we analyze the optimal advertising strategy in the presence of zero search costs. All consumers are able to acquire all necessary information, costlessly. This renders advertising ineffectual. With advertising being costly, no firm advertises.

Now consider persuasive advertising, wherein firms advertise in an attempt to create the perception that products are differentiated. Let us assume that search allows the consumer to verify the authenticity of the advertising claims. As consumer search costs fall to zero, advertising should also fall to zero, given that consumers are now able to freely and correctly assess the value of the product attributes of all firms. Again, the result is that advertising and search are substitutes, and so advertising should decrease as search costs decrease.

Both abbreviated analyses above produce results that suggest that the thesis that falling search costs can lead to increased advertising is simply wrong. But in both analyses we made the usual big assumption – the relationship between advertising and search costs is strictly monotone. At high search costs, firms advertise a lot; at zero search costs, firms undertake zero advertising. Strict monotonicity would then imply that advertising should continuously decrease with search costs. In this paper I find that the relationship is not monotone. This result parallels another monotonicity result in a related model. Consider a market with homogeneous firms that choose price and consumers that are differentiated by search
costs. High search costs support high profits; zero search costs support a zero profit equilibrium, à la a pure Bertrand model. Strict monotonicity would then imply that profits should fall with search costs. In this paper I find that this relationship is also not monotone.

I use the modeling framework of Doraszelski and Markovich (2007) (D&M) to combine advertising choice with a sequential search model, similar to that of Carlson and McAfee (1983). I find that the monotonicity breaks for the relationship between search costs and advertising, and also for the relationship between search costs and firms’ profitability, if search costs fall within a particular range. Sufficiently low search costs cause price competition to intensify enough, so that the market can only support a highly asymmetric structure, where the large firm enjoys significantly higher profits than the small firm. But sufficiently high search costs ensure that the market does not collapse to the pure Bertrand case of marginal cost pricing, where no firm earns a profit. If search costs remain in this delicate range long enough, firms are likely to aggressively advertise to try to secure the dominant position in a market that is likely to become highly asymmetric in the long run.

D&M uses a dynamic model to show that a firm can gain a sustainable competitive advantage through advertising. They establish this claim under both goodwill advertising and awareness advertising, stressing the difference in the source of the competitive advantage in each case. Though I make use of the general framework of D&M, this paper has substantive departures from it, both in terms of theoretical modeling and interpretation.

The most fundamental departure of my model from that of D&M is the introduction of consumer search into the product market. In my model, consumers search sequentially; while firms produce a homogeneous good, advertise to influence their sampling probabilities, and choose prices, given these sampling probabilities. This product market model represents a version of Carlson and McAfee’s (1983) search model that provides an equilibrium model of price dispersion with sequential search. Carlson and McAfee allow marginal costs to vary in a static model; whereas I allow sampling probabilities to vary in a dynamic setting, with the firms choosing advertising to affect their state variable, the market exposure level. I show that in this product market model a pure strategy Nash is unlikely to exist,
even in the limited duopoly setting, if we assume that consumer search costs are distributed any other way but uniformly.

I use this hybrid model to answer my main question with a qualified yes. Falling search costs can lead to increased advertising in a market where firms produce a homogeneous product, but only under the right conditions: (1) Search costs need to fall within a delicate range. It needs to fall low enough so that price competition significantly intensifies, but it needs to remain high enough so that one firm can enjoy significant economic profits and (2) All firms need to have low exposure levels and have roughly equal sampling probabilities. Under such conditions, firms aggressively advertise in response to a reduction in search costs to try to become the long-run dominant firm.

The paper proceeds as follows: The next three sections are used to describe the model, broken-down into an overview section (2.2), a description of the per-period product market competition (2.3), and a description of the dynamic choices (2.4). The results section (2.5) presents, among other thing, the results for the product market competition (2.5.1), the value function and policy function (2.5.2), and welfare implications (2.5.5). Concluding remarks are provided in section 2.6 and importantly, the appendices discuss the question of the existence of equilibria, while also providing additional details to supplement the results section.

2.2 Model Overview

I seek to answer the question of whether lower search costs, by itself, can serve as the impetus for increased non-price advertising in a market where firms produce a homogeneous product? To answer this question I use the modeling framework of Doraszelski and Marcovich (2007) (hereafter D&M). Similar to D&M, I am interested in the advertising decisions of firms over time, and their transitional and long-run implications, and so I must use a dynamic model. The significant departure that I make from D&M is that I use a very different product market, one that focuses attention on consumer search, which intentionally is not included in D&M. In my model the product market competition resembles that of Carlson and McAfee (1983), a model later adapted by Hortaçsu and Syverson (2004) (hereafter H&S).
2.2.1 Environment

I consider a duopoly setting, where firms produce homogeneous products and have the same constant marginal production cost. Firms are potentially differentiated by their levels of market exposure. The vector of all firms’ exposure levels is the state variable, which determines the likelihood that a randomly chosen consumer will visit a given firm on his or her next search. There is a continuum of consumers, who are differentiated only by their constant marginal search costs. Firms make their decisions over a discrete infinite time horizon.

2.2.2 Timing

The timing of the game is as follows:

1. Firm $j \in \{1, 2\}$ chooses its advertising level, $x_j$, to influence its market exposure, $z_j$. $z_j$ can take on one of $L$ distinct values, labeled $1, 2, \ldots, L$. The advertising effect is stochastic and $z_j$ potentially suffers from depreciation.

2. Before the advertising effects or depreciation is realized, firms compete in the product market by choosing their prices, $(p_1, p_2)$, given $(z_1, z_2)$.

3. Consumers then engage in costly sequential search, where the sampling probability for firm $j$, $\rho_j$, is a function of the realized market exposure of all firms.

4. Market shares and profits are realized, given optimal searching by consumers and the optimally chosen prices of firms.

5. The game transitions to the next period based on a Markov process of firms’ market exposures.

As is standard, such games are solved using backward induction, and so I discuss the second stage of the game first.

---

1 Though a duopoly market structure may appear to be an oversimplification or an overly restrictive setting, the duopoly problem itself is non-trivial and provides satisfactory results with very useful insights into the interaction between advertising decisions and changes in consumer search costs.

2 Where necessary, $z_{jk}$ will denote that firm $j$ has a market exposure level of $k$. 
2.3 Stage II: The Product Market Model

2.3.1 Set-up

The product market is the duopoly-equivalent model of Carlson and McAfee (1983). Firms produce products that have homogeneous non-price attributes, \((v)\), but firms are possibly differentiated by their level of market exposure. Consumers have heterogeneous constant marginal search costs, drawn from the cumulative distribution, \(G(\cdot)\), with support \([0,\bar{c}]\). As is discussed in Appendix 2.B, though \(G(\cdot)\) can take many forms, I will restrict my discussion to a uniform distribution due to its non-decreasing density function. This guarantees the existence of at least one pure strategy Nash equilibrium in all states. Consumers are homogeneous on all other dimensions, and thus have identical utility functions when we ignore search costs. Assuming linear utility functions, a consumer that purchases from firm \(j\) derives utility,

\[
u_j = v - p_j, \quad (2.3.1)\]

2.3.2 Consumer Search and Firm Demand

Consumers are aware of the distribution of prices (and consequently, the distribution of available indirect utilities) in the market. However, a consumer must engage in costly sequential search to find out the exact net benefits to be derived from a particular firm. Consumers search under the following assumptions:

1. Consumers know the array of prices but do not know which is associated with a given firm before searching.

2. The first search is costless. This ensures that each consumer searches at least one, and consequently, that all consumers purchase one unit of the good.

3. Consumers search with replacement.

4. Consumers have constant marginal search costs.

\(^3\)For the generalized version of this model with \(N\) firms, see Carlson and McAfee (1983) or H&S.

\(^4\)In a later paper I hope to add heterogeneity in consumer risks. This would allow me to speak to an important feature of insurance industries where firms cannot advertise prices because prices vary greatly across consumers.
These assumptions are consistent with those of Carlson and McAfee (1983) and H&S, who make the last two assumptions for tractability in their N-firm case. In the duopoly setting I do not need these assumptions for tractability, but I maintain them for ease of comparison with the generalized case. Additionally, the assumptions of search with replacement and constant marginal search costs combine to do a good job of mimicking the search behavior of consumers in an industry with many firms. Except for the firm-specific cutoffs, the demand functions would be the same if we were to assume search with replacement in the duopoly setting.  

Let $\rho_j$ be the probability that firm $j$ is sampled on each search and assumes the functional form

$$\rho_j = \frac{z_j^\alpha}{z_1^\alpha + z_2^\alpha}. \quad (2.3.2)$$

Also, let $c_j$ denote the lowest search cost that a consumer can have that purchases from firm $j$. A rational consumer will decide to search again as long as the expected benefit from the next search is at least as great as the expected cost of that search. Therefore, consumers who find the better firm, will not search again, regardless of their search costs. However, the consumer who finds the worse firm will continue to search if the price difference and the probability of finding the better firm are sufficiently high relative to his or her search cost. Under the assumptions above, the standard optimal decision rule for sequential search can be used to generate critical cutoffs, $c_j$, in search costs for each firm. Suppose $p_2 < p_1$, so that firm 2 is the better firm. The search cost cutoffs would be

1. $c_2 = 0$
2. $c_1 = \rho_2(u_2 - u_1) = \rho_2(p_1 - p_2)$

Given these cutoffs we can write the market share equations as functions of the distribution function at the cutoffs:

1. $q_1 = \rho_1 - \rho_1 G(c_1) = \rho_1 (1 - G(c_1))$

See Carlson and McAfee (1983) who discuss the equivalence of the set of assumptions listed above and assuming search occurs without replacement when consumers are not certain about the precise values of the prices that have been set.

The first term in the equation is the fraction of consumers that find firm 1 first. The second term is the fraction that continue to search until they find firm 2. Given the assumptions of search with replacement and constant marginal search costs, the second term can be written as $\rho_1 G(c_1) \left( \rho_2 + \rho_1 \rho_2 + \rho_1^2 \rho_2 + \rho_1^3 \rho_2 + \cdots \right) = \rho_1 \rho_2 G(c_1) \sum_{k=0}^{\infty} \rho_1^k = \rho_1 \rho_2 G(c_1) \frac{1}{1 - \rho_1} = \rho_1 G(c_1)$.  

---

5See Carlson and McAfee (1983) who discuss the equivalence of the set of assumptions listed above and assuming search occurs without replacement when consumers are not certain about the precise values of the prices that have been set.

6The first term in the equation is the fraction of consumers that find firm 1 first. The second term is the fraction that continue to search until they find firm 2. Given the assumptions of search with replacement and constant marginal search costs, the second term can be written as $\rho_1 G(c_1) \left( \rho_2 + \rho_1 \rho_2 + \rho_1^2 \rho_2 + \rho_1^3 \rho_2 + \cdots \right) = \rho_1 \rho_2 G(c_1) \sum_{k=0}^{\infty} \rho_1^k = \rho_1 \rho_2 G(c_1) \frac{1}{1 - \rho_1} = \rho_1 G(c_1)$.  

---
2. \( q_2 = \rho_2 + \rho_1 G(c_1) \)

Consumers purchase from firm 1, the worse firm, only if they find firm 1 first and they have search costs that are high enough to dissuade them from searching again. Firm 2’s demand comprises those consumers that visit firm 2 first, and those that find firm 1 but have sufficiently low search costs, so that they continue to search until they eventually find firm 2.

In this model, market demand is perfectly inelastic. A firm can increase its market share only by “stealing” customers from its rival. On the flip side, if the firms jointly increase their prices, consumers still must buy from one of these firms. By jointly increasing prices the firms increase their profit margin without sacrificing on quantity. As such, at first glance, it might seem plausible that firms would charge the monopoly price if firms have equal sampling probabilities. But this is true only if the fraction of high-search costs consumers is sufficiently big. Otherwise, both firms would have an incentive to undercut, because in doing so it would induce enough consumers to continue to search for and to find the low-price firm to outweigh the reduction in the profit margin per unit.

For a better understanding of the demand functions in this model, I sketch a general demand curve for this model in figure 2.1 juxtaposed against a demand function from the standard Bertrand duopoly pricing model. Recall that in the pure Bertrand model a discontinuity occurs as the two firms’ prices converge. Suppose firm 2’s price is fixed at \( \bar{p}_2 \). Firm 1 gains the entire market by charging a price less than \( \bar{p}_2 \), but loses the entire market by choosing a price even just slightly above \( \bar{p}_2 \). Without any refinement, firm 1’s market share can be anywhere in the closed interval of \([0, 1]\) when firms charge the same price.

One interpretation of our model is that it is a Bertrand pricing model with consumers who are differentiated by their search costs. The firm that charges a slightly higher price does not lose all its customers, but only those customers that

\[ q_0 = \rho_0 (1 - G(c_0)) \]

\[ q_1 = \rho_1 \left[ 1 + \frac{\rho_0}{1 - \rho_0} G(c_0) - \frac{1}{1 - \rho_0} G(c_1) \right] \]

\[ q_2 = \rho_2 \left[ 1 + \frac{\rho_0}{1 - \rho_0} G(c_0) + \frac{\rho_1}{(1 - \rho_0)(1 - \rho_0 - \rho_1)} G(c_1) \right] \]

\[ \text{For those curious, under the same search assumptions, a market with three firms with } p_2 < p_1 < p_0 \text{ would have the following demand functions:} \]

(a) \( q_0 = \rho_0 (1 - G(c_0)) \)

(b) \( q_1 = \rho_1 \left[ 1 + \frac{\rho_0}{1 - \rho_0} G(c_0) - \frac{1}{1 - \rho_0} G(c_1) \right] \)

(c) \( q_2 = \rho_2 \left[ 1 + \frac{\rho_0}{1 - \rho_0} G(c_0) + \frac{\rho_1}{(1 - \rho_0)(1 - \rho_0 - \rho_1)} G(c_1) \right] \)
have low enough search costs to induce them to search again. As the gap between the prices widens, the savings from purchasing from the low-price firm increases, causing more consumers to be able to afford additional searches. The fraction of the consumers that choose to search again, having found the high-price firm, depends on the size of the price gap, the probability of finding the low-price firm, and critically, on the search costs distribution function.

Figure 2.1 illustrates a representative demand function in our model when we assume that search costs are uniformly distributed over $[0, b]$. Holding the price of firm 2 fixed at $\bar{p}_2$, firm 1 secures the entire market if it charges a price that is sufficiently lower than $\bar{p}_2$. That is, $q_1 = 1$ if $(\bar{p}_2 - p_1) > \frac{b}{\rho_1}$. If firm 1 increases its price beyond $(\bar{p} - \frac{b}{\rho_1})$, it loses market share to firm 2 at a rate proportional to the density function of the search cost distribution. Firm 1 loses the consumers with the lowest search cost first and if it increases its price beyond $(\bar{p} + \frac{b}{\rho_2})$, its market share falls to zero. The heterogeneity in consumer search costs creates kinks in the demand function, but it gets rid of the discontinuity that exists in the pure Bertrand model. This in turn, ensures that we have continuous profit functions, which I discuss next in section 2.3.3.
2.3.3 Profit-maximization and Equilibrium

For the supply side, the profit-maximization problem for firm $j \in \{1, 2\}$ is

$$\max_{p_j > 0} q_j(p_1, p_2; z_1, z_2)(p_j - mc_j)M,$$

(2.3.3)

where $M$ is the total market size and $mc_j$ is the marginal cost of production for firm $j$. A specific example may be useful in giving us a better understanding of the profit-maximizing decision of the firms when faced with the type of kinked demand functions that we have in this model. Figure 2.2 shows a demand function and the associated profit function for firm 1. These functions are drawn under the assumptions that $c \sim U[0, 15]$, the firms have equal sampling probabilities, firm 2’s price is fixed at $\$50$, the marginal cost is zero, and the market size is 10.

If firm 1 charges a price of less than $\$20$, it gets the entire market. Firm 1 can therefore increase its price up to $\$20$ with no loss in its market share. This gives us the linear portion of the profit function. Continuing to increase $p_1$ beyond $\$20$ would then cause firm 1 to lose market share to firm 2, as consumers increasingly stay with firm 2 rather than search again when $p_1 < p_2$ (or consumers increasingly search again having found firm 1 when $p_2 < p_1$). This gives us the strictly concave portion of the profit function when $p_1 \in [20, 80]$, which corresponds to the standard downward-sloping linear segment of the demand function. Increasing price beyond $\$80$ causes $p_1$ to be sufficiently higher than $p_2$ for all consumers to search until they find the lower-priced firm 2.

Generally, a pure strategy Nash equilibrium, $(p_1^*, p_2^*)$, may not exist. If one exists in the interior, the Nash prices are the solutions to the first order conditions (FOCs)

$$q_j + (p_j - mc_j) \frac{\partial q_j}{\partial p_j} = 0,$$

(2.3.4)

$$\frac{\partial q_j}{\partial p_j} = -\rho_1 \rho_2 g(c_1).$$

(2.3.5)

$g(c_1)$ is the search cost density function at the cutoff for firm 1, the high-price firm. Recall that $c_1$ depends on the difference between $p_1$ and $p_2$ and on the sampling probability of the better firm; in this case firm 2. The existence of a pure strategy Nash thus depends critically on the shape of the search costs density function,
especially near to zero, as the price gap shrinks.

Carlson and McAfee (1983) explains “With customers assumed to be using a sequential, reservation-price, search strategy and to have correct perceptions of the price distribution, firms need to be different in some respects in order to sustain a persistent dispersion of prices.” They choose differences in marginal costs; I choose differences in sampling probability because of my interest in advertising choice. Carlson and McAfee go on to prove the existence of a pure strategy Nash equilibrium where the order of the optimal prices follows the order of the firms’ marginal production costs. In the N-firm case they find that a non-decreasing search cost density function is a sufficient condition for the existence of such an equilibrium.

Even in a model with only two firms, the existence of a pure strategy Nash is not guaranteed, and still depends on the shape of the density function. Actually, under different search cost distributional assumptions and in different states, I find numerically that the absence of a pure strategy Nash is the norm, rather than the exception. To ensure that a pure strategy Nash exists in all states of the game, I assume a uniform search cost distribution throughout the remainder of the paper.\footnote{In Appendix 2.B I illustrate the non-existence of a pure strategy Nash equilibrium in a state with equal sampling probability but where search cost is log-normally distributed.} Without loss of generality, I also assume $mc = 0$ for both firms, and we can now rewrite the FOCs as
\[ \rho_1 - \rho_1 G(c_1) - \rho_1 \rho_2 g(c_1) p_1 = 0 \]  
\[ \rho_2 + \rho_1 G(c_1) - \rho_1 \rho_2 g(c_1) p_2 = 0 \]  

Under the assumption that search costs are uniformly distributed over the interval of \([0, b]\), we can further reduce the FOCs and solve for the best response functions for each firm:

\[ p_1 = \frac{b}{2\rho_2} + \frac{p_2}{2} \]  
\[ p_2 = \frac{b}{2\rho_1} + \frac{p_1}{2} \]  

Notice that the optimal price of a given firm is monotone in the price of the other firm. This is a standard result in most pricing models. In this model, the reason is straightforward – the demand function and profit function are consequences of how much consumers search, and the intensity of search is a direct consequence of the size of the price gap. As such, if one firm increases its price, the rival firm must also increase its price to maintain that optimal gap. This is also a function of the perfectly inelastic market demand, because one firm can lose market share only to the other firm.

Simultaneously solving the best response functions gives us the Nash equilibrium prices as functions of the upper limit of the search costs distribution and the sampling probabilities of the firms:

\[ p_1 = \frac{b}{2\rho_2} + \frac{1}{2} \frac{b(\rho_1 + 2\rho_2)}{3\rho_1\rho_2} \]  
\[ p_2 = \frac{b(\rho_1 + 2\rho_2)}{3\rho_1\rho_2}. \]  

This result allows us to treat market exposure as the state variable because, ignoring the exogenous parameter values, the static pricing decisions entirely depend on the firms’ sampling probabilities, which in turn are deterministic functions of the firms’ levels of market exposure.
A few points warrant our attention here. First, recall that the demand functions were written assuming that \( p_2 < p_1 \). Combining the FOCs in equation 2.3.6 establishes that if an interior Nash equilibrium exists where \( p_2 < p_1 \), it must be true that \( \rho_2 < \rho_1 \). This is illustrated in figure 2.3 which shows the Nash equilibria for the pricing game in three different states. Carlson and McAfee (1983) prove that prices are monotone in marginal production costs in the N-firm case of the sequential game of the type described above. Here I show that prices are monotone in sampling probability for the duopoly case.

This monotone relationship comes from the fact that a larger share of the high-search cost consumers will purchase from the high-exposure firm. This in turn results from two complementary sources of a competitive advantage: (1) There is a greater likelihood that the high-exposure firm will be visited first and (2) The expected benefits of continued search for a consumer that has found the high-exposure firm is lower than that of those that find the low-exposure firm, for all pairs of prices. These effects combine to ensure that a high-price, high-exposure firm retains more of the high-search costs consumers than would a high-price, low-exposure firm. In short, the high-exposure firm has greater incentives to charge a higher price, and so an equilibrium with different prices must have the high-exposure firm charging the higher price.

Second, the larger the difference in sampling probabilities, the greater the disparity in Nash prices, with the price gap converging to zero as the sampling probabilities converge. This follows directly from the quantity-price derivative, \( \frac{\partial q_1}{\partial p_1} = -\rho_1 \rho_2 g(c_1) \). It is easy to see that the more equal the sampling probabilities are, the greater the drop-off in market share when a firm increases its price, all else constant. As we would expect, price competition is more intense in states where firms are more homogeneous. Towards the other end of the spectrum, if the difference in sampling probabilities is sufficiently big, the high-exposure firm will charge the monopoly price. In these cases the low-exposure firm still is unlikely to charge the monopoly price because that would reduce the incentive for consumers to continue to search. The probability of finding the low-exposure firm is so low that it must induce some of the consumers that find the high-exposure firm to search again. To induce these consumers, the low-exposure firm must charge a sufficiently lower price. These results are summarized in figure 2.3 which show
the best response functions of firms 1 and 2 in three different states.

Figure 2.3: Best Response Functions

![Best Response Functions](image)

(a) \( \rho_1 = 0.27 \) and \( \rho_2 = 0.73 \).
(b) \( \rho_1 = 0.5 \) and \( \rho_2 = 0.5 \).
(e) \( \rho_1 = 0.9 \) and \( \rho_2 = 0.1 \).

Third, the equilibrium is symmetric. Let \( p_j^*(k,l) \) denote the Nash equilibrium price of firm \( j \) in state \((z_1k, z_2l)\). A symmetric Nash equilibrium implies that \( p_1^*(k,l) = p_2^*(l,k) \). This is most easily seen in the best response functions in equation 2.3.7. This is a useful result as it leads to the profit, value, and policy functions also being symmetric, which allows me to exclusively focus on symmetric Markov-perfect equilibria. Fourth, under the assumptions that I have made, and will maintain throughout the paper, if a Nash equilibrium exists, it is always unique as the response functions are linear in the interior of the price space and have different slopes.

Fifth, the Nash prices are increasing in \( b \), the upper bound on consumer search costs. This is a somewhat obvious result where firms charge higher prices when the market comprises a higher proportion of high-search costs consumers. A more subtle result, and one that is critical to us understanding the impact of changing search costs on advertising choices is that the Nash price (and profits) of the high-exposure firm is less sensitive to changes in \( b \) than is the price (and profits) of the low-exposure firm when the initial level of search costs is relatively high. All the results mentioned above are confirmed numerically.

Having solved the static pricing game, I now discuss the dynamic pricing decisions of the firms that know that the period profit function for firm \( j \) in state

\footnote{Recall that the first element in the pair of exposure index always represents the exposure for firm 1.}
(z_1, z_2) is given by
\[ \Pi_j(z_1, z_2) = q_j(p^*_1, p^*_2, z_1, z_2)p^*_j M. \]  
(2.3.9)

2.4 Stage I: The Dynamic Advertising Choice

2.4.1 State-to-State Transitions

The state of the game is described fully by \((z_1, z_2)\). The state variables for firms 1 and 2 transition in the manner described by D&M. The exposure level of firm \(j\) in the next period, \(z_j'\), can move up or down by one level, or remain unchanged. Firm \(j\) influences its likelihood of moving to one of the adjacent exposure levels by how much it advertises this period, \(x_j\). We can interpret this process as firms engaging in one-period advertising campaigns that can succeed or fail. The probability that the campaign succeeds is given by \(\theta \equiv \frac{x_j}{x_j + 1}\), so the success rate increases with \(x_j\) but at a decreasing rate.

There is also a probability that the firm’s exposure level suffers a one-level depreciation. That is, the firm’s exposure level may actually remain unchanged when the advertising campaign is successful, or worse yet, the exposure level may fall one level when the campaign is unsuccessful. \(\delta\) denotes the probability that the firm’s market exposure suffers depreciation. I assume that the success or failure of the advertising campaign is independent of whether or not exposure depreciates. As such, the probability that the firm remains at its current exposure level would be

\[ \frac{x_j}{x_j + 1} \delta + \frac{1}{x_j + 1} (1 - \delta) = \frac{1 + \delta(x_j - 1)}{x_j + 1}. \]

This probability is calculated by recognizing that the firm’s exposure can remain unchanged if one of two sets of events occur:

1. The advertising campaign is unsuccessful and there is no depreciation.
2. The advertising campaign is successful but there is depreciation.

I assume that the success rate of advertising campaigns is zero if the firm is currently at the maximum exposure level, \(z_j = z_j L\); and that the probability
of depreciation is zero if the firm is at the minimum exposure level, \( z_j = z_{j1} \). \( \text{Prob}(z'_j = z_{jl} | z_j = z_{jk}; x_j) \) represents the probability that firm \( j \) transitions to exposure level \( l \) given it is currently at level \( k \) and given that it chooses advertising level \( x_j \) this period. I further simplify this notation to \( \text{Prob}(k'|k; x_j) \).

If the firm is at the lowest exposure level, \( k = 1 \), the transition probabilities are given by

\[
\text{Prob}(k'|1; x_j) = \begin{cases} 
0 & \text{if } k' = 0 \\
\frac{1}{x_j + 1} & \text{if } k' = 1 \\
\frac{x_j}{x_j + 1} & \text{if } k' = 2 
\end{cases}
\]

If the firm is at the maximum exposure level, \( k = L \), the transition probabilities are given by

\[
\text{Prob}(k'|L; x_j) = \begin{cases} 
\delta & \text{if } k' = L - 1 \\
(1 - \delta) & \text{if } k' = L \\
0 & \text{if } k' = L + 1 
\end{cases}
\]

If the firm is at an interior point in the exposure space, \( 1 < k < L \), the transition probabilities are given by

\[
\text{Prob}(k'|k; x_j) = \begin{cases} 
\frac{\delta}{x_j + 1} & \text{if } k' = k - 1 \\
\frac{1 + \delta(x_j - 1)}{x_j + 1} & \text{if } k' = k \\
0 & \text{if } k' = k + 1 
\end{cases}
\]

### 2.4.2 Bellman Equation and Policy Function

Given the per-period profit function and the transition probabilities, we can write down the value function for each firm given the current state. The functional equation simply adds the current profits, net of the advertising spending, to the infinite sum of the stream of future profits, discounted. \( \beta \) is the discount factor. The value of firm 1 given the current state \((z_1, z_2)\) is given by

\[
V_1(z_1, z_2) = \max_{x_1 \geq 0} \Pi_1(z_1, z_2) - x_1 + \beta \sum_{k=1}^{L} W_1(z'_{1k}) \text{Prob}(z'_{1k}|z_1, x_1) \quad (2.4.1)
\]
Realize that the continuation value for firm 1 of having a given exposure level next period, $W_1(z'_{1k})$, is calculated assuming that firm 2 chooses the optimal advertising level given the current state, $(z_1, z_2)$. The policy function, $x_j(\cdot, \cdot)$, is the solution to the maximization problem in equation 2.4.1. It is easy to see that the first order condition is

$$\sum_{k=1}^{L} W_1(z'_{1k}) \frac{\partial \text{Prob}(z'_{1k} | z_1, x_1)}{\partial x_1} = \frac{1}{\beta}. \quad (2.4.3)$$

Increased advertising increases the probability of moving to a better state, which increases the firm’s future value. The cost to the firm of increased advertising is simply the monetary cost of the advertising today. At an interior solution the marginal benefits of increased future returns must be equal to the marginal cost in future value, $\frac{1}{\beta}$. Recall that the transition probabilities depend on the firm’s current exposure level, and that the functional form differs when the firm is either at the minimum or maximum exposure level. D&M show that we can solve for the following solution when $1 < k < L$, and the solutions when $k \in \{1, L\}$ can be analogously derived.\(^{10}\)

$$x_1(z_{1k}, z_{2l}) = \max \left\{ 0, -1 + \sqrt{\beta \left[ (1 - \delta)(W_1(z_{1(k+1)}) - W_1(z_{1k})) \right]} + \beta \delta \left[ W_1(z_{1k}) - W_1(z_{k-1}) \right] \right\} \quad (2.4.4)$$

### 2.4.3 Equilibrium and Parameterization

Recall that the per-period profit functions are symmetric under the assumption of equal marginal production costs, which I assume to be zero. Given the symmetry in profit functions and the fact that the firms have the same transitional probabilities, I focus exclusively on finding a symmetric Markov-perfect equilibrium (MPE). That is, $V_1(k, l) = V_2(l, k)$ and $x_1(k, l) = x_2(l, k)$. As such, the value functions

\(^{10}\)See D&M for a discussion of issues of existence of the solution.
and policy rules are not firm-specific, allowing us to write the value function and policy rule as $V(k, l)$ and $x(k, l)$, respectively, when firm 1 has an exposure level of $z_{1k}$, and firm 2 has exposure $z_{2l}$. Doraszelski and Satterthwaite (2003) prove the existence of a symmetric Markov-perfect equilibrium in this type of model as long as advertising is bounded above. They however, find that uniqueness is not guaranteed. I use the Gauss-Jacobi variant of the Pakes and McGuire’s (1994) algorithm to solve for the MPE, which produces the same value function and policy rule for different initial points.

The main focus of this paper is to analyze how the advertising behavior of firms are impacted by a reduction in consumer search costs. As such, where possible I adopt the parameter values for the dynamic portion of the model from D&M. I set $L = 15$ and the maximum exposure level equal to 100. That is, the firm’s market exposure, $z$, can take on one of fifteen possible exposure levels, which are uniformly taken from the interval $[1, 100]$. Given the sampling probability function in equation 2.3.2, this discretization results in sampling probability ranging from 0.025 to 0.975. I use this relatively small number of grid points because the results are qualitatively the same when $L$ is equal to 21 or even 30.

I assume a uniform distribution for consumer search costs; specifically I assume that $c \sim U[0, b]$. I will reduce the upper bound of the distribution to analyze the effects of a reduction in search costs. This gives the effect that the lowest possible search cost cannot be reduced, but that those with search costs above a given level is uniformly assigned a search cost below that threshold. Throughout the paper, I will present results for $b \in \{40, 30, 20\}$, usually in that order to produce the effect of decreasing search costs. These values are reasonable given that the value of 100 is chosen for $v$, each consumer’s value of consuming the product. This is consistent with Williams (2013), who find the median search cost to be 24% of the value of the auto insurance coverage. These values for $b$ also provide sufficient range to give us a good picture of the likely effects of falling consumer search costs.

Following D&M, I interpret a period as a quarter and set the rate of depreciation of market exposure, $\delta = 0.3$. The discount rate is 0.02, so that the discount factor is $\beta = \frac{1}{1.02}$. $\alpha$, a measure of how market exposure affects sampling probabilities, is set equal to 0.8. This is consistent with the established result that diminishing marginal returns will set in as a firm attempts to increase its market exposure.
And the market size ($M$), a scale variable, is set equal to 10.

### 2.5 Results

#### 2.5.1 Product Market Competition

To fully appreciate the nuances of the dynamic choices of the firms, we must first have a good understanding of the pricing decisions that are taken each period. Many of the pricing results are discussed in section 2.3.3 which examines the Nash equilibrium in the product market. Here, I expand on these results, particularly focusing on how the firms’ pricing behavior differs across states and across different levels of search costs. This allows us to better understand the implications of falling search costs on advertising decisions, my key objective.

The firms’ pricing decisions are summarized in figure 2.4. The Nash prices depend exclusively on the state of the game. Both firms charge the same price in all states that have equal market exposure. That is, $p^*(k, k) = p^*(l, l) \in (mc, v)$ for all $k, l = 1, 2, \ldots, 15$. The result that the prices are equal in states of equal exposure follows from the symmetry of the profit functions. That the price is the same for all equal-exposure states holds because there is no outside firm, and it is the sampling probabilities that drive the firms’ pricing decisions. There is therefore no gain from increased market exposure if the rival firm has a matching increase. This abstraction allows us to focus on the combative advertising effect, without having to disentangle it from the market-expanding effect. The result that the Nash prices are bounded away from marginal costs holds because a non-degenerate search cost distribution takes us out of the pure Bertrand pricing world, and ensures that firms earn strictly positive profits. Finally, prices are bounded away from the monopoly price because search costs are sufficiently low, so that each firm would steal enough customers by undercutting to make deviating from monopoly pricing worthwhile.\(^{11}\)

As we move to states with unequal exposure levels, the high-exposure firm charges a higher price than the low-exposure firm. Loosely, the greater the disparity

---
\(^{11}\)For high enough search costs both firms charge the monopoly price ($p^*(k, k) = v$), but I ignore these less interesting cases.
in exposure, the greater the gap in Nash prices. When the asymmetry in exposure is acute enough, prices become strictly monotone in exposure until the price hits the upper bound. That is, the high-exposure fully capitalizes on any increase in its dominant exposure position by further increasing its price. The low-exposure firm, at the same time, reduces its price to survive. The intuition is straightforward; the low-exposure firm can attract consumers only by charging a sufficiently lower price, and this price gap must increase as the low-exposure firm loses more ground. For example, in the states of extreme asymmetry, (1, 15) or (15, 1), 97.5% of the customers will visit the high-exposure firm first. The low-exposure firm must give consumers a reason to search again. It does so by charging a significantly lower price than the high-exposure firm.

The firm’s own-price elasticities ($\epsilon$) at the equilibrium prices do not add much
information with regards to the degree of competition in a given state. This is the case because, at the interior Nash equilibrium both firms are maximizing their profits on the linear downward-sloping portion of the demand function. As such, when \( mc = 0 \), maximizing profits is equivalent to maximizing total revenues, and this occurs at the price where demand is unitarily elastic, \( |\epsilon| = 1 \). We do see the expected results in states that are highly asymmetric. The high-exposure firm’s price is a corner solution, so \( |\epsilon| \neq 1 \), and \( |\epsilon| \) decreases for the high-exposure firm as the game moves to states where its exposure advantage increases. This simply reflects the fact that the firm would prefer to increase its price but cannot.

The overall pricing pattern discussed above is consistent for all levels of search costs, as is illustrated in figure 2.4. There are two important differences that show up in the pricing behavior of firms as search costs fall. First, as one would expect, prices in all states for both firms fall with search costs. Lower search costs induces consumers to search more, which forces firms to lower their prices. The average share-weighted price falls from $\$87.73$ to $\$72.75$ to $\$53.99$ as \( b \) falls from 40 to 30 to 20, respectively.

Second, a more subtle difference is that the reduction in price is more dramatic for the low-exposure firm. Using the share-weighted prices averaged across the relevant states, I find that prices fell a little over 1% more for the low-exposure firm than for the high-exposure firm when \( b \) falls from 40 to 30. The difference jumps to over 3% when \( b \) falls from 30 to 20. The difference can be more easily seen by considering one of the extreme states, \((15, 1)\) or \((1, 15)\). At all three levels of search costs, the high-exposure firm charges the monopoly price of \( v = \$100 \), whereas the low-exposure firm reduces its price from $\$71$ to $\$65$ and then to $\$60$ when \( b \) falls from 40 to 30 and then from 30 to 20, respectively.

The effects of lowering search costs on prices translates directly to effects on profits. For the vast majority of states, the profits of both firms fall. The average producer surplus is $\$877.33$ when \( b = 40 \), $\$727.49$ when \( b = 30 \), and $\$539.86$ when \( b = 20 \). Similar to prices, the profits of the low-exposure firm fall more than those of the high-exposure firm. When \( b \) falls from 30 to 20, the profits of the low-exposure firm averaged across all states fall by 27.5%, a little over 3% more than the reduction for the high-exposure firm.

The main reason for this is that the reduction in search costs generates many
more marginally profitable states. To illustrate, consider states of equal market exposure. Recall that in each of these states firms charge the same price. When $b = 40$, firms charge $\$80$, and split the market equally, so that each firm earns a profit of $\$400$. When consumer search costs fall to $b = 20$, both firms lower their prices to $\$40$ and now each earns a profit of only $\$200$. When search costs are relatively high, states that have close to equal exposure levels, are able to support two reasonably profitable firms, with firms earning profits upwards of 40% of the total consumer value created in the market. States that are extremely asymmetric do not change much as search costs fall. Consider state $(15, 1)$. Firm 1 earns at

\[ \text{An obvious question may arise: Why would the firms reduce their prices rather than keeping them fixed at $\$80$ or even increase prices jointly to $\$100$? I assume that the game is non-cooperative, in which case firms cannot collude. With this assumption, given the lower search costs, a firm would gain by undercutting, if its rival’s price exceeds $\$40$.} \]
least $928 in profits for the three search costs levels, whereas firm 2 earns at most, $43. Firm 1’s profits are at least 21-times that of firm 2, the low-exposure firm, regardless of search cost.

2.5.2 Value Function and Policy Function

Figure 2.6: Value Function and Policy Function for $b = 40$ and $b = 30$

I now turn attention to the dynamic decisions of the firms by analyzing the value functions in conjunction with the policy functions. Recall that the value function is simply the sum of the stream of discounted future profits, net of the optimally chosen advertising outlays. The policy function gives the optimal advertising level in each state. Figure 2.6 depicts the value of a firm in the different states of the game along with the corresponding optimal advertising level when search costs
are relatively high, $b = 40$ or $30$. In line with the results of the profit function, the value function is increasing in the firm’s own exposure level and decreasing in its rival’s exposure level. This occurs because of two complementary reasons: (1) The period profits are higher in states where the firm has a higher sampling probability and (2) A firm with high exposure level this period is more likely to have high-exposure in the future.

For relatively high search cost levels, $b = 40$ or $b = 30$, though each firm would prefer to have high exposure, there is no sharp drop in the firm’s value for being in a state where it has low exposure. This again follows from the shape of profit function in different states. A low-exposure firm recognizes that a marginal increase in exposure increases its profits next period. It is therefore willing to invest in advertising to increase the probability of moving up. Each firm ultimately would prefer to be a maximum-exposure firm competing against a minimum-exposure firm, but each earns enough in the intermediate states to be comfortable with a gradual move up. The firms are therefore willing to be patient in their advertising strategies. Each firm advertises a moderate amount each period to increase their probability of moving up, and advertising tapers off as a firm approaches the maximal exposure level, as the effects of diminishing returns increase. Ultimately, the policy functions are likely to take the game to a state where both firms have high exposure, a result that we will return to in section 2.5.3.

For the case of relatively low search costs, $b = 20$, the shape of the value function is dramatically different. It is still increasing in the firm’s own exposure level and decreasing in its rival’s exposure level. However, the disparity between the value of the firm in the most profitable states and its value in the least profitable states is significantly greater when $b = 20$ than at higher levels of search costs. The range in valuations when $b = 40$ is $\$5,229$, whereas it is $\$18,818$ when $b = 20$. Also, though the means are not the same, the difference in the standard deviations are also noteworthy. The standard deviations are $\$1,139$ and $\$7,522$ when $b = 40$ and 20, respectively. See table 2.A.1 in appendix 2.A for additional statistics.

The major takeaway here is that at low enough search costs levels very few states offer a moderately high value. Instead the vast majority of states fall into one

---

13 These relationships are not strictly monotone, but the non-monotonicity here is not acute enough to invalidate the arguments presented.
of two categories: Those that are extraordinarily profitable or those that generate meager profits. Though the value function for $b = 30$ lies almost completely below that of $b = 40$, as one would reasonably expect; the monotonicity does not hold true for very low levels of search costs. In fact, at $b = 20$, 63 of 225 states offer a value greater than the maximum value at $b = 40$ ($23,032$), while 149 states offer
a value less than the minimum value at \( b = 30 \) (\$13,748)\(^{14}\). This is depicted in figure 2.8.

Broadly speaking, a bad state for firm 1 is one in which it has lower market exposure than firm 2, or firm 2 has a relatively high exposure level, which would make it unlikely that firm 1 would ever be able to secure an overwhelmingly dominant position. Based on the parameterization in this paper, when \( b = 20 \), states where firm 1’s sampling probability is less than 53% or firm 2’s exposure level is more than 7, would constitute bad states for firm 1. These states are bad not simply because they have low per-period profits, but perhaps more importantly, in these states there is a low probability that the firm will be able to move to a good state in the future. One may reasonably ask, “Why would a firm in a bad state, not simply increase its advertising in an effort to increase its probability of moving to a good state?” The answer to that question comes from the shape of the policy function when \( b = 20 \).

Recall that when search costs are high, both firms advertise enough for each to eventually achieve close to maximum exposure. This is not the case when search costs are sufficiently low, as is illustrated in figure 2.7b. As is discussed in section 2.5.3, when \( b = 20 \), the long-run equilibrium market structure is highly asymmetric, with a high-exposure firm competing against a significantly overmatched low-exposure firm. At low levels of search costs, consumer search intensifies, which in turn leads to increased price competition. This intensified product market competition in each period ensures that only a highly asymmetric state with only one highly profitable high-exposure firm can be sustained in the long run. Obviously, both firms would prefer to be the high-exposure firm in such a state, and each firm chooses its policy function to achieve this end, if possible.

Figure 2.7b illustrates that there is a select group of states where firms aggressively advertise in an attempt to become that dominant firm. For the remaining discussion, I label a firm’s advertising strategy as aggressive if its advertising-to-sales ratio exceeds 10%. Given this definition, firms aggressively advertise in only

\(^{14}\)Going forward, when \( b = 20 \), I will classify the states offering a value greater than \$23,032 (the maximum value at \( b = 40 \)) as good states and those offering less than \$13,748 (the minimum value at \( b = 30 \)) as bad states. Note that these thresholds are somewhat arbitrary, but that the subsequent arguments would follow through with any set of cutoff values chosen with a little care. Also, sometimes the arguments will be made with respect to firm 1 or firm 2. This is simply for ease of exposition. Given the symmetry of the game, firm identities are completely irrelevant.
30 of the 225 states. During the period of aggressive advertising, firms typically spend in excess of 20% of total revenue on advertising, and sometimes even more than 30%. This type of advertising intensity appears even more impressive when compared to advertising-to-sales ratio of a little over 1% during non-aggressive periods.

The group of states in which a firm aggressively advertises meets two general criteria: (1) The firm must have the same or slightly higher exposure than its rival and (2) Both firms must have relatively low market exposure levels. To understand why firms aggressively advertise only in this particular set of states, let us consider the other three broad sets of states.

First, consider the states where both firms have high exposure. Though these states offer each firm a low value when compared to the maximum value that can be achieved in a highly asymmetric state, each firm’s value function is increasing in its own exposure level (at least with respect to close neighboring states). As such, each firm has an incentive to maintain its high exposure, and hopes that the other firm has a string of unfortunate advertising and depreciation shocks, so that it would gain the dominant position. Therefore, if both firms have high exposure, neither will aggressively advertise, but instead each would practice maintenance advertising: In an attempt to prevent its exposure from falling below a certain level, a firm advertises just enough. Given the average advertising outlay in non-aggressive periods based on simulations, maintenance advertising amounts to about 1.5% of total revenues.

Second, consider the states where one firm has a significant exposure advantage over the other. Given the discussion above, it is clear that the firm with the advantage need not aggressively advertise, it simply needs to spend just enough to reduce the probability of its exposure falling. The disadvantaged firm likewise chooses to practice maintenance advertising. It would aggressively advertise if it

---

15I exclude 10 states where a firm actually has an advertising-to-sales ratio in excess of 10%, but it is clear that the firm is not trying to achieve a dominant position in the market. The value of the firm in a state where it has minimum exposure but its rival has relatively high exposure is so depressed that the firm advertises significantly, simply to escape these states of doldrums. Once the firm gets out of one of these states, its advertising falls back to normal levels, rather than it actively trying to ascend to a dominant position. Also, this cutoff of 10% is somewhat arbitrary, but the argument remains intact for any reasonably chosen cutoff value. For example, with a cutoff of 5% the number of states where a firm aggressively advertises increases from 30 only to 46.
expected that in doing so, the competitive landscape would likely flip, so that it would become the dominant firm. However, given that the dominant firm will optimally practice maintenance advertising, an aggressive advertising campaign by the small firm would, at best, take the game to a state with two high-exposure firms. When search costs are low enough, such a state offers very little gain relative to the current state. The small firm, knowledgeable of the large firm’s policy function, chooses to advertise very little. Third, by symmetry, neither firm has an incentive to aggressively advertise if the identity of the firms were reversed.

Having eliminated the other broad groups of states, we return to the original statement that firms aggressively advertise only in states where both firms have low exposure levels and where the firms’ exposure levels are roughly equal. Each firm recognizes that there is an opportunity to become the long-run dominant firm, where it would earn extraordinarily high profits. Both firms start out with aggressive advertising campaigns. Eventually, one firm gains a sufficiently big advantage, so that the identity of the more likely long-run dominant firm is revealed. Given that advertising is costly, future profits are discounted, and the probability that the advantage will eventually flip is low, the low-exposure firm reduces its advertising significantly and settles into maintenance advertising. Figure 2.9 shows a representative advertising profile over time, given the policy function and the transition probabilities.

The advertising strategy here when $b = 20$ resembles the preemption race described by D&M in their model of awareness advertising, when the value of the product is high. However, there are substantive differences between their results and mine. In their model, the preemption race arises because of an increase in the value of the firms’ product, relative to the outside good. This causes an increase in the number of consumers that prefer the inside good, which leads to more intense price competition between the inside firms. The long-run equilibrium therefore, becomes more asymmetric and the incentives to advertise increase. In my model there is no outside firm. The increased price competition results from a significant reduction in consumer search costs, and ultimately leads to a dramatic reduction in profits in all but the highly asymmetric states for the high-exposure firm. The source of increased advertising is thus the reduction in consumer search costs, which D&M do not directly address.
Figure 2.9: Simulations when $b = 20$ and Initial State is $(2, 2)$.

To allow for a complete comparison, while avoiding cluster, I include the simulation graphs when $b = 30$ and 40 in Appendix 2.A.

Additionally, D&M, find that the firms aggressively advertise as long as firms have roughly equal awareness levels. In my model, firms aggressively advertise only in states where both firms have relatively low exposure. The crucial difference lies in the difference between the value of a firm in a state with two large firms, relative to the firm’s value when its state variable is lower. In D&M, the value of firm 1 in state $(15, 15)$ is significantly greater than its value in almost all states where it has a lower market exposure. Therefore, each firm always has great incentives to pursue maximal exposure. In my model with low search costs, the firm’s value

\[16^1\text{There are important differences between D&M’s state variable of awareness and mine, market exposure. But for ease of comparison, I treat them as been equivalent whenever it is not necessary to make a distinction.}\]
is significantly depressed in all states, except for those states where the firm has
disparately higher exposure than its rival. If both firms have relatively high ex-
posure, each recognizes that neither is likely to ever gain a significant advantage.
Given that state (15, 15) is not significantly better than other bad states, firms
have little incentive to advertise heavily.

The presence of an outside firm in D&M also helps to explain this difference
in advertising when both firms already have relatively high exposure. In D&M,
from this position firms still gain from moving up because there is a market-
expansion effect from advertising – the firms steal consumers from the outside
firm. In my model, there is no market-expansion effects from advertising, so the
pricing game in state (1, 1) is equivalent to that of state (15, 15). When both
firms have high exposure, neither anticipates ever becoming the overwhelmingly
dominant firm, and each simply practices maintenance advertising, rather than
aggressively advertising.

2.5.3 Limiting Distribution

Our discussion in section 2.5.2 partially relies on the long-run equilibrium market
structures that evolve for the different levels of search costs. In this section, I show
how we solve for these long-run equilibria. The limiting distribution of a Markov
chain can be interpreted as the long-run equilibrium probability distribution. Let
\( \pi(k, l) \) denote the probability of being in state \((k, l)\) at some randomly chosen time
in the distant future. We can then write the probability distribution, \( \pi \), as a row
vector with \( L^2 \) elements, \( \pi(k, l) \) for \( k = 1, 2, \ldots, L \) and \( l = 1, 2, \ldots, L \). Also, let
\( P \) denote the \( L^2 \times L^2 \) transition matrix. To solve for the limiting distribution I
invoke the well-established theorem that “Every irreducible Markov chain with a
finite state space is positive recurrent and thus has a stationary distribution.” The
unique probability distribution, \( \pi \), that solves the system of \( L^2 \) linear equations,
\( \pi = \pi P \), is that distribution.

The theorem’s first condition of an irreducible Markov chain is satisfied. This
is the case because there is a single communicating class – all states of the game
can be reached from every other state, in a finite number of periods, with a strictly
positive probability. This follows from the fact that the probability that market
exposure depreciates ($\delta$) and the probability of having advertising success ($\theta$) are both strictly positive in all periods and in all relevant states of the game for at least one firm.\footnote{Recall that I assume that $\delta = 0$ if the firm has the minimum exposure level, and that $\theta = 0$ if the firm has the maximum exposure level. These assumptions do not invalidate the statement above, as a firm with minimum exposure can still move up and a firm with maximum exposure can still move down, which is the essential requirement for irreducibility of the Markov chain. The statement that $\theta > 0$ in all relevant states requires that advertising be strictly positive in these states for at least one firm. This is true for the derived policy functions.}

**Figure 2.10:** Limiting Distribution for Different Levels Consumer Search Costs

![Figure 2.10: Limiting Distribution for Different Levels Consumer Search Costs](image)

(a) $b = 40$

(b) $b = 30$

(c) $b = 20$

The result that there is only one communicating class can be easily illustrated with an extreme example. Though highly improbable, it is possible for the game that starts in state $(15, 1)$ to eventually flip and reach state $(1, 15)$. This could occur if firm 2 enjoys a sufficiently long sequence of good fortune, where it has...
more advertising success than market exposure depreciation; while simultaneously firm 1 suffers a similarly long sequence of misfortune.

The theorem’s second condition of finite state space is trivially satisfied. The limiting distributions are illustrated in figure 2.10 for three levels of search costs. Firstly, note that though all states are positive recurrent, only very few states have a significant probability of been visited in the long-run for each level of search costs. In fact, when $b = 40$ only 10 of the 225 states have a probability greater than one-tenth of a percent. In the case where $b = 40$, the long-run equilibrium has two high-exposure firms, with the most likely state being $(15, 15)$, where both firms have maximum exposure. There is a 60% probability of getting to this state in the long-run. Here firms charge the same price, share the market equally, and enjoy relatively high profits.

This symmetric long-run equilibrium with two large firms follows directly from the pricing rule and the profit function when search costs are relatively high. At high search costs levels, a firm has less incentive to charge a significantly lower price than its rival because fewer consumers can afford additional searches. Given that a firm cannot induce many consumers to search again, both firms will attempt to profit as much as possible from those customers that do purchase from it. This creates a less intense pricing environment and the difference in equilibrium prices and profits are therefore considerably smaller at higher levels of search cost. This reduced product market competition allows the market to be able to support two high-exposure, high-priced firms in the long run.

Regardless of the initial state, both firms advertise enough to achieve maximum exposure. Each firm hopes to get to a long-run equilibrium where it is the dominant firm in a highly asymmetric state; but each also recognizes that it would not lose greatly if it were forced into the more likely equilibrium with two high-exposure firms. Note that when search costs fall from $b = 40$ to $b = 30$, the general long-run structure of the market does not change. The limiting distribution still places high probability on symmetric states with two high-exposure firms. As we would expect though, the probability distribution is a bit more dispersed. For example, when $b = 30$, $\pi(15, 15) = 0.57$ and $\pi(15, 14) = 0.16$, compared with $\pi(15, 15) = 0.60$ and $\pi(15, 14) = 0.15$ when $b = 40$.

The long-run equilibrium market structure changes dramatically when $b$ falls
from 30 to 20. At relatively low search costs a highly asymmetric market structure is the most likely long-run outcome. Very little probability mass is placed on symmetric states. The most likely states are jointly (15, 3) and (3, 15), each with a probability of 0.29, and \( \pi(2, 15) = \pi(15, 2) = 0.14 \). In state (3, 15) the dominant firm has a market share of 60% and profits of $495, whereas the low-exposure firm earns profits of $210. The profits in the second most likely long-run equilibrium are even more unequal. The dominant firm has a market share of 68% and profits of $680, whereas the low-exposure firm has profits of $195. When search costs fall low enough, the increased price competition depresses the profits in all states, except for those that are highly asymmetric. Firms therefore aggressively advertise only in a select group of states, which is discussed in sections 2.5.2 and 2.5.4 in greater details. These policy functions guide the market to an asymmetric state in the long run with a very high probability. In short, the reduced search costs creates a price competition environment that cannot support a symmetric equilibrium.

2.5.4 Can falling search cost lead to increased advertising?

The main question that I seek to answer in the paper is whether or not a significant reduction in consumer search costs can cause firms to significantly increase their advertising spending in a market where firms produce a homogeneous product. Given our discussion in section 2.5.2 we can now answer, “Yes, given the right conditions.” The first condition requires that search costs fall within the right range. If the reduction in search costs is too small, such that the market is still able to support two highly profitable firms, the policy function would fall with search costs. Recall that when \( b \) falls from 40 to 30, the policy function falls in such a way that advertising is lower in almost all states. And for the states where this is not true, the increases are insignificant. The reason for this is that value function falls in all states, with very little change in the shape across the states. As such, the same advertising strategy applies in each state, just scaled down.

But if the reduction in search costs is too big, such that profits (and the firm’s value) are negligible in all states, then the policy function would shift down dramatically in all states. A big reduction in search costs would take the market close to the pure Bertrand case, where search costs are so low that all consumers search
until they find the low-price firm. At such a low level of search costs, market exposure is almost irrelevant, so that all firms would advertise very little in all states. For advertising to increase significantly in some states of the game, search costs need to fall low enough for price competition to significantly intensify, but remain high enough so that one firm can enjoy substantial profits.

Assuming the first condition is satisfied, there is still an important second condition: Both firms must have relatively low market exposure and roughly equal sampling probability. Section 2.5.2 outlines the argument that if this condition is not met, firms would simply practice maintenance advertising, rather than increase advertising in an attempt to achieve the dominant position in the market. This is so because in all the other groups of states the probability that the firm would be able to move to one of the few good states is too low, regardless of the advertising spending.

### 2.5.5 Welfare Discussion

The model allows us to analyze the effects of falling search costs on consumer and producer welfare. I will do so by analyzing changes in consumer surplus, producer surplus, and total surplus as the market transitions to a highly asymmetric long-run market structure, as search costs fall to a sufficiently low level. Note firstly, that gross total surplus, which ignores search costs incurred by consumers, does not vary by state. This is the case because I model total demand as perfectly inelastic over the relevant price range. As such, when prices change there is no change in the equilibrium quantity, and hence there is no loss in gross total surplus; only a redistribution of that surplus. Given the chosen parameter values, the gross total surplus is always $1,000, regardless of the state of the game. As the market moves to states with higher share-weighted prices, surplus is simply transferred from consumers to firms. As such, I will focus on the direction and magnitude of this redistribution as search costs fall.

Aggregate firm welfare is measured as the sum of the firms’ profits, which is equivalent to the sum of total revenues, given the assumption that $mc = 0$. I use consumer surplus, net of search costs, to measure aggregate consumer welfare. Consumer surplus is therefore calculated as
\[ CS = [v - (q_1p_1 + q_2p_2)] M - SC, \]

(2.5.1)

where SC, the total search costs incurred by consumers, is given by

\[ SC = \rho_1 G(c_1) M \int_{c_1}^{c_2} cf(c) dc. \]

(2.5.2)

The search costs equation above is derived under two main assumptions: (1) The first search for each consumer is free and (2) The consumer finds the low-price firm (firm 2 in this case) on the second search, if the consumer decides to search again. The second assumption here differs from the assumption maintained throughout the paper, where consumers are assumed to search with replacement and have constant marginal search costs. I make this amendment here to allow for easier calculation of the consumer surplus, and given that this amendment has no qualitative effect on the results. With this modified assumption, the calculated consumer surplus should be interpreted as an upper bound on the actual surplus derived by consumers.

Figure 2.11 illustrates the calculated search costs for each state, given the Nash equilibrium prices. First, we note that in equilibrium there is only a limited amount of search by consumers. At search costs level \( b = 20 \), the maximum search costs incurred by consumers is $6.42, which represents only 0.64% of the total value that consumers derive from purchasing the good. The numbers are even smaller when \( b = 40 \), with maximum search costs of $3.85. This is the expected result, as firms will optimally adjust prices until consumers have little incentive to actually search. By doing so, the firms effectively create more captive consumers for each firm. This result is consistent with observed search behavior in many industries, such as auto insurance, where consumers engage in very limited amounts of search, even though significant differences exist in prices.

In the extreme cases where firms have equal exposure, firms charge equal prices and consumers have no incentive to undertake a second search. Hence, consumers incur zero search costs in states where firms have equal exposure levels. And

---

18 Equation 2.5.2 is written assuming that \( p_2 < p_1 \), and it recognizes that consumers of firm 1 (the high-price firm) pay zero search cost and of firm 2’s customers, only the ones who find firm 1 first, actually incur strictly positive search costs. The corresponding equation for consumer search costs incurred in the states where \( p_1 \leq p_2 \) can be analogously derived.
though consumers incur higher levels of search costs in more asymmetric states, the amounts still reflect only a limited amount of actual searching by consumers. Given the low levels of actual search costs in all states, the total surplus, net of search costs, does not differ significantly from the gross total surplus, which in turn, does not vary by state.

The same is not true for consumer surplus or producer surplus, as surplus moves from consumers to producers as the market becomes more asymmetric. This pattern is evident in figure 2.12, which shows that consumer surplus (producer surplus) is highest (lowest) in states where firms are most equally matched. This is a standard result and the intuition is obvious. Price competition is most intense when firms have equal sampling probability. With Nash prices driven down and consumers incurring very little search costs, consumers enjoy very high net gains in states where firms are most equally matched.

For example, when $b = 40$, consumers enjoy surplus of $200 in states where firms have equal exposure, whereas they receive net benefits of only $12 in states $(15, 1)$ or $(1, 15)$. The drop-off is even greater when $b = 20$, where consumer surplus is $600 in equal-exposure states, compared to a paltry $29 in the most asymmetric states. This is the case as consumers with lower search costs have greater incentives to continue searching at all pairs of prices. Firms therefore have greater incentives to lower price, depressing prices in all states and driving producer surplus down and consumer surplus up.

When we compare surpluses in corresponding states at different levels of search costs, we observe the standard, expected results – a decrease in search costs is unambiguously good for consumers but bad for firms. However, the result is not
necessarily as clear-cut when we compare the long-run equilibrium outcomes for different levels of search costs. Recall that for relatively high search costs environments, the market evolves to a roughly symmetric long-run equilibrium, where consumer surplus is maximized for a given search costs level. However, for low enough search costs, the long-run market structure is highly asymmetric, with a high-exposure firm competing against an over-matched low-exposure firm. Consider the most likely long-run state for each level of search costs: (15, 15) for $b = 40$ and $b = 30$, but state (15, 3) for $b = 20$.

In the long run consumer surplus increases from $200 to $400 when $b$ falls from 40 to 30. This effect is consistent with standard theory. However, consumer surplus falls from $400 to $289 when $b$ falls from 30 to 20. This result that consumer welfare falls as search costs fall appears to be counter-intuitive at first glance or when viewed in the context of traditional search costs-advertising models. However, it is easy to explain when viewed in light of the main results of this paper. In the long run consumer surplus can fall with search costs, if search costs fall within the region, where the market gravitates towards a highly asymmetric long-run structure. The reduction in search costs drives prices down in all states, but the increased price competition actually makes it more likely that one firm will gain a long-term advantage and exploit its ascendancy at the cost of consumers.
In some cases the latter effect dominates, resulting in consumer welfare falling in the long run as consumer search costs fall.

This result has significant policy implications. High consumer search costs create imperfections in markets and confer significant market power to firms. In industries that are plagued with persistent high search costs, a compelling argument can be made for some form of consumer protection. However, as consumer search costs fall, this argument becomes less tenable, and towards the extreme where search costs is close to zero, the case for firm protection may actually gain some traction. The results here suggest that policymakers in industries where consumer protection has been warranted due to high search costs, should be very judicious in the removal of such protection when search costs fall. Given that the market may become highly asymmetric as search costs fall, consumers may be forced to endure long periods of sustained high prices, which come with having a dominant firm in a highly asymmetric state.

2.6 Concluding Remarks

In this paper I seek to answer one question: Can falling search costs serve as a catalyst for increased advertising activity in a market where firms produce a homogeneous product? This question is important because if the answer is yes, then there are significant policy implications about whether consumers or firms should be afforded some type of regulatory protection as search costs continue to fall. I use the modeling framework of Doraszelski and Markovich (2007) to combine advertising choice in a duopoly setting with a sequential search model, similar to that of Carlson and McAfee (1983).

I find that it is possible for falling search costs to cause advertising to increase. This can occur if three general conditions are satisfied: (1) Search costs must fall to a sufficiently low level, such that the intensified pricing competition give rise to a market environment that can support only one highly profitable firm, (2) Search costs should remain high enough, so that competition is not so fierce that firms adopt marginal cost pricing, and (3) Neither firm should have a significant market exposure advantage before the reduction in search costs occurs.

With these conditions satisfied, the market evolves to a highly asymmetric
structure, and firms pursue aggressive advertising campaigns in a typical rent seeking manner. This is a significant result in light of the unmistakable reduction in consumer search costs that has occurred over the last two decades and that is likely to continue for the foreseeable future. For markets that are significantly shaped by consumer search, many researchers and industry experts envisioned moving continuously from a world where consumer protection could be easily justified to a world where policymakers perhaps empathize with firms that barely break-even in increasingly competitive markets.

The main result from this paper suggests that as consumer search costs continue to fall; through advertising, markets possibly will transition to a highly asymmetric structure before becoming highly competitive. This result should be a cautionary note to consumer protection groups, who may blindly welcome decreasing consumer search costs, and to policy-makers, who may too eagerly embrace policies to protect firms against “excessive” competition. One potential for future work in this area is to more closely investigate the likely effects of decreasing search costs under different distributional assumptions. This would likely necessitate allowing the model to admit mixed strategy Nash equilibria in prices. This would build on the work here and would continue to add to our understanding of the interaction between falling search costs, non-price advertising, and market structure.

References


2.A Supplemental Figures and Tables

2.A.1 Pricing and Advertising when $b = 50$

In sections 2.5.1 and 2.5.2, I describe the pricing and advertising behavior of the firms when search costs fall by analyzing the game when $b = 40, 30, \text{ and } 20$. Here, I present the results for $b = 50$ to show that the overall pattern is consistent. For $b = 50$, the firms’ pricing and advertising decisions are in keeping with the discussion in section 2.5.1 for high search costs levels ($b = 30 \text{ and } 40$). Firms continue to charge equal price in symmetric states, whereas the high-exposure firm charges the higher price in asymmetric states. As would be expected, the effect that higher search costs lead to higher Nash equilibrium prices is even more pronounced at $b = 50$ than at $b = 40 \text{ or } 30$. For example, the share-weighted average price across states of $96.80$ is strikingly high when compared to the monopoly price of $100$. But though this average may appear high, it is to be expected when $b = 50$. With extremely high search costs, even huge disparities in prices, do not induce many consumers to search again, which significantly reduces the incentives of both firms to aggressively compete on prices. This ultimately results in all states having relatively high Nash prices. In fact, when $b = 50$, a firm will charge the monopoly price in all states where it has equal exposure or greater.

This pricing behavior translates into the profit and value functions for $b = 50$ lying above those for $b = 40$ in all states. Again, consistent with the discussion in section 2.5.2, the drop-off in the firm’s value for being in a bad state, relative to
being in a good state, is significantly less at higher values of $b$. For example, firm 1 in its worst state, (1, 15), has a reasonably high value, earning 78.3% of its already high value in its best state, (15, 1). The effect on advertising, as we saw with other cases of high search costs, is that both firms advertise moderately in all states, so that both firms almost always attain the maximum exposure level. Advertising decreases with the firm’s exposure level and the advertising function lies above the function for $b = 40$ for all states. This is consistent with the standard theory, where higher search costs create incentives for firms to advertise more.

2.A.2 Simulation Results when $b = 40$ and 30

In section 2.5.2 I describe the advertising behavior of firms based on a representative simulation of the game when $b = 20$. Here, for comparison purposes, I
present the high search costs cases, $b = 30$ and 40. Recall that when $b = 30$ or 40, we obtain a symmetric long-run equilibrium with two high-exposure firms. Given that the game spends the vast majority of time in the state where both firms have maximum exposure, I choose the most asymmetric state, $(15, 1)$, for the initial state.

**Figure 2.A.2:** Simulations with $b = 40$ and Initial State, $(15, 1)$

For both $b = 40$ and $b = 30$, the high-exposure firm practices maintenance advertising, where it ramps up advertising activity when its exposure level falls, but advertises very little otherwise. The low-exposure firm advertises enough initially to almost guarantee advertising success in each period. The firm’s advertising spending then tapers off as its exposure level increases, but the firm continues to maintain high enough advertising until it achieves close to maximum exposure level. Once the low-exposure firm catches up with the high-exposure firm, it starts to practice maintenance advertising. Figures 2.A.2 and 2.A.3 illustrate that even starting from state $(15, 1)$, where firm 2 is at a great disadvantage, it still advertises enough to achieve maximum exposure within roughly 30 periods when $b = 30$ or 40. Contrast this with the game when $b = 20$, where a firm with even a slight exposure disadvantage would simply try to maintain its exposure level, rather
than aggressively compete in the market exposure dimension.

**Figure 2.A.3:** Simulations with $b = 30$ and Initial State $(15, 1)$

---

2.A.3 Summary Results Tables: Value and Policy Functions for $b = 40$, 30, and 20

Table 2.A.1 summarizes the results for the value function for the three levels of search costs that are analyzed through-out the paper ($b = 40$, 30, and 20). A key finding of the paper shows up in this table – the value of the firm falls in all states when search costs fall from $b = 40$ to $b = 30$. However, when there is a further reduction in $b$ from 30 to 20, though the value of the firm still significantly falls in most states, the firm’s value actually increases dramatically in highly asymmetric states, for the firm with the exposure advantage.

To illustrate, the mean of the firm’s value in high-exposure states relative to the mean in low exposure states is marginally greater than 1, at 1.09 and 1.10 when $b = 40$ and 30, respectively. That same ratio jumps to 2.11 when $b = 20$. Similarly, the firm’s maximum value is almost 3-times its minimum value when $b = 20$, compared to roughly only 1.33-times when $b = 40$ or 30. This reiterates a
point made in section 2.5.2. When search costs are sufficiently low, they create a competitive landscape with a very small number of states that are extraordinarily profitable, alongside many states that provide roughly equal, but meager returns.

Table 2.A.1: Summary Results for Firm 1’s Value Function

<table>
<thead>
<tr>
<th></th>
<th>$b = 40$</th>
<th>$b = 30$</th>
<th>$b = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum (State)</td>
<td>23,032</td>
<td>18,301</td>
<td>28,823</td>
</tr>
<tr>
<td>Minimum (State)</td>
<td>17,802</td>
<td>13,748</td>
<td>10,005</td>
</tr>
<tr>
<td>Max-Min Ratio</td>
<td>1.29</td>
<td>1.33</td>
<td>2.88</td>
</tr>
<tr>
<td>Mean</td>
<td>19,980</td>
<td>15,124</td>
<td>15,547</td>
</tr>
<tr>
<td>Median</td>
<td>19,751</td>
<td>14,799</td>
<td>10,262</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>1,139</td>
<td>1,004</td>
<td>7,522</td>
</tr>
<tr>
<td>Mean for High-Exposure States</td>
<td>20,894</td>
<td>15,882</td>
<td>21,510</td>
</tr>
<tr>
<td>Mean for Low-Exposure States</td>
<td>19,099</td>
<td>14,413</td>
<td>10,187</td>
</tr>
<tr>
<td>Mean for Equal-Exposure States</td>
<td>19,750</td>
<td>14,799</td>
<td>11,332</td>
</tr>
<tr>
<td>High-low Means Ratio**</td>
<td>1.09</td>
<td>1.10</td>
<td>2.11</td>
</tr>
<tr>
<td>High-equal Means Ratio</td>
<td>1.06</td>
<td>1.07</td>
<td>1.90</td>
</tr>
<tr>
<td>Low-equal Means Ratio</td>
<td>0.97</td>
<td>0.97</td>
<td>0.90</td>
</tr>
</tbody>
</table>

** High-low means ratio represents the fraction \( \frac{\text{Mean for High-exposure States}}{\text{Mean for Low-exposure States}} \). The high-equal means ratio and low-equal means ratio are similarly defined.

Table 2.A.2 presents some of the key results for the policy function when the search costs distribution parameter $b = 40, 30, and 20$. The main highlight is that advertising activity decreases when $b$ decreases from 40 to 30, but increases when $b$ decreases from 30 to 20. This can be seen in the overall means for the different values of $b$, but it shows up more clearly in the states where firms advertise aggressively. In keeping with section 2.5.2, I classify advertising as aggressive when a firm’s advertising-to-sales ratio is greater than 10%. We see that the mean advertising spending in states where the firm advertises aggressively when $b = 20$ is almost twice as much as when $b = 40$ or 30. The more general and more important point is that when search costs are sufficiently low, the firm either advertises heavily or a minuscule amount, dependent on the state of the game; whereas at higher levels of search costs, the firm advertises more moderately in all states. This is consistent with the results from the value function, which has a
similar gaping disparity between the good states and the bad states for a firm.

**Table 2.A.2: Summary Results for Firm 1’s Policy Function**

<table>
<thead>
<tr>
<th></th>
<th>(b = 40)</th>
<th>(b = 30)</th>
<th>(b = 20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum (State)</td>
<td>21.01</td>
<td>18.43</td>
<td>70.84</td>
</tr>
<tr>
<td>Minimum (State)</td>
<td>1.17 (15, 15)</td>
<td>0.95 (15, 15)</td>
<td>0 (3, 7), (7, 15)†</td>
</tr>
<tr>
<td>Mean</td>
<td>10.62</td>
<td>8.76</td>
<td>12.01</td>
</tr>
<tr>
<td>Median</td>
<td>9.93</td>
<td>7.93</td>
<td>5.66</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>5.11</td>
<td>4.12</td>
<td>16.18</td>
</tr>
<tr>
<td>Mean for Aggressive Advertising States</td>
<td>19.06</td>
<td>15.63</td>
<td>36.24</td>
</tr>
<tr>
<td>Mean for Non-Aggressive Advertising States</td>
<td>9.79</td>
<td>8.31</td>
<td>6.77</td>
</tr>
<tr>
<td>Aggressive Ad/Non-aggressive Ad (Means)</td>
<td>1.95</td>
<td>1.87</td>
<td>5.35</td>
</tr>
</tbody>
</table>

† When \(b = 20\) the low-exposure firm has little to no incentive to advertise in states where it is significantly far behind the high-exposure firm. And the value of the firm is similarly depressed in these states, so that the firm optimally chooses not to advertise in many of them. This is true for firm 1 in all states \((k, l)\), where \(3 \leq k \leq 7\) and \(l - k \geq 3\). There are 39 such states.

Table 2.A.3 provides information about the market structure that arises in the state that is most likely to occur based on the limiting distribution. I present results for \(b = 40\), \(b = 30\), and \(b = 20\). Table 2.A.3 focuses attention on the likely asymmetry that arises when search costs fall to a sufficiently low level, \(b = 20\). This stands in stark contrast to the symmetric long-run equilibrium that is realized at the higher search costs levels, \(b = 40\) and \(b = 30\). We observe that the most likely state to occur are \((15, 15)\), \((15, 15)\), and \((15, 3)\) for search cost levels \(b = 40\), \(30\), and \(20\), respectively. The disparity in the outcome shows up not simply in the levels of exposure but in all of the variables that depend on the equilibrium state of the game. For example, when \(b = 20\) in the limiting state, the high-exposure firm is 5-times more likely to be visited than the low-exposure firm; controls 60% of the market; and has a lifetime value that is about 3-times greater than that of the low-exposure firm.
Table 2.A.3: Limiting Distribution Market Structure

<table>
<thead>
<tr>
<th></th>
<th>( b = 40 )</th>
<th>( b = 30 )</th>
<th>( b = 20 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limiting State(^\ddagger)</td>
<td>(15, 15)</td>
<td>(15, 15)</td>
<td>(15, 3)</td>
</tr>
<tr>
<td>Limiting Probability</td>
<td>0.60</td>
<td>0.57</td>
<td>0.29</td>
</tr>
<tr>
<td>Sampling Probabilities</td>
<td>(0.5, 0.5)</td>
<td>(0.5, 0.5)</td>
<td>(0.82, 0.18)</td>
</tr>
<tr>
<td>Nash prices</td>
<td>(80, 80)</td>
<td>(60, 60)</td>
<td>(82, 53)</td>
</tr>
<tr>
<td>Market Shares</td>
<td>(0.5, 0.5)</td>
<td>(0.5, 0.5)</td>
<td>(0.6, 0.4)</td>
</tr>
<tr>
<td>Share-weighted Average Price</td>
<td>80</td>
<td>60</td>
<td>71</td>
</tr>
<tr>
<td>Per-period Profits</td>
<td>(400, 400)</td>
<td>(300, 300)</td>
<td>(495, 210)</td>
</tr>
<tr>
<td>Producer Surplus</td>
<td>800</td>
<td>600</td>
<td>705</td>
</tr>
<tr>
<td>Consumer Surplus</td>
<td>200</td>
<td>400</td>
<td>289</td>
</tr>
<tr>
<td>Search Costs Incurred</td>
<td>0</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Firm Valuations</td>
<td>(19,767; 19,767)</td>
<td>(14,817; 14,817)</td>
<td>(28,214; 10,246)</td>
</tr>
<tr>
<td>Advertising Expenditures</td>
<td>(1.17, 1.17)</td>
<td>(0.95, 0.95)</td>
<td>(2.21, 0)</td>
</tr>
</tbody>
</table>

\(^\ddagger\) I use the term limiting state to refer to the state with the highest probability in the limiting distribution at each search costs level. The values in each column relate to the respective limiting state for \( b = 40, 30, \) and 20. The first element in a pair of values refers to firm 1 and the second refers to firm 2.

2.B Distributional Assumptions and the Existence of a Pure Strategy Nash

Throughout the paper I maintain the simplifying assumption that search costs are uniformly distributed over the interval \([0, b]\). I do not make this assumption for purposes of tractability, because though closed-form solutions are useful when available, they are unessential when numerical solutions are being used, as is the case here. Instead, I assume search costs are uniformly distributed because a pure strategy Nash equilibrium in prices is not guaranteed for other distributional assumptions. Here, I illustrate this point by assuming that search costs are log-normally distributed, with scale parameter, \( \mu \), and shape parameter, \( \sigma \). Specifically, I choose \( \mu = 1 \) and \( \sigma = 0.5 \). For ease of reference, the cumulative distribution function (CDF) and probability density function (PDF) are shown in figure 2.B.1. Of particular interest, note that the CDF reaches approximately 0.995 when \( c = 10 \), and that the PDF is maximized at approximately \( c = 2 \).
Figure 2.B.1: CDF and PDF when $c \sim \ln \mathcal{N}(1, 0.5)$

To get a good understanding of the non-existence of a pure strategy Nash equilibrium in some states under the assumption that $c \sim \ln \mathcal{N}(1, 0.5)$, it is necessary to first understand the shapes of the demand and profit functions given this distributional assumption. Recall that, assuming $p_2 < p_1$, the demand functions are

1. $q_1 = \rho_1 - \rho_1 G(\rho_2(p_1 - p_2))$

2. $q_2 = \rho_2 + \rho_1 G(\rho_2(p_1 - p_2)).$

And the slope of the demand functions is given by $\frac{\partial q_j}{\partial p_j} = \rho_1 \rho_2 g(\rho_2(p_1 - p_2))$. Figure 2.B.2 shows the demand function and profit function for firm 1 that arise under the assumptions that $c \sim \ln \mathcal{N}(1, 0.5)$, firms have equal sampling probabilities, and that firm 2 charges a price of $18.88.

It is obvious that when the price gap is zero, the firms share the market equally. From this position, if $p_1$ increases, firm 1 becomes the worse firm and consumers with sufficiently low search costs, who would have purchased from firm 1 initially, now continue to search until they find firm 2. As the price gap widens, firm 1 loses consumers to firm 2 at a rate determined by the density function of the search costs distribution. Given the assumed distribution function, the density function is first concave upwards, so that as $p_1$ increases, firm 1’s market share falls at an increasing rate, at first. There is less mass in the upper tail of the density function, and so for a wide enough price gap, the reduction in market share is negligible, but so too is firm 1’s overall market share at this point.
Figure 2.B.2: Demand and Profit Functions when $c \sim \ln \mathcal{N}(1,0.5)$, $\rho_1 = 0.5$ and $p_2 = $18.88.

Given the symmetry of the game, the portion of firm 1’s demand function with $p_1 < p_2$ is the same as firm 2’s demand function with $p_2 < p_1$. For example, from the point where $p_1 = p_2$, $q_1$ falls by roughly 0.012 when $p_1$ increases by $\$2$. This reduction in $q_1$ matches the increase that would result if firm 1 were to have reduced its price by the same $\$2$. It is also important to note that when the price gap is roughly within this range, the demand function is markedly flatter. Recall that the search costs density function approaches zero when the price gap shrinks towards zero. Therefore, charging a slightly higher or slightly lower price than its competitor has very little impact on the firm’s market share. Given that demand is very inelastic when the two prices are close, a firm’s total revenues (and profits) move in the same direction as the price change, for a small enough price change. The density function increases with search costs, which implies that the firm loses (gains) consumers at an increasing rate as the firm continues to increase (reduce) its price.

The profit function eventually gets to a local maximum if the firm continues to increase its price or eventually gets to a local minimum if the firm continues to decrease its price. This relationship leads to the profit function having two local maximums on either sides of the price gap of zero. Which of these is the global maximum depends on the price charged by the other firm. And as we would expect, when firms are otherwise homogeneous, a firm’s profits are lowest when it charges a significantly different price from its competitor.
The best response functions are depicted in figure 2.B.3. They show that in the case where sampling probabilities are equal, firm 1 wants to be the high-price firm if \( p_2 \leq \$18.88 \), but wants to be the low-price firm if \( p_2 > \$18.88 \). The symmetry in the game allows us to make a similar argument for firm 2. And each firm wants a price gap of at least \$2. The intuition for this is obvious. If firm 2 is charging a relatively high price, then firm 1 can undercut \( p_2 \) and induce enough consumers to continue to search for firm 1, without sacrificing too much on its profit margin. On the other hand, if firm 2 is charging a relatively low price, then there is enough room for firm 1 to increase its price such that it gains more from the increase in the profit margin than it loses from the reduction in market shares.

As is illustrated by the best response functions in figure 2.B.3, no pure strategy Nash equilibrium exists when firms have the same sampling probabilities and the same marginal costs, and consumers search sequentially, with \( c \sim \ln N(1, 0.5) \). In this case where firms have equal sampling probability, the reason for the non-existence of a pure strategy Nash is easy to pinpoint. Given the type of strategic interaction and the symmetry of this game, a pure strategy Nash requires that firms charge equal prices if they have equal sampling probability. But given the assumed
distribution function, firms cannot choose the same price in a pure strategy Nash equilibrium. Such an equilibrium requires that a slight unilateral increase in price would cause the firm to lose enough consumers to the low-price firm, so that the firm would have no incentive to increase its price.

This in turn requires that the density function be sufficiently high when the price gap is very small. Recall that the density function approaches zero as the price gap shrinks towards zero. As such, when firms have equal sampling probability, a firm always does better by charging a slightly higher price than its competitor. The constant density function for the uniform cumulative distribution function solves this problem. Similar results can be shown for different values of $\mu$ and $\sigma^2$, different sampling probability combinations, and for different families of search costs distribution. The important take-away is that a pure-strategy Nash equilibrium is not guaranteed for distributions other than the uniform distribution, the distributional assumption that I maintain throughout the paper.
3.1 Introduction

The last decade has witnessed the near extinction of bounced checks.\(^1\) The majority of banks, especially the larger ones, now use automated overdraft programs to authorize the processing of many more non-sufficient funds transactions (NSF). An automated overdraft program is any computerized system that authorizes the payment of an overdraft, based on a minimal set of criteria, at a fixed fee.\(^2\) The 2008 FDIC Study of Bank Overdraft Programs reports that 76.9 percent of large banks had an automated overdraft program. It also reports that 74 percent of the $2.66 billion in service charges on deposit accounts resulted from NSF-related fees in 2006.

Banks suggest that this service has been extended as a courtesy to their consumers, who would otherwise be adversely affected by bounced checks. However, many consumers and consumer affairs groups argue that banks are using courtesy overdrafts merely to increase fee income at the expense of, especially young, low-

\(^1\) Bounced checks refer specifically to written checks that have been returned to the payer without being processed due to an insufficient checking account balance. Here I use the term more broadly to describe any transaction, including point of sale purchases, that has been declined due to insufficient funds.

\(^2\) Though tiered pricing is common, I ignore this feature in the discussion below.
income, households. Their contention is that a large fraction of consumers would actually prefer if the transactions were denied rather than to incur the “exorbitant” flat fee for each overdraft.

Limited formal work has tackled this problem directly. Gabaix and Laibson (2006) (hereafter, G&L) propose a model that explains the existence of industries where firms shroud information about add-on goods and price them above marginal cost. In their model banks are able to do so if a sufficiently large fraction of consumers are myopic in the sense that they do not consider these overdraft services when choosing the base product. The authors show that in this equilibrium no firm has an incentive to “educate” the myopic consumers partly because it causes the pool of potential consumers to become less profitable.

G&L fits into the strand of literature that addresses the monopolization of aftermarkets. However, one of its main results runs contrary to well-established theory that in relatively competitive markets, firms have an incentive to provide consumers with all salient information about their products. Given this conflict, it is important that we empirically substantiate or refute the findings of G&L.

I choose to conduct this investigation using overdraft service because it appears consumers are less knowledgeable about this product and because of the recent changes to how banks process overdraft services.

Additionally, it would be imprudent of us to overlook overdraft services based simply on the fact that only a small proportion of consumers use this product. If Gabaix and Laibson’s hypothesis is correct, then the existence of this small fraction of consumers has significant effects on the overall functioning of the industry and the welfare implications are not restricted to only those who use overdraft services. This appears to be at least part of the reasoning behind the recent proposals by the Federal Reserve that requires all banks to, among other things, obtain affirmative consent from consumers before charging overdraft fees. We expect that this proposal, which becomes effective July 1, 2010, will significantly influence not only how banks price overdraft services but also how they price other add-ons and checking accounts.

3 Based on a non-random sample, the 2008 FDIC Study of Bank Overdraft Programs reports that 84 percent of all NSF-related fees were charged to 4 percent of all accounts.

4 This result of G&L was also established by Ellison (2005) in a more general framework. See, for example, Shapiro (1995) for an apposing view.
To date, empirical work on overdraft fees has mainly been reduced-form analyses that investigate the main determinants of the fee. However, these regressions are usually devoid of a cogent theoretical foundation. In this paper I attempt to test the implications of Gabaix and Laibson’s model using six years of repeated cross-sectional data on the overdraft fees charged by U.S. banks and credit unions.

Specifically, I attempt to establish whether the presence of myopic consumers is a significant factor in the pricing of overdraft fees by U.S. banks and credit unions. I observe the overdraft fees charged by a representative sample of firms in different geographic markets over six years. Importantly, I also observe the relevant characteristics of consumers that are served by these firms in each market. Using age, income, and education to proxy for the fraction of myopic consumers, I find that age and income significantly influence overdraft fees in the direction predicted by G&L.

The remainder of the paper proceeds as follows. Section 3.2 provides a brief description of the theoretical model and its testable implications. Section 3.3 presents the regression model, while the key features of the data are described in section 3.4. Section 3.5 discusses the main findings, and concluding remarks are provided in section 3.6.

### 3.2 A Model of Add-on Pricing: Gabaix and Laibson (2006)

Gabaix and Laibson put forward an explanation for why firms would choose to shroud relevant product information even in competitive markets with little or no cost of disseminating such information. In so doing the authors provide a reasonably compelling pricing theory for a particular class of products - complementary products, such as add-on goods and after-sale services, that are demanded subsequent to purchasing some base good. G&L suggest that the model applies directly to information suppression and pricing decisions of certain services that accompany bank accounts. My aim here is to test the applicability of this model to the pricing of overdraft services in the banking industry.

---

5 See, for example, Hannan (2006).
The key assumption of G&L is that a strictly positive fraction of the consumers are “myopic” as opposed to being “sophisticated.” They describe myopic consumers as those that do not consider the cost of add-ons, or the cost of avoiding add-ons, when purchasing the base goods. In our context, the base good is a checking account, whereas the add-on good is overdraft services. A brief summary of the timing of model and the optimization problem is given below. I will discuss the extension of the base model that considers the case of continuous add-on demand.

**Period 0**

Having observed the fraction of consumers that are myopic, $\alpha$, each firm chooses whether or not to shroud the price of the add-on. G&L defines a shrouded attribute as “a product attribute that is hidden by a firm, even though the attribute could be nearly costlessly revealed.” Shrouding can thus be interpreted as the case where firms ensure that the cost of learning the price of the add-on is sufficiently high such that no consumer has incentive to attempt to learn the add-on price. Each firm then chooses the price of the base good ($p$) and the price of the add-on good ($\hat{p}$) simultaneously. Note that firms cannot choose prices that are consumer type-specific.

**Period 1**

Each consumer chooses to open a checking account with the firm that gives him or her the highest expected utility. Having chosen a bank the consumer then decides on the level of effort to exert in monitoring his/her checking account balance. The choice of effort level also includes choosing the level of precautionary balance to maintain. Increased effort reduces the probability of overdrafting. The optimal effort level for a consumer depends on the price of overdrafts, which if shrouded, the consumer views as being drawn from a known distribution. The expected utility of a sophisticated consumer from purchasing the base good of firm $i$ is given by

$$U^s_i = u_i - p_i - \min_{e_i} E_{\tilde{p}_i} \{ pr(e_i) \hat{p}_i + c(e_i) \}$$  \hspace{1cm} (3.2.1)$$

where $u_i - p_i$ is the net surplus from purchasing the base good from firm $i$ if there
is zero probability of overdrafting. \( e_i \) is the effort level exerted if firm \( i \)'s base good is chosen. \( \text{pr}(\cdot) \) and \( c(\cdot) \) are the probability of overdrafting function and the cost of effort function, respectively. The probability function is negatively related to the effort level, while the cost is a positive function of the effort level. Note that the optimal effort level will vary across firms only if the price of the add-on varies. It is necessary to consider expected add-on prices since firms may choose to shroud these prices. In such a case sophisticated consumers are assumed to form rational expectations.

Assume that there is a unique solution, \( e^*_i \), to the sophisticated consumer’s minimization problem such that the indirect probability function can be written as \( \text{pr}^S(\hat{p}_i) \equiv \text{pr}(e^*_i(\hat{p}_i)) \) evaluated where \( E[\hat{p}_i] = \hat{p}_i \). If firms choose not to shroud the price of the add-on, a fraction of the myopic consumers, \( \lambda \in (0, 1] \), become informed. The behavior of informed myopes is identical to that of sophisticated consumers.

An uninformed myope, on the other hand, chooses a firm to maximize \( u_i - p_i \). G&L describes this situation as myopes passively choosing a default level of effort, which results in a baseline probability of overdrafting, \( \text{pr}^M \).

**Period 2**

All shrouded prices are revealed. Demand for the add-on by sophisticates and myopes are realized with probabilities \( \text{pr}^S(\hat{p}_i) \) and \( \text{pr}^M \), respectively.

**Equilibrium behavior**

Assuming that firms know the probability and cost of effort functions of the sophisticated consumers, each firm chooses the monopoly add-on price, \( \hat{p}^m \), by

---

\(^6\) Here I depart from G&L slightly by writing the payoff from consuming the add-on strictly as an expected cost. In this version of the model the sophisticated consumer chooses effort level to determine the probability of overdrafting based on the announced or expected price of the add-on good. Demand for the add-on good is then realized in Period 2 based on this probability. In G&L the consumer reoptimizes in period 2 by choosing the optimal quantity of the add-on given his or her effort level and the, by then, revealed add-on price. Thinking of the demand for overdraft services in a probabilistic sense is a more natural interpretation.
solving the following problem:

\[
\max_{\hat{p}} (\hat{p} - c)(1 - \alpha)\text{pr}^S(\hat{p}) + \alpha\text{pr}^M
\]

G&L discuss symmetric equilibria. Define the average probability of overdrafting as

\[
\hat{\text{pr}}(\hat{p}) \equiv (1 - \alpha)\text{pr}^S(\hat{p}) + \alpha\text{pr}^M
\]

The overall profit function for a firm that charges add-on price \(\hat{p}^m\) and base price \(p\) when all other firms charge \(\hat{p}^s\) and \(p^s\), respectively, is given by

\[
\Pi = (p + (\hat{p}^m - c)\hat{\text{pr}}(\hat{p}^m))D(p^s - p) \tag{3.2.2}
\]

where \(D(x)\) is the demand of a firm that offers an average perceived surplus of \(x\) greater than the average perceived surplus of next best competitor. G&L show that under relatively weak conditions a unique symmetric equilibrium exists where the average mark-up per consumer is given by

\[
p - c = \frac{D(0)}{D'(0)} \equiv \mu^7 \mu \text{ thus represents a measure of competition, with } \mu = 0 \text{ when the base good market is perfectly competitive. I seek to test the implications of the following proposition proved by G&L in Appendix 2.}

**Proposition:** Suppose that unshrouding makes all consumers sophisticated (\(\lambda = 1\)). Then the price vector

\[
p = \mu - (\hat{p} - c)\hat{\text{pr}}(\hat{p})
\]

\[
\hat{p} = \hat{p}^m
\]

supports a Shrouded Prices Equilibrium if and only if \(B \geq 1\), where \(B\)

---

7 We assume that there is sufficient concavity in the probability function to guarantee a unique internal solution. Otherwise, we revert to G&L’s base model, which considers discrete demand for the add-on. In the *shrouded equilibrium* of that model, firms charge the exogenously given maximum price for the add-on.

8 Additional details of the microfoundations of the demand function, \(D(x)\), is presented in Appendix 1 of G&L. Also, without loss of generality the marginal cost of the base good is assumed to be zero.
is the debiasing ratio:

\[
B = \frac{\text{cross subsidy to sophisticates from myopes}}{\text{loss of social surplus (for sophisticate demand) due to add-on mark-up}}
\]

\[
= \frac{\alpha(\hat{p}^m - \hat{c})(pr^M - pr^S(\hat{p}^m))}{pr^S(\hat{p}^m)\hat{p}^m - pr^S(\hat{c})\hat{c}}
\]

Given reasonable functional forms for the probability of overdrafting and cost of effort functions, G&L provide a number of empirically testable predictions. (1) Shrouding of information, specifically the price of add-ons, is more likely to occur in markets with a relatively larger fraction of myopic consumers. (2) The absolute difference between the price of the add-on and the price of the base good should be larger in markets with a relatively larger fraction of myopic consumers. (3) G&L also conjecture that a fraction of myopic consumers is likely to learn over time, especially those who repeatedly consume the add-on good. If in fact myopic consumers are important, we would expect to see a downward trend in the price difference as consumers gain experience with the good.

I test the latter two implications by using market-level consumer demographic characteristics to proxy for the fraction of myopic consumers, \(\alpha\), since \(\alpha\) is unobservable to us. Each proxy is chosen because it can be plausibly linked to the fraction of myopic consumers and because it is available. I choose as proxies the following city-specific demographic characteristics: median age, median income, and the fraction of the population above age twenty-five (inclusive) that has at least a high school diploma or its equivalent.

Ideally, I would like to directly examine the behavior of consumers to determine whether this dichotomous classification of consumers as myopic or sophisticated can be inferred from their behavior. This would involve an analysis of the consumer’s choice of bank, choice of overdraft protection, the consumer’s spending patterns, and most importantly the consumer’s demand for overdraft services over time. Note that the time dimension is critically important since it would allow us to evaluate whether or not consumers learn and consequently adjust their behav-

---

9 The model reasonably assumes that firms perfectly observe \(\alpha\). Financial institutions are likely to have a fairly accurate estimate of the fraction of myopic consumers based on historical banking data that is unavailable to us.
ior over time. The data on hand, however, does not allow me to conduct such an analysis.

Additionally, a complete test of G&L would require empirically testing implication (1) above. This would, however, require that I identify the equilibrium, shrouded or unshrouded, that is being played in each market. One means of doing so would be to estimate a switching regression model. However, this and other methods of distinguishing between the equilibria proved infeasible given the available data.

There is significant value in the empirical exercise that is undertaken. A sound theoretical model forms the basis of our reduced-form analysis. Also, the results of this exercise should provide additional guidance for further theoretical work and structural empirical examination of aftermarkets with the appropriate data. Admittedly, the proposed test is not foolproof. Failure to find evidence for the importance of myopic consumers may simply be a result of poor proxies. Secondly, if we find support for Gabaix and Laibson (2006) in the sense that our proxies for the fraction of myopic consumers have the required coefficients, it does not preclude the validity of other explanations. Our results should therefore, be interpreted in light of these limitations.

3.3 The regression model

The goal of the empirical model is to explain the variation in overdraft fees across banks and credit unions with a combination of variables that measure the fraction of inexperienced or uninformed consumers in the geographic area served by the bank. We will also control for observable characteristics of banks and credit unions that measure heterogeneity in product quality and cost.

I estimate the following regression equation using OLS:

$$Y_{int} = \alpha_0 + \beta X_{int} + \zeta Z_m + \delta D_t + \eta M_{mt} + \epsilon_{int}$$

where $Y_{int}$ is the real overdraft fee charged by firm $i$ in city $m$ in year $t$. $X_{int}$ is a vector of firm specific characteristics that are allowed to vary over time and across branches (cities). These include firm type (bank or credit union), the size
of firm, the availability of overdraft protection options, and whether or not checking accounts are free. We assume that firms choose the price of overdraft after having decided on what overdraft protection options to offer and also after having determined the price of opening a checking account.

$Z_m$ is a vector of time-invariant demographic variables, which are used to proxy for the fraction of myopic consumers in each market. $D_t$ is a vector of year dummies. $M_{mt}$ is a vector of market-specific variables, which include the Herfindahl-Hirschman Index (HHI), city size, and state dummies. $\epsilon_{int}$ is the error term, which is assumed to be iid across markets, years, and firms, has mean zero, and constant variance. The vectors of parameters to be estimated are $\alpha_0, \beta, \zeta, \delta,$ and $\eta$. The expected relationships are discussed along with the results in section 3.5.

### 3.4 Data description

I utilize bank and credit union data on overdraft fees from a repeated cross section of branches conducted by Moebs $\text{ Services Inc.}^{10}$ Each cross section is obtained by using a stratified sampling procedure to reflect the population in terms of firm size and regions of the country. The data set covers the six-year period 2004 to 2009 and provides 7,903 total observations on 5,206 unique city-bank pairs$^{11}$.

I do not, and in many instances cannot, distinguish between different branches of the same firm that are located in the same city. This makes it infeasible to use a finer market definition than a city. Additionally, this definition seems a reasonable approximation since I find very little variation in overdraft fees within a city in a given year for the same firm$^{12}$. On the other hand, even within the same state there is significant variation in the fees for a firm across cities, and so I choose not to aggregate up to the state level.

I would have preferred to have a panel, which would allow for the inclusion

---

10 Moebs $\text{ Services Inc.}, an economic research firm that started in 1983, focuses on collecting and analyzing primary pricing and expense data for financial institutions. Moebs Inc. is an integral player in the banking industry as it provides consultation services to financial institutions in creating or modifying their fee structures and other features such as automatic overdraft services.

11 The survey on overdraft fees dates back to 1995 but data on some crucial variables is unavailable pre-2004.

12 The data provides a few instances where I am able to observe multiple branches with the same firm-city-year triple.
of firm fixed effects. This is, however, not possible in our data set given that 64 percent of the city branches were observed only once. Also, though the main objective is to explain cross sectional variation, I retain the time dimension and include time dummies in order to indirectly test for learning effects.

Table 3.1 presents summary statistics on overdraft fees along with the three demographic variables. A median overdraft fee of $26.60 is curious when compared to the following statistic. Based on a non-random sample, the 2008 FDIC Study of Bank Overdraft Programs concludes that the median transaction size of a check that resulted in an NSF transaction was $66. At these medians, if the total amount of $92.60 is paid within a month, then the consumer actually pays an APR of 483.64 percent.\textsuperscript{13}

The variation in our dependent variable is considerable; 15 percent of the branches in our sample charged a fee of at most $20, while 25 percent charged at least $30. This variation is especially noteworthy given the fact that there is little or no differentiation in overdraft services across firms. Variation across years is minimal, and hence I focus on trying to explain the significantly more important variation across geographical markets.

The sample is representative of the size distribution of the population of banks with only 9 percent of observations coming from large firms, while over 50 percent are of small firms. This is important since there is a noticeable difference of over $3 in the mean overdraft fees charged by large firms as compared to small firms. Similarly, the mean fee of banks are about $3 greater than that of credit unions. Two-thirds of our sample are of branches that offer free checking account to consumers, while the remaining branches charge a positive price for this product. This information allows us to include a reasonably good control variable for the price of the base good, which we do not observe. It is also worth noting that in over 95 percent of the observations the branch offered at least one of the two overdraft protection options that are of interest to us.

We augment the Moebs data set with demographic data from the 2000 Decen- nial Census conducted by the U.S. Census Bureau. I utilize three demographic characteristics: median age, median income, and the fraction of the population,\textsuperscript{13} Technically, an overdraft is not considered a loan, and so there is no requirement for banks to report APR for this product.
twenty-five and older that has a high school diploma or its equivalent. These choices are based on the assumption that they strongly correlate with the fraction of myopic consumers in a geographic market.

It was my intent to use annual city level demographic data to complement our repeated cross section on banks and credit unions. However, the most appropriate source of such data, the American Community Survey, is conducted only for census areas with populations in excess of 65,000. Confining our analysis to these cities would significantly reduce our sample size and introduce obvious sample selection issues. Also, I am reasonably comfortable using a time-invariant measure of the demographic characteristics over the six-year period since large relative cross sectional changes in demographic characteristics are unlikely over a short period of time.

Summary statistics are provided for the three demographic variables in table 3.1. Similar to the overdraft fees, we find that there is considerable variation in these variables. The median household income ranges from a low of $16,000 to a high of $283,000. Twenty percent of the cities have a median age of forty or more, while six percent have a median age of twenty five or less.

The Herfindahl-Hirschman Index (HHI) is used to control for differences in the level of competition across cities. This index is not very demanding on data and there is no clear advantage in using alternative measures of concentration. The HHI is calculated for each city-year pair using FDIC data on the deposits of bank branches.

### 3.5 Results

Table 3.1 reports the results of regression Equation 1. The table contains information on variables that have been retained based on significance levels or compelling economic reasons why they should be included in the regression. My interest lies mainly in the demographic variables that jointly serve as a proxy for the fraction of myopic consumers, \( \alpha \). However, I report and discuss a number of important control variables that were included.

---

14 For earlier years the lower bound on population for surveyed census areas was actually 250,000. Some special cases of areas with at least 65,000 were also included during this period.
Table 3.1: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overdraft fee ($)</td>
<td>25.98</td>
<td>26.60</td>
<td>1.03</td>
<td>45.43</td>
<td>5.75</td>
</tr>
<tr>
<td>Median Age</td>
<td>35.26</td>
<td>35.00</td>
<td>20.70</td>
<td>68.80</td>
<td>4.60</td>
</tr>
<tr>
<td>Median Income ($'000)</td>
<td>55.68</td>
<td>51.30</td>
<td>16.44</td>
<td>282.92</td>
<td>19.92</td>
</tr>
<tr>
<td>Education*</td>
<td>0.805</td>
<td>0.812</td>
<td>0.240</td>
<td>0.985</td>
<td>0.091</td>
</tr>
</tbody>
</table>

* Fraction of the population, 25 years and older, that has at least a high school diploma or its equivalent.

3.5.1 Control variables

Bank and size dummies

The bank dummy, which is included to allow for differences between banks and credit unions, is found to be positive and significant at the 1 percent level. On average the overdraft fee charged by banks is $3.50 greater than that charged by credit unions. This may reflect some inherent differences between banks and credit unions such as ownership structure and tax obligations.

Firms are classified into three size categories based on their asset holdings within a given year. These classes are loosely termed as “small,” “medium,” and “large.” The estimated coefficient on the dummy for small firms is negative and significant at the 1 percent level, whereas the dummy for large firms is positive but significant only at the 10 percent level. These signs are consistent with intuition - large firms have greater market power and are thus able to extract greater rent. That is, consumers on average derive greater utility from the differentiated base product of larger firms.

The data does not allow me to determine whether or not the firm’s size is the true driving force. Economies of scale or scope could allow larger firms to provide a broader ATM network, more branches, more customer service representatives, or other desirable banking features. Alternatively, there may be some underlying characteristic that is correlated with size that gives larger firms a competitive

15 All dollar values are in 2009 constant dollars.
advantage. The effect of size also shows up in the interaction term between size and HHI. This is discussed below.

**Free checking account dummy**

In keeping with our expectations, firms that offer free checking accounts, on average charge $0.72$ more than those that offer checking accounts at a strictly positive price. This is consistent with the predictions of Gabaix and Laibson (2006); sophisticated consumers are willing to pay a higher price for the add-on if they pay a lower price for the base cost. Said differently, the bank is willing to sacrifice profits on the base good only if it can profitably charge a higher price on the add-on goods. It is worth noting at this point that there are additional fees that a bank may charge, such as minimum balance fees and ATM fees, in an attempt to recoup losses that it may sustain from underpricing checking accounts.

**The availability of overdraft protection options**

The cost to consumers for each overdraft is significantly reduced if the consumer purchases an overdraft protection option before the incidence of the overdraft. The two main overdraft protection options are linked accounts and line of credit. When a transaction account is linked to another of the consumer’s accounts, such as a savings account or a credit card account, the bank agrees to transfer funds or charge the credit card to cover the amount of the overdraft. Under line of credit, the bank agrees to pay a consumer’s overdraft obligations at a predetermined APR. Here the parties are effectively entering into a standard loan contract. Both options usually have an upper limit on the amount that is covered, and both charge a fixed upfront fee.

Dummies are included in the regression to control for availability of linked accounts only, line of credit only, or both. When interpreting the coefficients on these variables we need to be cognizant of the fact that we do not know when consumers are made aware of these protection options, nor do we know the total cost of these options. All three dummies are positive. The dummies on line of

---

16 I am not overly concerned about the possible upwards bias on firm specific characteristics, such as size, that may result from the omitted variables problem.

17 Overdraft protection has varied definitions in the banking industry. I define it as any option that is available to the consumer that guarantees the processing of overdraft transactions at a predetermined cost that is lower than the overdraft fees that we have discussed to this point. Linked accounts and line of credit fit neatly into this definition.
<table>
<thead>
<tr>
<th>Regressor</th>
<th>Coefficient</th>
<th>Robust Standard error</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>14.512</td>
<td>2.728</td>
<td>0.000</td>
</tr>
</tbody>
</table>

**Demographic variables**

- Median Age: -0.056, 0.014, 0.000
- Median Income: 0.962, 0.245, 0.000
- Education*: -0.614, 0.921, 0.505

**Firm-specific controls**

- Bank dummy: 3.497, 0.214, 0.000
- Small firm dummy: -0.678, 0.198, 0.001
- Large firm dummy: 0.560, 0.336, 0.096
- Free checking dummy: 0.725, 0.126, 0.000

**Alternatives to overdraft services**

- Linked account only dummy: 0.476, 0.321, 0.138
- Line of credit only dummy: 1.997, 0.342, 0.000
- Both overdraft protection dummy: 1.653, 0.299, 0.000

**Market-specific controls**

- HHI: 0.960, 0.692, 0.166
- HHI x bank dummy: -1.918, 0.633, 0.002
- HHI x small firm dummy: -1.778, 0.489, 0.000
- HHI x big firm dummy: 2.754, 1.114, 0.013

**Year dummies**

- 2005: 0.173, 0.214, 0.420
- 2006: -0.478, 0.205, 0.020
- 2007: -1.232, 0.215, 0.000
- 2008: -0.799, 0.197, 0.000
- 2009: -0.057, 0.195, 0.771

* Fraction of the population, 25 years and older, that has at least a high school diploma or its equivalent.
credit only and both options are significant at the 1 percent level, whereas the dummy for linked accounts only is insignificant.

The positive coefficients are consistent with Gabaix and Laibson (2006) in the following way. The authors’ main argument is that sophisticated consumers consider the possibility of overdrafting, while myopic consumers do not. As such, only the behavior of sophisticated consumers is affected by the availability of overdraft protection options. For a given fraction of myopic consumers, firms attempt to exploit myopic consumers without losing significant market share of sophisticated consumers. By giving sophisticated consumers a cheaper option of overdrafting, it allows firms to charge a higher overdraft fee without severely sacrificing its sophisticated consumer base.

We find that on average, a bank that offers only line of credit has an overdraft fee that is $2.00 higher than a bank that offers neither line of credit nor linked accounts. Interestingly, we find no statistical difference between banks that offer only linked accounts and those that offer neither options. One interpretation is that line of credit is seen as a better overdraft protection option than linked accounts for the fraction of sophisticated consumers that are in fact concerned about the possibility of overdrafting. This interpretation assumes that banks are essentially using second degree price discrimination. Consumers that use overdraft services are of two types; myopic or sophisticated. Banks offer the same service to different groups of consumers at different prices – sophisticated consumers receive lower prices because they consider overdraft prices when choosing a bank.

This finding appears to provide substantial support for G&L. This is the case since overdraft protection options can be considered as substitutes for overdraft services. Standard theory asserts that overdraft fees should be lower in the presence of cheap, easily accessible substitutes. The existence of myopic consumers appears to be a very plausible reason for us observing the opposite effects.

Additionally, if in fact line of credit is a better overdraft protection option than linked account, we can assume that sophisticated consumers who consider the

---

18 Note here that some sophisticated consumers may in fact not consider the cost of overdrafting because they correctly anticipate that they will never overdraft. This may result from the consumer having sufficiently high income and in turn, high account balances or him/her having especially good financial management skills. Based on a nonrandom sample, the FDIC Survey of Overdraft reports that 75 percent of consumers never overdraft.
possibility of overdrafting are those that may have limited alternative resources. This would support the finding of Fusaro (2007) that some amount of overdrafting results from consumers seeking short term loans. I am, however, unable to quantify what proportion of overdrafts can be classified as such.

**State and firm dummies**

I include state (and firm) dummies for those states (and firms) that are observed sufficiently many times for the fixed effects to be estimated with a reasonable degree of precision. Dummies are included for states that are observed at least twenty times, while I include dummies for firms that are observed at least twelve times. We expect state dummies to pick up factors that vary by state but are not associated with the three demographic characteristics of interest to us. These factors include but are not limited to differences across states in banking regulations (for example, taxation), costs, and demand shifters. Nineteen of the thirty-one state dummies and one (Farmers State Bank) of the five firm dummies proved significant at the 5 percent level.

**Concentration**

I attempt to control for the level of competition in a given market by including the city-level HHI. Initially, the model was estimated under the assumption that the coefficient on HHI was constant across firms. In this model we obtained a negative and significant coefficient on HHI, which could be interpreted as support for the efficiency argument with respect to concentration. That is, concentrated markets could result from more efficient firms driving out less efficient ones. The surviving firms in concentrated markets would then be able to operate profitably at lower prices than firms in less concentrated market. Note that cost differences are not isolated to the production of overdraft services only, but also includes the base good and all add-ons.

However, under the structure-conduct-performance (SCP) interpretation we expect that firms in more concentrated markets (higher HHI) would possess greater market power, and hence charge higher prices. Additionally, Shepherd (1982) argues that only firms with large market shares are able to exploit market power in concentrated markets. As such, I interact the HHI variable with the bank and

---

19 The deposits of credit unions are not included in the calculation of HHI.
size dummies, individually. These dummies all proved to be significant at the 5 percent level.

Consistent with the efficiency hypothesis, we find that at the 1 percent significance level, banks charge significantly less in more concentrated cities. Credit unions, on the other hand, charge insignificantly more. The interaction term between size and HHI provides support for the SCP. Small firms in concentrated markets have lower fees while large firms in concentrated markets have higher fees. The coefficient is significant at the 1 percent and 5 percent levels of significance for small and large firms, respectively. These results are consistent with the idea that smaller firms compete more aggressively for market share in more concentrated markets. The behavior of large firms follows standard theory – firms in more concentrated markets exploit their increased market power by charging higher prices. The overall coefficient for banks is negative mainly because over 50 percent of the banks in the data set are small.

**Year dummies**

Having controlled for general price changes by deflating overdraft fees using the general consumer price index, systematic changes by banks over time should be picked up by the year dummies. Three of the five year dummies are significant but the coefficients do not suggest any downward trend in overdraft prices. This result, however, is not necessarily at odds with the notion that consumers learn over time, and thus become better able to avoid using overdraft services. This is the case since the theory only speaks to the effect of aggregate learning. If new myopic consumers replace previously myopic consumers who have learned, then no significant change in overdraft price is to be expected. My data set does not include annual demographic characteristics at the city-level, which possibly would allow us to speak more definitively on the likely changes in the fraction of myopic consumers over time.

### 3.5.2 Variables of interest: Proxies for the fraction of myopic consumers

Recall that my main interest lies in determining whether or not the fraction of myopic consumers is a significant determinant of overdraft fees. Furthermore, if
significant, is the direction of its effect consistent with the theory? Since we cannot
directly observe this key variable, I use demographic proxies, namely, age, income,
and education. I discuss these in turn.

**Median age**

Age proves to be negative and significant at the 1 percent level. This is consistent
with the prediction of the model if younger consumers are more likely to be myopic.

At the most basic level, we can measure financial experience by a consumer’s age.

Even if this is a poor approximation at the individual level, it is likely to be a
reasonable approximation at the city level. Consequently, I assume that older
persons are more financially savvy and thus are more likely to carefully consider
the repercussions of their financial choices. Specifically, an older person is more
likely to have had an overdraft or non-sufficient funds transaction (NSF) in the
past. Thus, at worst, he/she is more likely to be aware of the overdraft policies
of the bank than a younger consumer. Since G&L predict that the equilibrium
price of the add-on will be higher in markets with a higher fraction of myopic
consumers, we expect higher prices for overdraft services in cities with a younger
population. The significant negative relationship that we find is thus in accordance
with Gabaix and Laibson’s model.

**Income**

Consumers with high income are less likely to require overdraft services, and con-
sequently, they are also less likely to consider overdraft services when deciding on
a bank. Thus, technically, we expect to have a higher fraction of myopic con-
sumers in high-income cities. Note, however, that greater myopia of this kind does
not necessarily imply greater demand for overdraft services, which is implicitly
assumed by Gabaix and Laibson. Another channel through which income could
affect the price of overdraft services is that high-income consumers, sophisticated
or myopic, are likely to be less price sensitive. Consequently we would expect to
observe higher overdraft fees in high-income cities, where we have more inelastic
demand.

Both channels described above suggest a positive relationship between income

---

20 This would be true if high-income individuals suffer greater loss (loss of reputation, in-
convenience, and loss of utility from forgone consumption) from bounced transactions.
and overdraft fees. We find that income is positive and significant at the 1 percent level. This suggests that at least one of the following is true: (1) High-income individuals are less price sensitive or (2) High-income individuals are more likely to be myopic but are not significantly less likely to use overdraft services. This result does not necessarily provide evidence in favor of Gabaix and Laibson (2006) but it does suggest that further testing of G&L is warranted.

**Education**

Similar to experience, I assume that more educated consumers are more likely to be financially astute. That is, I expect a negative relationship between the education level of a population and the fraction of myopic consumers in that city. Therefore, a negative relationship between the fraction of the population with at least a high school diploma and the overdraft fee charged would be consistent with G&L. Though education has the predicted negative sign, it is insignificant at any reasonable level of significance. I considered other measures of educational attainment such as the fraction of the population with at least an undergraduate degree. These alternative measures also proved insignificant.

### 3.6 Concluding remarks

This paper tests the implications of the model proposed by Gabaix and Laibson (2006), which seeks to explain the existence of information suppression and above-marginal pricing of add-ons in seemingly competitive markets. The authors make a non-trivial assumption that consumers are of two types: myopic or sophisticated. Myopic consumers are those who ignore the possibility of subsequent demand for the add-on when purchasing the base good.

Here, I empirically test the importance of the presence of myopic consumers in the pricing of overdraft services using a repeated cross-section of U.S. banks and credit unions. Having used demographic proxies for the fraction of myopic consumers, I find that the relationships between overdraft fees and two of the three proxies are consistent with the theoretical model. However, the strongest support

---

21 Given this dual interpretation of a positive coefficient it may be more appropriate to treat income simply as a control variable, rather than one of the proxies for the fraction of myopic consumers.
for Gabaix and Laibson (2006) may be the result that banks that offer consumers a cheaper alternative to overdraft services have higher overdraft fees. The overall findings suggest that structural estimation of Gabaix and Laibson’s model is likely to provide useful insights to help us better understand the pricing of add-ons and after-sale services. I did not structurally estimate the model because of lack of suitable data.

References


MARLON L. WILLIAMS
PhD Candidate
Department of Economics
The Pennsylvania State University
University Park, PA 16802
Email: mlw300@psu.edu

SELECTED TEACHING EXPERIENCE (INSTRUCTOR)

Lock Haven University of Pennsylvania, Lock Haven, PA, 17745
2011 Aug. – Present

The Pennsylvania State University, University Park, PA, 16802

The University of the West Indies (UWI), Mona, Kingston, Jamaica

EDUCATION

The Pennsylvania State University, University Park, PA
Ph.D. Economics Candidate (currently in ABD status and focus on Industrial Organization)
2007 August – Present

The University of the West Indies, Mona, Kinston, Jamaica
M.Sc. Economics (Graduated with distinction)
2004 Aug. – 2006 May

B.Sc. Economics major and Accounting minor (Graduated with First Class Honors)
2001 Aug. – 2004 May

RESEARCH

Publications

Revise and Resubmit (The Economic Journal)
• Williams, Marlon L. (2013), “Advertising by Auto Insurers when Consumers Search Sequentially.” (Chapter 1 of Dissertation research)

Working Papers