

Demand Estimation with Differentiated Products: An Application to Price Competition in the U.S. Brewing Industry

CHRISTIAN ROJAS

Dissertation Submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

in

Economics

Everett Peterson, Chair

Nancy Lutz, Co-Chair

Catherine Eckel

Hans Haller

Richard Ashley

September, 2005

Blacksburg, Virginia

Keywords: Differentiated Products, Demand, Beer, Competition

Copyright 2005, Christian Rojas

ABSTRACT

Demand Estimation with Differentiated Products: An Application to Price Competition
in the U.S. Brewing Industry

CHRISTIAN ROJAS

A large part of the empirical work on differentiated products markets has focused on demand estimation and the pricing behavior of firms. These two themes are key inputs in important applications such as the merging of two firms or the introduction of new products. The validity of inferences, therefore, depend on accurate demand estimates and sound assumptions about the pricing behavior of firms. This dissertation makes a contribution to this literature in two ways. First, it adds to previous techniques of estimating demand for differentiated products. Second, it extends previous analyses of pricing behavior to models of price leadership that, while important, have received limited attention. The investigation focuses on the U.S. brewing industry, where price leadership appears to be an important type of firm behavior.

The analysis is conducted in two stages. In the first stage, the recent Distance Metric (DM) method devised by Pinkse, Slade and Brett is used to estimate the demand for 64 brands of beer in 58 major metropolitan areas of the United States. This study adds to previous applications of the DM method (Pinkse and Slade; Slade 2004) by employing a

demand specification that is more flexible and also by estimating advertising substitution coefficients for numerous beer brands.

In the second stage, different pricing models are compared and ranked by exploiting the exogenous change in the federal excise tax of 1991. Demand estimates of the first stage are used to compute the implied marginal costs for the different models of pricing behavior prior to the tax increase. Then, the tax increase is added to these pre-tax increase marginal costs, and equilibrium prices for all brands are simulated for each model of pricing behavior. These “predicted” prices are then compared to actual prices for model assessment.

Results indicate that Bertrand-Nash predicts the pricing behavior of firms more closely than other models, although Stackelberg leadership yields results that are not substantially different from the Bertrand-Nash model. Nevertheless, Bertrand-Nash tends to under-predict prices of more price-elastic brands and to over-predict prices of less price-elastic brands. An implication of this result is that Anheuser-Busch could exert more market power by increasing the price of its highly inelastic brands, especially Budweiser. Overall, actual price movements as a result of the tax increase tend to be more similar across brands than predicted by any of the models considered. While this pattern is not inconsistent with leadership behavior, leadership models considered in this dissertation do not conform with this pattern.

Acknowledgements

I would like to thank my main dissertation advisor, Everett Peterson, for his guidance, helpful comments and support, and for his collaboration on research stemming from this work. His insights have enriched the content of this dissertation.

I thank Nancy Lutz for helpful comments and guidance. I thank Richard Ashley and Hans Haller for comments and helpful discussions.

I would like to express my special gratitude to Catherine Eckel. I thank Catherine for all the many ways in which she helped me accomplish this dissertation and, most importantly, for always directing me towards the right path. I am deeply indebted to Catherine for putting everything aside from her busy schedule to offer me her continuous encouragement, advice and support in all ways possible. Financial support as her research assistant and Lab Manager helped me achieve much more productive research. I certainly look forward to long and productive collaborative work with Catherine in the future.

I want to offer my deepest gratitude to my beloved parents Edmundo Rojas and Elena Acosta for their love and their constant sacrifice to give me the best opportunities in life. I know my father is proudly watching from above.

Finally, I want to thank my beloved wife, Alexandra Andino for her patience, love, support and sacrifice during all these years. I thank her for giving up everything to be with me far away from home.

Contents

ABSTRACT	ii
Acknowledgements	iv
List of Tables	viii
List of Figures	x
Chapter 1. Introduction	1
1.1. The U.S. Brewing Industry	5
Part 1. Demand Estimation	14
Chapter 2. The Empirical Model	15
2.1. An Overview of Empirical Models of Demand for Differentiated Products	15
2.2. The Model	20
Chapter 3. Data and Estimation	29
3.1. Data	29
3.2. Distance Measures	34
3.3. Own-Price and Own-Advertising Interactions	41
3.4. Estimation Details	41
Chapter 4. Results	45
4.1. Preliminary OLS Regressions	45

4.2. Brand Share Equation	48
4.3. Elasticities	52
4.4. Semi-Parametric Results	57
4.5. Summary of Demand Results	60
Part 2. Price Competition	62
Chapter 5. Background	63
5.1. Introduction	63
5.2. An Overview of Strategies for Assessing Pricing Behavior	65
5.3. The Federal Excise Tax Increase	66
Chapter 6. The Supply Model	68
6.1. Equilibrium Conditions	68
6.2. Implied Marginal Costs	71
6.3. Predicted Prices with Higher Excise Taxes	77
6.4. Actual Price Increases	78
Chapter 7. Results of Model Comparisons	80
7.1. Implied Price-Cost Margins	80
7.2. Predicted vs. Actual Price Increases	82
Chapter 8. Conclusion	96
References	99
Appendix A. Derivation of Elasticity Formulas	103
A.1. Price Elasticities	103
A.2. Advertising Elasticities	104

Appendix B. Data and Programming Details	106
B.1. Demand	106
B.2. Implied Marginal Costs	118
B.3. Simulation	124
Appendix. Vita	146

List of Tables

1.1	Percent of Beer Drinkers, Various Demographic Groups, 2001	7
1.2	Two and Four-Firm National Concentration Ratios in the Beer Industry, Selected Years, 1950-2002	8
1.3	Classification of Beer Brands by Product Type and Price Segment	11
3.1	Description of Variables, IRI Infoscan Database	30
3.2	Selected Brands and their Brewers [acronym, country of origin] (Brand ID)	31
3.3	Summary Statistics	34
4.1	Estimated Coefficient on Weighted Prices and Weighted Advertising from separate OLS Regressions	46
4.2	Results of GMM Estimation of Demand Model	50
4.3	Sample of Median Own- and Cross-Price Elasticities	55
4.4	Sample of Median Own- and Cross-Advertising Elasticities	56
7.1	Summary Statistics of Price Cost Margins for Different Scenarios, (1988-1990)*	81
7.2	Summary Statistics of Actual Price Increases for Selected Brands	90

7.3 Summary Statistics of Actual and Predicted Price Increases and
Performance Metrics of Models 95

List of Figures

1.1	U.S. Consumption of Beer, 1940-2002 (millions of barrels)	6
1.2	U.S. per capita consumption of Beer, 1940-2002 (gallons)	6
1.3	Market Shares of Largest Brewers	9
3.1	Fraction of Cities in which a Brand appears (mean of 20 quarters)	37
3.2	Location of Common Boundaries of Brands in Coverage-Size Space (Chicago, 4th Quarter of 1992)	40
4.1	Function g for Nearest Neighbors (<i>NNAC</i>)	58
4.2	Function g for Same-Brewer Beers (<i>BREW</i>)	58
4.3	Function g for Same-Product Segment Beers (<i>PROD2</i>)	59
4.4	Function g for Beers with Similar Geographic Presence (<i>REG</i>)	59
5.1	Quarterly Mean Prices, Various Segments (1988-1992)	67
7.1	Predicted Price Increases by Bertrand-Nash behavior vs. Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)	84
7.2	Predicted Price Increases by Leadership of Anheuser-Busch and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities).	85

7.3	Predicted Price Increases by Leadership of Budweiser and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)	86
7.4	Predicted Price Increases of Collusive Price Leadership by Budweiser and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)	87
7.5	Predicted Price Increases by Collusive behavior between 3 largest firms and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)	88
7.6	Predicted Price Increases by Collusive behavior between leading brands of 3 largest firms and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)	89
7.7	Mean of Actual Price Increases vs. Mean of Predicted Price Increases, Bertrand-Nash Competition	91

CHAPTER 1

Introduction

Differentiated products are prominent in many consumer goods markets. A large part of the empirical work on these markets has focused on demand estimation and the pricing behavior of firms as these are two key inputs in important applications such as merger simulation and the analysis of new product introductions. These studies construct a partial equilibrium model by combining demand estimates with an assumption about firms' pricing behavior. The usual practice is to assume an environment in which firms price their products competitively and then simulate counterfactual equilibrium conditions that resemble the cases of interest.¹ Clearly, the validity of using this kind of inference depends on accurate demand estimates and on how close firms' behavior is to the assumed model of how firms price. This dissertation makes a contribution to the literature in two areas. First, it adds to previous techniques of estimating demand for differentiated products by considering a more flexible functional form that also includes the effects of advertising. Second, it extends the analysis of pricing behavior to models of price leadership that, while important, have received limited attention.

This investigation follows the New Empirical Industrial Organization (NEIO) literature by estimating structural demand parameters and making use of game theory to capture the strategic interactions among firms (Bresnahan, 1989). One of the advantages of this approach is that structural parameters of demand are readily interpretable and

¹In merger simulation the case of interest would be the equilibrium prices and quantities after two firms merge. For new product introductions, the case of interest would be the welfare change due to the introduction of a new brand.

tightly linked to theory. With this approach, the implied marginal costs can be recovered for different pricing games without actually observing them. The general strategy consists of two stages.

In the first stage brand-level demand parameters are estimated (Part I). Demand estimation with differentiated products is a non-trivial endeavor given the numerous price-substitutability parameters that need to be estimated. Chapter 2 discusses some methods that have allowed researchers to deal with this dimensionality problem in various ways. This dissertation employs a recent technique to estimate the demand for 64 brands of beer in 58 major metropolitan areas of the United States. Unlike most previous work on demand for differentiated products, the demand model is based on the neoclassical “representative consumer” approach rather than on a “discrete choice” approach. While for some products such as automobiles the discrete choice assumption seems logical, for others it appears less likely. This study provides two contributions to the literature of demand estimation with differentiated products. First, it extends previous applications of the Distance Metric (DM) method (Pinkse and Slade; Slade 2004) by replacing the linear demand specification with a demand system that is more flexible. Second, this is the first attempt at estimating advertising substitution coefficients for numerous beer brands. The inclusion of advertising in the demand specification is a unique feature of this study given the importance of advertising in many differentiated products industries.

Historical accounts of the existence of a leading brand or firm whose pricing decisions are closely followed by competitors have been reported in well scrutinized industries: the cigarette industry in the 1920’s and 1930’s, the U.S. automobile industry in the 1950’s, the breakfast cereals industry between the 1960’s and 1970’s, and the U.S. brewing industry

since the 1970's.² In the U.S., Anheuser-Busch is the leading beer producer with a 50% market share. Anheuser-Busch has been identified as a price leader especially through its heavily marketed "King of Beers" brand Budweiser (Greer; Tremblay and Tremblay; and references therein).

The importance of price leadership goes beyond these frequent informal encounters. Price leadership not only allows firms to earn larger profits than competitive pricing but it may also be used as a collusive device producing large welfare losses. Rotemberg and Saloner, for example, show that with their dynamic "collusive price leadership" model, society could incur in higher welfare losses than with overt collusion. Despite the economic importance of price leadership, formal empirical assessments of such behavior in differentiated goods markets are largely absent.³

The second stage of this investigation compares and ranks different models of pricing behavior in the U.S. brewing industry (Part II). While the task of assessing different models of firm behavior often involves constructing complex non-nested econometric tests of the competing hypotheses, here an exogenous variation in the data or "natural experiment" is exploited to evaluate various models of pricing behavior. The natural experiment in this dissertation is the 1991 100% increase of the federal excise tax on beer. This change in policy caused all beer producers and importers to simultaneously adjust their prices to the new per unit tax. This behavior is captured in the unique data set which spans before and after the tax increase (1988-1992).

To perform the assessment of pricing behavior, the structural demand system estimated in Part I is used to compute the implied marginal costs for the different models

²See Scherer and Ross for accounts in cigarettes, automobiles and breakfast cereals. See Tremblay and Tremblay for accounts in brewing.

³Two exceptions are Kadiyali, Vilcassim and Chintagunta, and Gasmi, Laffont and Vuong. These studies, however, have been limited to a small number of differentiated products (4 and 2, respectively).

that prevailed before to the tax increase. The exogenous increase in the federal excise tax from \$9 to \$18 per barrel levied on all brewers is added to the implied marginal costs in the pre-tax increase period and the “predicted” equilibrium prices for all brands are simulated for each model. Several model fit measures are used to assess the closeness of each model’s “predicted” prices to the prices that were actually observed in the period that followed the tax increase.

Two types of leadership models are considered. The first is a “collusive price leadership” model in which followers match Budweiser’s price changes. The second is a Stackelberg model. In one variant of the Stackelberg model Anheuser-Busch uses its brand Budweiser as the price leader. In the other variant of the Stackelberg model, all brands produced by Anheuser-Busch act as price leaders. These leadership models are compared to alternative hypotheses of Bertrand-Nash and collusion. Although price leadership appears to be an important type of pricing behavior in some differentiated products industries, this is the first study to econometrically analyze this model at this level of product disaggregation.

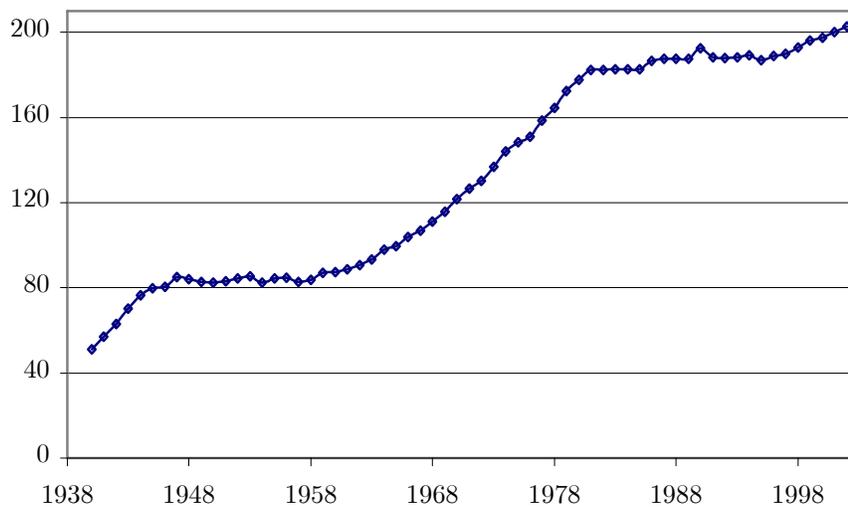
Bertrand-Nash predicts the pricing behavior of firms more closely than other models, although Stackelberg leadership models produce results that are not substantially different from the Bertrand-Nash model. However, Bertrand-Nash tends to under-predict prices of more price-elastic brands and to over-predict prices of less price-elastic brands. This results suggests that Anheuser-Busch could exert more market power by increasing the price of its highly inelastic brands, especially Budweiser. Overall, actual price movements as a result of the tax increase tend to be more similar across brands than predicted by any of the models considered. While this pattern is not inconsistent with leadership behavior, the specific leadership models considered in this paper do not conform with this pattern.

1.1. The U.S. Brewing Industry

Commercial brewing began during the American colonial period. By 1810 there were 132 breweries producing 185,000 barrels of mainly English-type beers (i.e. ale, porter and stout). Lager beer was later introduced in the mid nineteenth century. A common technical classification for beers divides them into lagers and ales. Lagers are brewed with yeasts that ferment at the bottom of the fermenting tank while ales are brewed with yeasts fermenting at the top of the fermenting tank and at higher temperatures. Porter and stout are also regarded as beer types but are generally darker and sweeter than ale. Today, over 90 percent of the U.S. brewing industry's output consists of lagers while the remaining 10% is distributed among ales, porters and stouts. Draught beer, which does not go through a pasteurization process and has to be maintained at low temperatures, accounts for less than 10% of beer production (down from a 50% share in 1940).

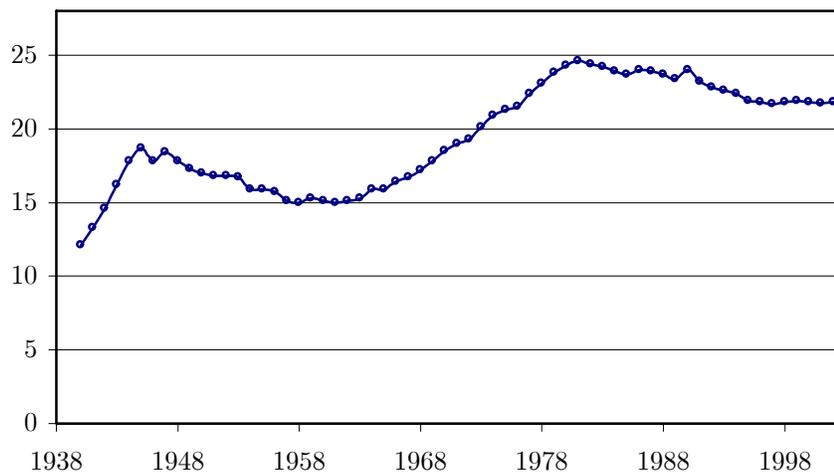
Overall, total demand for beer in the U.S. has increased from 51 million barrels in 1940 to 203 million in 2002 (figure 1.1). However, aside from a slight recovery in recent years, for the last three decades total demand for beer has remained relatively constant (between 180 and 210 million barrels per year) compared to the much stronger growth patterns between 1960 and 1980. On a per capita basis, consumption has fluctuated between 15 and 25 gallons (figure 1.2). Per-capita consumption declined until the early 1960's after which, it sharply escalated to an all time high of 24.6 gallons in 1981. This increase was fueled by increased numbers of young adults (i.e. the baby boomers), lower drinking age requirements and broader acceptance of beer among women (Elzinga 2000: 87). Opposite forces in the last decade, namely an aging population and stricter drinking age requirements, have brought per capita consumption down to approximately 22 gallons.

Figure 1.1. U.S. Consumption of Beer, 1940-2002 (millions of barrels)



Source: Brewers Almanac

Figure 1.2. U.S. per capita consumption of Beer, 1940-2002 (gallons)



Source: Brewers Almanac

Per capita consumption of beer ranks fourth among commercial beverages, behind soft drinks, coffee and milk. Of beer, spirits and wine, beer accounts for approximately 90% of

all volume sales.⁴ In terms of total volume in 1998, beer was the second largest beverage category with 19.4% of sales volume, after carbonated soft drinks which accounted for 49% (Beverage World 1999, cited in Dhar, Chavas and Gould).

Consumption of beer varies among different groups of consumers. More than half of the adult population does not drink beer, while 80% of all beer is consumed by 13% of the population. Younger adults and males account for the majority of beer consumption. Table 1.1 displays the fraction of beer drinkers in different demographic categories.

Table 1.1. Percent of Beer Drinkers, Various Demographic Groups, 2001

<i>Gender</i>	
Males	57.5*
Females	36.5*
<i>Age Brackets</i>	
18-24	50.5
25-34	51.6
34-44	49.6
45-54	44.2
55-64	35.0
65+	26.0
<i>Race</i>	
White	47.3*
African-American	42.1*
Other	43.8*

Source: Tremblay and Tremblay (2005: 19).

* Data for 1995

Consumption patterns also vary geographically. Nine states (CA, TX, FL, NY, PA, IL, OH, MI, NC) account for over 50% of total consumption of beer in the U.S. (Brewers Almanac). The lowest annual per-capita consumption is registered in Utah with 12.9 gallons while Nevada has the highest at 32.1 gallons (Tremblay and Tremblay: 20).

Concentration

⁴In 1997, per capita consumption of beer was 22 gallons whereas per capita consumption of wine and distilled spirits was 1.23 and 1.9, respectively (Brewers Almanac)

Many features of the U.S. brewing industry have interested researchers for decades. Arguably, the dramatic change of the industry from a fragmented structure to a concentrated oligopoly has drawn the most attention. Before World War II, most beer was supplied by smaller firms that would seldom sell beyond their local (a city or a metropolitan area) or regional markets (one or a few contiguous states). Today most production is concentrated among national or “mass-producing” firms.

The number of mass-producing brewers declined from 350 in 1950 to 24 in 2000 with a corresponding increase in the Herfindahl index from 204 to 3612 (Tremblay and Tremblay: 187), making this industry one of the most concentrated in U.S. manufacturing. To illustrate, the Herfindahl indices for cigarettes, breakfast cereals and automobiles are 3100, 2446 and 2506 respectively. For all manufacturing industries, the index is 91 (U.S. Census Bureau, 1997 concentration ratios).⁵ Table 1.2 shows how the sharp increase in the two- (C2) and four-firm (C4) concentration ratios which have escalated from 11.9 and 22 in 1947 to 67.9 and 83 in 2002, respectively. Figure 1.3 shows how the market shares of the four largest producers have evolved.

Table 1.2. Two and Four-Firm National Concentration Ratios in the Beer Industry, Selected Years, 1950-2002

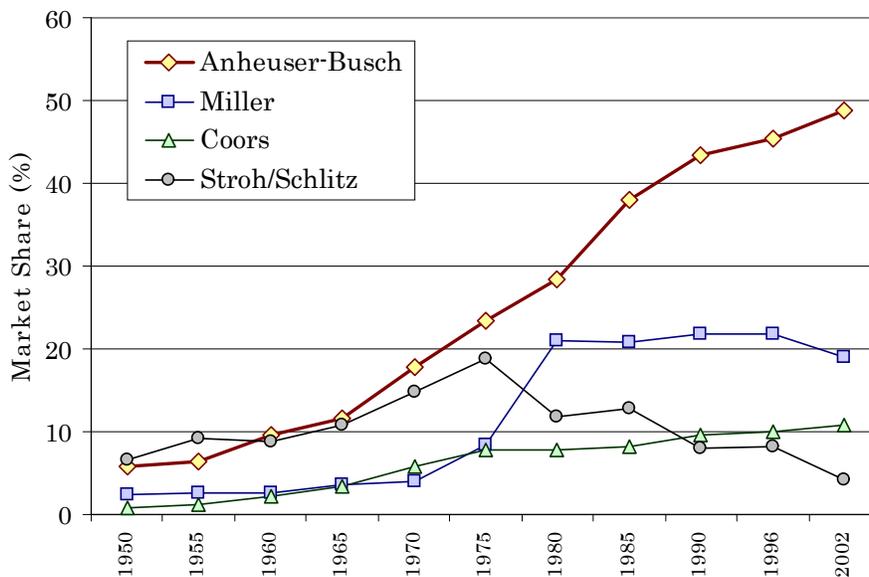
	1950	1955	1960	1965	1970	1975	1980	1985	1990	1996	2002
A-B	5.8	6.5	9.6	11.6	17.8	23.4	28.4	38.1	43.4	45.4	48.9
Miller	2.5	2.6	2.7	3.6	4.1	8.5	21.1	20.8	21.8	21.8	19.0
Coors	0.8	1.2	2.2	3.5	5.8	7.9	7.8	8.3	9.7	10.0	10.9
Stroh/Schlitz	6.7	9.2	8.8	10.8	14.8	18.9	11.9	12.9	8.1	8.3	4.2 ⁶
C2	11.9	13.3	16.0	20.1	29.9	38.8	49.5	58.9	65.2	67.2	67.9
C4	22.0	22.1	27.0	34.4	44.2	57.7	66.4	80.8	83.0	85.5	83.0

Source: Greer, Beer Marketer’s Insights

⁵These numbers correspond to domestic manufacturers and hence do not reflect sales of imported goods.

⁶Stroh exited the market in 1999. This number corresponds to Pabst’s market share. Pabst acquired some of Stroh’s brands (Elzinga 2000).

Figure 1.3. Market Shares of Largest Brewers



Source: Greer (1998), Beer Marketer's Insights.

Note: Stroh exited the market in 1999. The 2003 number corresponds to Pabst's market share (the acquirer of some of Stroh's brands).

Anheuser-Busch became the largest beer producer in 1960 and has continued to increase its market share (Table 1.2 and Figure 1.3). In 2003 Anheuser-Busch (49.8%), SABMiller (17.8%) (formerly Miller and owned by Philip Morris) and Coors (10.7%) accounted for nearly 80% of all beer sales in the U.S.. Budweiser and Bud Light, Anheuser-Busch's two leading brands, capture approximately one third of beer sales nationwide.⁷ Besides the marked growth in market shares by these three firms, a remarkable aspect is the rise of Miller's market share in the 1970's from a mere 4.1% in 1970 to 21.1% in 1980. This leap has been attributed to the success of the Miller Lite brand.

While imports and specialty beers have increased their combined market share from less than 1% in the 1970's to approximately 12% and 3%, respectively, their impact in the industry as a whole remains limited. The reason is that imports and specialty beers

⁷Based on 2001 production estimates (Tremblay and Tremblay: 13)

tend to compete less directly with traditional mass-producers since they target different types of consumers.

Larger production, and hence concentration, has been stimulated by considerable economies of scale (Elzinga, 2000). Elzinga suggests that significant unit costs reductions are achieved by plants that operate with a capacity of around 1.25 million barrels per year and that even further, though less pronounced, gains can be realized up to 8 million barrels.⁸ These benefits are mainly attributable to faster packaging equipment (which can only be used by large scale plants), automation and labor specialization. The national presence of these large producers may have also contributed to concentration. Greer argues that national brewers have access to less costly (per viewer) advertising because it is more efficient to promote a product nationally than it is locally. This advantage can thereby further reduce unit costs for large producers.

Mergers have also played a role in the consolidation of the U.S. brewing industry. From 1950 until 1983, there were 170 horizontal mergers. However, as shown by Elzinga (2000), these numerous mergers have not led to increased market shares for the acquiring firms. The most active firms (in terms of mergers), Stroh and Hieleman, are no longer in the market while the lesser active, A-B, Miller and Coors, remain as the market leaders mainly by way of internal expansion.⁹

Beer Segments

The different brands of beer are usually grouped into several categories or segments. Although there is no complete agreement as to the most relevant segmentation of brands, most agree on dividing beers into various price segments. For example, Greer groups

⁸The 1.25 million gallons capacity can be roughly considered as the minimum efficient scale (MES).

⁹A more detailed analysis of mergers in U.S. brewing is beyond the scope of this dissertation and can be found in Elzinga (1973, 2000).

Table 1.3. Classification of Beer Brands by Product Type and Price Segment

	Imports	Super-premium	Premium	Popular
Lager/ Draft	Heineken Molson Fosters Corona	Blue Moon Killians Rolling Rock Michelob	Budweiser MGD Coors	Miller High Life Busch Old Milwaukee Keystone
Light	Amstel Light Molson Light Becks Light Corona Light	Michelob Light	Miller Lite Bud Light MGD Light Coors Light	Busch Light Keystone Light Strohs Light
Dry	Kirin Dry Molson Dry	Michelob Dry	Bud Dry Coors Dry	Keystone Dry Olympia Dry
Ice	Molson Ice Labatt Ice		Bud Ice Ice House Coors Artic Ice	Old Milwaukee Ice Pabst Ice Busch Ice

Source: Greer, and other industry sources

beers into (from more to less expensive): imports, super-premium, premium and popular. In addition, Greer allocates various brands to the different price segments according to whether they are lager, light, dry or ice (table 1.3). Other studies (e.g. Hausman, Leonard and Zona; and Langan) include light beer as another price segment (i.e. in the horizontal classification of table 1.3), and include imports as part of the superpremium segment. Conversely, an industry report by SABMiller adopts three segments: worthmore (imports, superpremium and specialty brands), mainstream (premium regular and premium light) and economy (popular).

In spite of price differences within a segment, studies of blind beer testing suggest that consumers can not differentiate between brands and that quality differences are, in general, not related to price.¹⁰ That is, more expensive beer, for a given type of beer (e.g. ice, dry), appears not to be an indication of higher quality.

¹⁰An experiment of blind beer testing (Consumer Reports June, 1996:10-17, cited in Greer) showed that some less expensive beers were ranked among some of the best in taste while some of the higher priced brands ranked near the bottom. In addition, it was found that no significant correlation existed between prices and taste rankings.

Part of this inconsistency is attributed to the origins of premium and light beers. The earlier expansion of regional brewers to a national market was carried out through the promotion of their brands as being superior in quality and adding a price “premium”. However, this premium was used to offset the high transportation costs from their plants (usually only one or two) to distant locations and not for more expensive production techniques. Once expansion took the form of multiplant operations, transportation costs reduced significantly, but the premium image and, to a lesser extent, the higher price remained. Similarly, when light beer was successfully introduced by Miller in the 1970’s it was marketed (and still is) at a higher price and as of finer quality than regular beer, even though its production costs were lower.¹¹ The premium and light beers stories portray the consumer’s problem of incomplete information about the quality of the product. As a consequence, consumers will rely on cues about the product’s quality. Greer suggests that for beer these are: price, seller’s market position (or reputation) and advertising.

Price-quality relationship aside, there have been important demand patterns at work among the price segments in the classification above. Light beer (non-existent 30 years ago) has become the most popular product with a market share of over 40%. Conversely, popular (economy) beer, once the leading segment, has lost ground to premium beer and imports. In 2003, people consumed 54.3% of premium beer (regular and light), 20.1% of imports and super-premium (including craft brands) and 25.6% of popular beer (Beer Marketer’s Insights). Overall, the U.S. brewing industry is characterized by numerous product introductions and, consequently, a large number of brands.

While the data available for this investigation correspond to an earlier period (1988-1992), generally speaking the described patterns of concentration and beer segments were already present 15 years ago. Perhaps the main difference between the industry today

¹¹Light beer is made with more water and less ingredients.

and the industry in the late 1980's and early 1990's is a larger market share of imported and specialty beers today.

Conduct and Other Features of the Industry

In terms of firm conduct, Anheuser-Busch is often assumed to be the price leader (Greer; Tremblay and Tremblay). In particular, the price of Anheuser-Busch's Budweiser is considered to be driving the prices of other brands. Also, since price reductions of the premium and super-premium brands might damage their image, price competition in the popular segment appears to be more frequent (Greer).

Anheuser-Busch has consolidated its position as the leader in this industry not only through constant expansion but also through active advertising strategies (Greer; Tremblay and Tremblay). In 2000, Anheuser-Busch, Miller and Coors spent \$744 million in advertising (Tremblay and Tremblay: 171). Currently, the advertising-to-sales ratio for beer is 8.7 percent compared to 2.9 percent for cigarettes, and 7.1 percent for other beverages.¹² National brewers have taken advantage of the more cost-effective marketing channel: national TV.

In addition to the features briefly presented here, the potential negative externalities of alcohol consumption expand the scope of research about this industry to additional policy and social arenas. A recent extensive economic overview of research on the many features and ramifications of this industry is presented in Tremblay and Tremblay and is beyond the scope of this dissertation.

¹²Source: Advertising Age, 2000, cited in Tremblay and Tremblay.

Part 1

Demand Estimation

CHAPTER 2

The Empirical Model

2.1. An Overview of Empirical Models of Demand for Differentiated Products

Product differentiation is present in nearly all markets. While numerous theoretical approaches address various issues of product differentiation, more recent developments attempt to deal with the non-trivial task of empirically modeling brand-level demand. The data available for demand models can be of two types: micro (or individual level) data and aggregate data. This section presents a brief historical overview of demand models with aggregate data¹, with an emphasis on contrasting the advantages and disadvantages of the two types of models most commonly used by applied empirical industrial organization researchers: neoclassical models and discrete choice (DC) choice models.

Neoclassical demand models establish that under a given set of conditions on preferences, individual demand schedules can be consistently aggregated into a unique ‘representative consumer’ demand function. DC models, on the other hand, are built upon the assumption that consumers purchase one unit of the good that provides them with the highest utility. Since consumers’ choices are not observed with aggregate data, a DC approach models the market shares of the brands as the aggregation of the discrete choices of consumers. As opposed to DC models, neoclassical models do not restrict consumers’ choices to either one brand or one unit of a product. However, as it is explained below, DC models have some advantages over neoclassical approaches. A comparison of these

¹See Reiss and Wolak for details on micro data models.

two approaches is based on two important features of differentiated products: the number of parameters and heterogeneity in consumer tastes.

In many differentiated products markets, numerous varieties of a product are available to consumers. With traditional demand estimation (i.e. estimating one equation for each variety or brand) the task of computing all substitution parameters becomes a challenging one. For example, there are over 200 brands of beer in the U.S. Without imposing the theoretical restrictions of homogeneity and symmetry the number of own- and cross-price coefficients to be estimated is $200^2 = 40,000$. Available databases would rarely be large enough to estimate such a system. Moreover, non-linearities in the demand equations may make estimation computationally unfeasible if the algorithm used does not converge to a solution.² Even if computational feasibility is resolved for a particular problem, applied work often relies on counterfactual estimation during which the obstacle may arise again (Reiss and Wolak). Second, an important fact about heterogeneity in consumption arises in the study of differentiated products: preferences over different product characteristics vary across consumers. This may be one of the main reasons why firms differentiate their products in the first place. Hence, for some researchers, this heterogeneity plays an important role and should be modeled appropriately.

Earlier attempts to estimate differentiated products with neoclassical models are, among others, Baker and Bresnahan; Hausman, Leonard and Zona; and Hausman. Baker and Bresnahan get around the dimensionality problem by calculating the *residual demand* of two (or three) beer producers to later evaluate the potential consequences of mergers (or collusion) in the industry. In practice, however, the researcher is often interested in calculating parameters that are beyond the capabilities of this approach.

²Several neoclassical models specify flexible demand equations that are highly non-linear (e.g. Christensen, Jorgenson and Lau).

Hausman, Leonard and Zona, and Hausman use a *multistage budgeting* approach (Gorman, 1971) in conjunction with an Almost Ideal Demand System (ALIDS) specification. In their model the consumer's decision is divided in three stages. First, consumers decide what portion of their income will be allocated to beer. Then, a decision is made as to how much of each segment (e.g. popular, light, premium and super-premium beer) will be consumed. Finally, how much of each brand to consume within a segment is determined. With multistage budgeting, substitution patterns exist only among brands within the same segment.³ However, this approach has several shortcomings. First, the cross-price elasticities are dependent on the assumed structure of the separable utility function. Given the large number of alternative separability structures, it becomes difficult to test the validity of this assumption.⁴ Another, perhaps more critical, limitation is the fact that as the number of brands increases the dimensionality problem arises again, defeating the original purpose of the exercise.

More recently, in the neoclassical literature, Pinkse, Slade and Brett, have proposed a *Distance Metric* (DM) method that allows a significant reduction of the dimensionality of parameters. The DM method handles the dimensionality problem by specifying the cross-price terms as a function of each brand's location in product space relative to other brands. A brand's location in product space is determined from its observed product characteristics. Various distance measures between brands may be constructed from their relative location in product space and used as weights to create cross-price indices (one for each distance measure). The cross-price coefficients and elasticities can then be computed using the estimated coefficients for the cross-price indices and the distance measures between brands. The advantage of DM method is that it is easier to estimate than the

³Substitution patterns between two segments are also recovered but not at the brand level.

⁴Hausman, Leonard and Zona provide a way to test the validity of segment definitions. However, they note that further research is required in this matter (p. 172).

random coefficients DC model (see below) and allows testing the existence and strength of different product groupings as potential sources of competition, instead of ad-hoc segment definitions as in the multistage budgeting approach and nested logit models. While this technique does not directly specify how product characteristics enter consumers' utility, it does recognize the role that a brand's location in product space plays in differentiated products industries.

Aggregate DC models (Berry; Berry Levinsohn and Pakes -BLP henceforth) have been more commonly used to analyze the demand for numerous differentiated products.⁵ One reason for their popularity is that these models handle the dimensionality problem in a parsimonious way. Specifically, DC models project the number of products on to a lower dimensional space, namely the product characteristics. Thus, the substitution patterns for 200 beer brands may (in principle) be recovered by estimating parameters only on product characteristics and price. Within this class of models, the simplest is the *logit* model, which has been widely used, but nevertheless possesses the non-desirable independence of irrelevant alternatives (IIA) attribute. In terms of demand estimation, the IIA condition implies that the cross-price elasticities in a given column of the elasticity matrix will be equal.

More general models have been constructed to deal with the restrictive IIA condition, namely the *nested logit* (McFadden, 1978) and the *random coefficients*, or *mixed*, logit (Berry, BLP, and McFadden and Train). The nested logit model partially relaxes the IIA condition by grouping, or nesting, the products into mutually exclusive sets and allowing variable substitution patterns across nests. However, the IIA assumption still

⁵These 'aggregate' DC models are slightly different than the individual choice representations (McFadden 1974, 1981, 1984). Here the dependent variable is given by the market share of the product instead of the actual individual choices.

holds within nests, thus constraining the substitution patterns in a given segment. Moreover, the a priori group definitions create an additional problem, similar to that of the multistage budgeting (see above). The most general of all DC models is the random coefficients logit. In this specification, more general substitution patterns are possible by allowing taste parameters to vary by consumer, while maintaining the low dimensionality feature of the logit model. Another important characteristic of the random coefficients model is that it captures heterogeneity in consumer preferences; namely, variations of the taste parameters according to different demographic characteristics of the consumer. Ultimately, the random coefficients on the taste parameters allow the relaxation of the IIA property. However, the more flexible and general random coefficients model requires significant computational costs due to the non-closed form of the market shares integral (a direct consequence of the random coefficients) which demands a time consuming search for the parameters.⁶

In sum, the model by BLP has one major advantage over neoclassical models (including those that use the DM method): it controls and models consumer heterogeneity explicitly. The costs of this feature are its computational complexity and, when unlikely to hold, the restrictive discrete choice assumption. Neoclassical models, on the other hand, relax the discrete choice assumption. Further, the DM technique handles the number of parameters in a parsimonious way while reducing the computational complexity. Since both models have their limitations and advantages, the choice of what model to use will vary depending on the specific application. In this study, the discrete choice assumption of purchasing a single unit of a single brand appears unlikely and hence a neoclassical model is employed.

⁶These and other estimation details of the random-coefficients model are surveyed in Nevo (2000a).

2.2. The Model

While the DM method reduces the dimensionality problem, it does not constrain the choice of functional form. This section is divided in two parts. First, the choice of functional form for demand is discussed and it is shown how this choice is more general than that of previous work that has applied the DM method (i.e. Pinkse and Slade and Slade 2004). The second part presents the DM method and the third part its semi-parametric version.

The Choice of Functional Form

The functional form is based on the representative consumer approach of neoclassical demand models. In this dissertation, Deaton and Muellbauer's Almost Ideal Demand System (ALIDS) is used. Some of the desirable properties of the ALIDS model are: (a) satisfies the principles of utility maximization, (b) accommodates exact (i.e. non-linear) aggregation, (c) it is a flexible form approximation to any expenditure function and (d) allows testing of theoretical properties of demand functions. Specifically, Deaton and Muellbauer (1980) start by specifying an expenditure function based on the logarithmic form of price independent generalized linear (PIGLOG) preferences. What characterizes PIGLOG preferences is that non-linear, instead of linear, aggregation across consumers is possible. The importance of PIGLOG preferences, and hence of property (b) above, rests upon the fact the ALIDS model is more general and thus more flexible than those that rely on linear aggregation.⁷

⁷Linear aggregation is feasible when indirect utility functions are of the Gorman Polar Form (Gorman, 1959). This particular form assumes that all consumers, both high-income as well as low-income, have the same marginal utility of income (this is also known as quasi-homothetic preferences or parallel Engel curves).

The models of Pinkse and Slade, and Slade (2004), on the other hand, rely on a linear aggregation assumption. This assumption is not overly restrictive if every consumer purchases all products being considered in the analysis, for example when modeling broad categories of goods such as food and clothing. However, with differentiated products consumers are likely to purchase at most a few varieties of the good. To circumvent this problem, Pinkse and Slade, and Slade (2004), assume that the income effect for all brands is zero.⁸

Formally, let $j \in (1, \dots, J)$ be the brand index, $t \in (1, \dots, T)$ the market index (in this application, the market is defined as a city-quarter pair), $q_t = (q_{1t}, \dots, q_{Jt})$ the vector of quantities demanded, $p_t = (p_{1t}, \dots, p_{Jt})$ the corresponding price vector and $x_t = \sum_j p_{jt} q_{jt}$ total expenditures. The ALIDS is given by the following formula:

$$(2.1) \quad w_{jt} = a_{jt}^* + \sum_k b_{jk} \log p_{tk} + d_j \log(x_t/P_t)$$

where $w_{jt} = \frac{p_{jt} q_{jt}}{x_t}$ is brand j 's sales share in market t and $\log P_t$ is a price index defined by:

$$(2.2) \quad \log P_t = a_0 + \sum_j a_{jt}^* \log p_{jt} + \frac{1}{2} \sum_j \sum_k b_{jk} \log p_{jt} \log p_{kt}$$

Equation (2.1) is non-linear in the parameters and may be estimated by non-linear methods. However, because of the large number of cross-price terms to be included, estimating a non-linear equation will be unpractical. Following Deaton and Muellbauer

⁸Pinkse and Slade; and Slade (2004) assume that if there exists a single consumer that does not purchase a particular brand, then the income effect for that brand should be zero. In practice it amounts to assuming that the income effect for each brand is zero, because it would be easy to find at least one consumer that does not purchase it.

(1980: 316), an approximation to the price index $\log P_t$ is employed, resulting in the model known as the *linear* ALIDS or LALIDS. Moschini shows that the Stone price index employed by Deaton and Muellbauer has the undesirable property of not being invariant to changes in units of measurement.⁹ Following Moschini, the loglinear analogue of the Laspeyres price index is used instead of the Stone price index to linearize the ALIDS:

$$(2.3) \quad \log P_t \approx \log P_t^L = \sum_j w_j^o \log(p_{jt})$$

where w_j^o is brand j 's 'base' share, that is defined as, $w_j^o = T^{-1} \sum_t w_{jt}$, or the average share of brand j over t .¹⁰

Pinkse and Slade and Slade (2004), in contrast, derive a non-linear demand equation from a quadratic indirect utility function, which is a flexible second order approximation, in prices and income, to any indirect utility function. In order to obtain a linear demand equation, they set the price index in the denominator to 1, losing (almost) no generality since their application is limited to a small panel (i.e. two-cross sections). Therefore, in addition to allowing nonlinear aggregation, the LALIDS specification places no restrictions on the length of the panel used for estimation.¹¹

Incorporating the new price index (2.3), the LALIDS in sales share form is written as

$$(2.4) \quad w_{jt} = a_{jt}^* + \sum_k b_{jk} \log p_{kt} + d_j \log(x_t/P_t^L)$$

⁹Essentially, the invariability principle states that when units of measurement are changed a price index may only vary up to a multiplicative constant. Violation of this property could lead to biased estimates.

¹⁰An advantage of this price index is that it does not contain w_{jt} directly, but rather a 'fixed' base (w_j^o). This moderates the problem of having an additional endogenous variable on the right hand side.

¹¹It may be argued, however, that the demand function of Pinske and Slade and Slade (2004) could be linearized by a price index of the sort chosen here to allow for longer panels. Unfortunately, approximation properties of price indices in such models have not been investigated.

Given that beer commercials provide little information about the product, advertising is assumed to be persuasive rather than informative. This dissertation focuses on traditional advertising (e.g. television, radio and press), rather than on local promotional activity (e.g. local paper, in-store promotions, and end-of-aisle product location), as the key advertising variable because it has played a crucial role in the development of the industry. Further, only the flow effects of advertising are considered with all lagged own- and cross-advertising terms being omitted from the demand equation.¹²

Advertising is incorporated into equation (2.4) through the intercept term, which is modified to equal:

$$(2.5) \quad a_{jt}^* = a_{jt} + \sum_k c_{jk} A_{kt}^\gamma$$

where A_{kt} represents advertising expenditures of brand k in market t .¹³ The parameter γ is included to account for decreasing returns to advertising. Following Gasmi, Laffont and Young, γ is set equal to 0.5. The constant term a_{jt} incorporates time, city and brand binary variables as well as product characteristics and other market specific variables (e.g. demographics). Substituting equation (2.5) into equation (2.4) and including an econometric error term ε_{jt} gives:

$$(2.6) \quad w_{jt} = a_{jt} + \sum_k b_{jk} \log p_{kt} + \sum_k c_{jk} A_{kt}^\gamma + d_j \log(x_t/P_t^L) + \varepsilon_{jt},$$

¹²The existence of possible stock effects was investigated by specifying lagged advertising expenditures. In all specifications, the coefficients for lagged advertising expenditures were found to be statistically insignificant.

¹³A logarithmic specification for advertising could not be used because of zero advertising expenditures for some brands across cities and time.

Equation (2.6) can now be interpreted as a first-order approximation in prices and advertising to the demand function that allows for unrestricted price and advertising parameters. The usual practice with an ALIDS is to estimate all coefficients by specifying $(J - 1)$ seemingly unrelated equations, one for each product. However, with 64 brands this task becomes problematic given the large number of parameters to be estimated. The DM method, as explained next, is utilized to reduce the dimensionality of the estimation.

The Distance Metric (DM) Method

The cross-price and cross-advertising coefficients (b_{jk} and c_{jk}) in equation (2.6) are specified as functions of different distance measures between brands j and k . These distance measures may be either continuous or discrete. For example, the alcohol content of a brand is an example of a variable that can be used to construct a continuous distance measure. Dichotomous variables that identify brands by product segment, such as light beer or regular beer, can be used to construct a discrete distance measure. The continuous distance measures use an inverse of the Euclidean distance, or closeness, in product space between brands j and k . This measure of closeness varies between zero and one, with a value of one if both brands are located at the same location in product space. The discrete distance measures take the value of 1 if j and k belong to the same grouping and zero otherwise.

Following Pinkse, Slade and Brett, the cross-price and cross-advertising coefficients are defined to be equal to $b_{jk} = g(\delta_{jk})$ and $c_{jk} = h(\mu_{jk})$, where δ_{jk} and μ_{jk} are the set of distance measures for price and advertising, respectively.¹⁴ Then, equation (2.6) can be written as:

¹⁴To simplify notation, distance measures (δ_{jk} and μ_{jk}) are depicted as market invariant. In the application, however, some distance measures vary by market.

$$(2.7) \quad w_{jt} = a_{jt} + b_{jj}p_{jt} + c_{jj}A_{jt}^\gamma + \sum_{k \neq j} g(\delta_{jk}) \log p_{kt} + \sum_{k \neq j} h(\mu_{jk}) A_{kt}^\gamma + d_j \log(x_t/P_t^L) + \varepsilon_{jt},$$

The functions g and h measure how the strength of competition between brands varies with distance measures. These functions are specified as a linear combination of the distance measures:

$$(2.8) \quad g = \sum_{l=1}^L \lambda_l \delta_{jk}^l$$

$$(2.9) \quad h = \sum_{m=1}^M \tau_m \mu_{jk}^m$$

where λ and τ are coefficients to be estimated, L and M are the number of distance measures for price and advertising, respectively. Because the distance measures are symmetric by definition ($\delta_{jk} = \delta_{kj}$ and $\mu_{jk} = \mu_{kj}$), symmetry may be imposed by setting λ and τ to be equal across equations. This implies that $b_{jk} = b_{kj}$ and $c_{jk} = c_{kj}$. The cross-price and cross-advertising coefficients (b_{jk} , c_{jk}) and elasticities are then recovered from the estimates of λ and τ , and the distance measures.

In principle, $(J - 1)$ seemingly unrelated equations can be estimated. However, if J is very large, as is the case in this dissertation with 64 brands, then it may become impractical to estimate such a large system of equations. One method to reduce the dimensionality of the estimation procedure is to assume that the own-price and own-advertising coefficients (b_{jj} and c_{jj}), as well as the coefficient on real expenditures (d_j), are constant across equations thereby reducing estimation to a single equation. Since this is too restrictive of an assumption, following Pinkse and Slade the coefficients b_{jj} , c_{jj} , and

d_j are specified as functions of each brand's product characteristics. For example, using alcohol content as the only product characteristic, the own-price coefficient in equation (2.7) would be defined as $b_{jj} = b_1 + b_2 ALC_j$, where ALC is brand j 's alcohol content. Thus $b_{jj} \log p_{jt} = b_1 \log p_{jt} + b_2 \log p_{jt} ALC_j$, effectively interacting price with the product characteristics.

Combining equations (2.7), (2.8), and (2.9) with the own-price and own-advertising interactions described above yields:

$$(2.10) \quad w_{jt} = a_{jt} + b_1 \log p_{jt} + \sum_{g=1}^G b_{g+1} \log p_{jt} PC_{gt}^p + c_1 A_{jt}^\gamma + \sum_{h=1}^H c_{h+1} A_{jt}^\gamma PC_{ht}^A + \\ + \sum_{k \neq j} \left(\sum_{l=1}^L \lambda_l \delta_{jk}^l \right) \log p_{kt} + \sum_{k \neq j} \left(\sum_{m=1}^M \tau_m \mu_{jk}^m \right) A_{kt}^\gamma + d_j \log(x_t/P_t^L) + \varepsilon_{jt},$$

where PC_{gt}^p is the g^{th} characteristic of product j interacted with the own-price, and PC_{ht}^A is the h^{th} characteristic of product j interacted with own-advertising. After regrouping cross-prices into L weighted terms and cross-advertising expenditures into M weighted terms, the empirical model is written as:

$$(2.11) \quad w_{jt} = a_{jt} + b_1 \log p_{jt} + \sum_{g=1}^G b_{g+1} \log p_{jt} PC_{gt}^p + c_1 A_{jt}^\gamma + \sum_{h=1}^H c_{h+1} A_{jt}^\gamma PC_{ht}^A + \\ + \sum_{l=1}^L \left(\lambda_l \sum_{k \neq j} \delta_{jk}^l \log p_{kt} \right) + \sum_{m=1}^M \left(\tau_m \sum_{k \neq j} \mu_{jk}^m A_{kt}^\gamma \right) + d_j \log(x_t/P_t^L) + \varepsilon_{jt},$$

Note that the number of independent cross-price parameters has been reduced from $J(J-1)/2$ to L . Similarly, the number of independent cross-advertising parameters has been reduced from $J(J-1)/2$ to M . In the analysis that follows, each cross-price and cross-advertising distance measure in each market is depicted as a $(J \times J)$ "weighing"

matrix with element (j, k) equal to the distance between brands j and k when $j \neq k$, and zero otherwise. Thus, when the $(J \times J)$ weighing matrix is multiplied by the $(J \times 1)$ vector of brand prices or advertising in each market one obtains the appropriate sum over k in the share equation.

It is important to note that unlike the nested logit or multistage budgeting models, with the DM method substitution patterns are not restricted a priori. Instead, many possible conjectures about the discrete and continuous distance measures can be tested. For instance, discrete measures that take the value of 1 if two brands belong to the same category, and zero otherwise, can be constructed to test different brand groupings and choose the one that is best supported by the data. Chapter 3 further develops these possibilities and also considers some computational issues that arise with the DM method.

Semi-Parametric Specification

The discussion above considered a parametric version of the functions g and h . Because this specification can be a restrictive version of the functions and not always consistent, Pinkse, Slade and Brett propose a semi-parametric variant based on series expansions of the continuous distance measures. For illustrative purposes, suppose that the only relevant product characteristic is alcohol content. Denote with δ_{jk}^{alc} the absolute difference, or distance, in alcohol content between brands j and k . As is common in the semi-parametric literature (e.g. Pagan and Ullah) one can specify a series estimator for an unknown function, b_{jk} , as:

$$(2.12) \quad b_{jk} = g(\delta_{jk}^{alc}) = \sum_{y=0}^Y \lambda_y (\delta_{jk}^{alc})^y = \lambda_0 + \lambda_1 \delta_{jk}^{alc} + \lambda_2 (\delta_{jk}^{alc})^2 + \dots + \lambda_Y (\delta_{jk}^{alc})^Y$$

where $g(\delta_{jk}^{alc})$ is the function being approximated, Y is the number of expansion terms and $(\lambda_0, \dots, \lambda_Y)$ are coefficients to be estimated. In this example, a polynomial series in

δ_{jk}^{alc} was chosen. Alternatively, Fourier series terms, which are based on the trigonometric functions sine and cosine, may also be utilized. After replacing equation (2.12) in (2.7), estimation of the λ_y coefficients are carried out in the same fashion as in the parametric case.

Because series expansions can not be performed with discrete distance measures, Pinkse, Slade and Brett suggest constructing one series expansion for each of the different groups of products defined by the discrete measures. For example, if a discrete measure takes a value of 1 if two beers are of the same type and zero otherwise, then one series expansion like (2.12) would be constructed for beers that are of the same type and another one for beers that are of different type. While this semi-parametric specification is considered in this dissertation, it should be noted that the number of parameters that need to be estimated increase as more discrete measures become available. Also, the inclusion of advertising increases this number. Collinearity problems become more severe as the number of expansion series increases. Therefore only results of restricted versions of this specification are presented in chapter 3. Essentially, these restricted semi-parametric versions are used to check that the parametric version does not yield substantially different estimates than a more semi-parametric flexible specification.

CHAPTER 3

Data and Estimation

3.1. Data

The main source of the data is the Information Resources Inc.(IRI) Infoscan Database, obtained from the University of Connecticut. IRI is a Chicago based marketing firm that collects scanner data from a large sample of supermarkets across various areas in the U.S. The definitions of the geographic areas are similar, but generally broader, than the Metropolitan Statistical Areas used by the Bureau of Labor Statistics. The sample of supermarkets is drawn from a universe of stores with annual sales of more than 2 million dollars. This universe accounts for 82% of all grocery sales in the U.S.. In addition, IRI data correlates well with private sources in the brewing industry. For instance, the correlation coefficient of market shares between data from IRI and data from the Modern Brewery Age Blue Book is 0.95.¹

Several hundred brands are available, for up to 63 geographic areas ("cities" henceforth) and 20 quarters (1988-1992). IRI collects data related to sales, markets and demographics. Table 3.1 provides a description of the variables from the IRI database. The IRI data base is aggregated by brand, that is, data for a given brand consists of sales of that particular brand in all sizes.

Although national brands tend to capture the largest share of beer sales in a given city, there exist cities and regions where regional brands are popular as well. In order to capture this feature, brands having at least a 3% local market share in **any** given city are

¹This calculation was performed for the top 10 brands over the 1988-1992 period.

Table 3.1. Description of Variables, IRI Infoscan Database

Variable	Description	Units
<i>Sales</i>		
Price	Average (per brand) Price	\$/288oz
Quantity	Volume sold	288 oz.
Units	Number of units sold, regardless of size	N/A
COV	Sum of all commodity value (ACV) sold by stores carrying the product / ACV of all stores in the city	%
<i>Market</i>		
SGROC	Supermarket / Grocery Store sales ratio	%
GCR4	Grocery four firm concentration ratio	%
LBCR4	Local, four-brand concentration ratio	%
<i>Demographics</i>		
POP	Population	(000)
HOUS	Median Household size	# persons
INC	Median Income	\$
U10K	Percentage of Households with income under \$10K/year	%
O50K	Percentage of Households with income over \$50K/year	%
HISP	Percentage of hispanic population	%
AGE	Median Family Age	Years

Source: IRI Infoscan Data Base, University of Connecticut, Food Marketing Policy Center

included in the analysis.² Once the set of brands is determined, a brand for a given city and quarter is included in the sample if it has a local market share greater than 0.025%.³ Brands that appear in less than 10 quarters are also dropped. In addition, if a brand appears only in one city in a given quarter, the observation for that brand in that quarter is not included. This is done because sometimes the price of a brand in other cities is used as an instrument. After the selection procedure 33,392 observations are included in the sample, corresponding to 64 brands and 13 brewers. Table 3.2 shows the chosen brands with the name, acronym and country of origin of the corresponding brewer.⁴ Given that many of these brands are only produced regionally, the average number of brands per city is 37 with a maximum of 48 and a minimum of 24.

²For example, if Heineken has a 3% local market share in Boston, then this brand is kept in all cities.

³By this criteria, if Heineken has a 0.001% market share in Chicago, this observation is dropped.

⁴Table 3.2 also displays an assigned brand ID number which will be used later in the analysis.

Table 3.2. Selected Brands and their Brewers [acronym, country of origin]
(Brand ID)

Brewer	Brand	Brewer	Brand
Anheuser-Busch: [AB, U.S.]	(1) Budweiser	Grupo Modelo [GM, Mexico]: Goya [GO, U.S.]: Heineken [H, Netherlands]: Labatt [LB, Canada]: Molson [M, Canada]:	(34) Corona
	(2) Bud Dry		(35) Goya
	(3) Bud Light		(36) Heineken
	(4) Busch		(37) Labatt
	(5) Busch Light		(38) Labatt Blue
	(6) Michelob		(39) Rolling Rock
	(7) Michelob Dry		(40) Molson
	(8) Michelob Golden Draft		(41) Molson Golden
	(9) Michelob Light		(42) Old Vienna
	(10) Natural Light		(43) Falstaff
	(11) Odouls		(44) Hamms
Adolph Coors [AC, US]:	(12) Coors	Pabst [P, U.S.]:	(45) Hamms Light
	(13) Coors Extra Gold		(46) Olympia
	(14) Coors Light		(47) Pabst Blue Ribbon
	(15) Keystone		(48) Red White & Blue
	(16) Keystone Light		(49) Genuine Draft
Bond Corp. [B, U.S.] ⁵ :	(17) Black Label	Philip Morris/Miller: [PM, U.S.]	(50) Meister Brau
	(18) Blatz		(51) Meister Brau Light
	(19) Heidelberg		(52) MGD Light
	(20) Henry Weinhard Ale		(53) Miller High Life
	(21) Henry Weinhard P. R.		(54) Miller Lite
	(22) Kingsbury		(55) Milwaukee's Best
	(23) Lone Star		(56) Goebel
	(24) Lone Star Light		(57) Old Milwaukee
	(25) Old Style		(58) Old Milw. Light
	(26) Old Style Light		(59) Pils
	(27) Rainier		(60) Schaefer
	(28) Schmidts		(61) Schlitz
	(29) Sterling		(62) Stroh
	(30) Weidemann		(63) Matts
	(31) White Stag		(64) Utica Club
Genesee [GE, US]:	(32) Genesee	FX Matts [W, U.S.]:	
	(33) Kochs		

⁵These brands correspond to G. Heileman Brewing Co., which was acquired in 1987 by the Australian Bond Corporation Holdings; it is classified as a domestic brewer mainly because this foreign ownership was only temporary.

IRI records the number of units sold regardless of size. A measure of average size is constructed as $SIZE = Quantity/Units$, which is the inverse of the average number of units that are contained in 288 ounces (a 24-pack). The variable COV is construed as the fraction of the potential market to which the brand is available (coverage).⁶ Beers with low coverage may be interpreted as specialty brands that are targeted to a limited fraction of the population.

$SGROC$ measures the fraction of grocery sales accounted for by supermarkets. $GCR4$ is the four firm concentration ratio of supermarkets in a given city and quarter and $LBCR4$ measures the local concentration of sales by the four largest brands. Demographic variables, $SGROC$ and $GCR4$ are only available annually.

In addition to the IRI data, several other sources were used. Advertising data (ADV) was obtained from the Leading National Advertising annual publication.⁷ These are quarterly data by brand comprising total national advertising expenditures for 10 media types. Alcohol content (ALC) was collected from the scientific study of alcohol content conducted by Case, Distefano and Logan. For brands not included in the scientific study several specialized websites (www.realbeer.com, www.brewery.org, www.ratebeer.com, www.beeradvocate.com, www.taproom.com) and the brewers' websites were used as sources. It is assumed that alcohol content has remained constant. While this may be a strong assumption, it is the best proxy available for this characteristic.⁸

⁶This is a more precise measure of coverage than that used by Pinske and Slade and Slade (2004), whose coverage variable is equal to the fraction of establishments carrying the product. Here, the fraction of establishments is adjusted by their size (the "all commodity value", or ACV, of store) with respect to the total market (ACV of all stores).

⁷IRI and LNA data was kindly provided by Ronald Cotterill, Director of the Food Marketing Policy Center at the University of Connecticut.

⁸In some states, low alcohol content limits exist (e.g. Oklahoma and Utah). In these cases, the alcohol content variable is a less accurate proxy for actual alcohol content. This problem is lessened by the inclusion of city dummies, which capture city-specific characteristics such as regulations.

Given the strategic nature of price and advertising, instruments are needed to identify the structural parameters. Data for price instruments were collected from additional sources. A proxy for supermarket labor cost was constructed by averaging the hourly wages (*WAGES*) of interviewed individuals from the Bureau of Labor Statistics (BLS) Current Population Survey (CPS) monthly earning files.⁹ For a given city-quarter combination, individuals working in the retail sector were selected for that city over the corresponding three months. Then, the average was calculated over the number of individuals selected. City density estimates (*DEN*), collected from Demographia and the BLS, were included to proxy for cost of shelf space.¹⁰ This variable is measured in population per square mile.¹¹

Table 3.3 displays summary statistics of the variables described above. The mean price for a 24-pack (288 ounces) is near \$12 while firms spend an average of \$3.54 million in advertising per brand during a quarter. The average package size is slightly smaller than a 12-pack (*SIZE*=0.38 corresponds to 9.6 twelve-ounce containers). Conversely, the average brand is carried by stores that represent 74% of all sales in a market. Beers in the sample have an average alcoholic content of 4.48% per volume and a small variability (standard deviation 0.94). To simplify the presentation of summary statistics, only one product grouping is considered in this table: budget, light, premium, super-premium and import. From all observed brands in different cities and quarters, 37% are budget, 23.5% are light, 18.5% are premium, while super-premium and imports account for 10% (each) of all observations.

⁹These files are available at the National Bureau of Economic Research (NBER), www.nber.org.

¹⁰As in Nevo (2001), it is assumed that more densely populated areas are correlated with more expensive shelf space.

¹¹Density estimates for IRI city definitions were not available. The proxy used was the density of the main metropolitan area in the IRI city definition.

Table 3.3. Summary Statistics

Variable	Mean	Std.	Min	Max	Variable	Mean	Std	Min	Max
Price	12.09	3.87	0.82	28.96	O50K (%)	23.34	6.05	10.3	44.8
Quantity	23.53	63.6	0.0014	2651.6	HISP (%)	7.69	12.47	0.4	71.6
ADV (Mill\$)	3.54	6.3	0	40.29	AGE (years)	33.02	2.37	24.1	41.8
SIZE	0.38	0.117	0.075	1.30	ALC (%/vol)	4.48	0.94	0.4	5.25
COV (%)	74.01	28.61	0.26	100	Budget	0.37	N/A	N/A	N/A
SGROC (%)	75.01	6.05	58.8	95.3	Light	0.235	N/A	N/A	N/A
GCR4 (%)	66.32	12.48	30.2	88.1	Premium	0.185	N/A	N/A	N/A
LBCR (%)	44.7	7.10	29.5	69.3	SPremium	0.10	N/A	N/A	N/A
POP (Mill)	2.64	2.89	0.31	15.7	Import	0.101	N/A	N/A	N/A
HOUS	2.6	0.15	2.3	3.2	WAGES (\$/hr)	7.3	1.17	3.58	12.3
INC (000)	31.996	6.9	18.09	53.4	DEN (000)	4.73	4.13	0.73	23.7
U10K (%)	15.4	3.2	8.6	25					

3.2. Distance Measures

As described in section 2.2, coefficients for rival advertising and prices (b_{jk} and c_{jk}) depend on discrete and continuous distance measures, δ_{jk} .¹² For empirical implementation, weighing matrices are constructed with typical element j, k containing the distance measure between products j and k . Several distance measures are created in an attempt to determine the ones that better explain price and advertising substitutability patterns. Although a priori measures are created, the modeler may test various possibilities, as opposed to other models (e.g. nested logit, multistage budgeting) that constraint the estimation to one grouping whose validity is difficult to test. To ease the interpretation of results, discrete and continuous measures are created in terms of the inverse of distance: closeness. Discrete measures of closeness capture local competition between brands since non-zero values (typically 1) are assigned only to close neighbors (e.g. neighbors that are of the same type or produced by the same brewer). Continuous measures, on the other

¹²The number of distance measures is limited by information on product characteristics.

hand, capture global competition by also allowing distant products to compete, though less strongly, with one another.¹³

3.2.1. Continuous Distance Measures

Three continuous product characteristics are utilized in this study: alcohol content (*ALC*), product coverage (*COV*), and container size (*SIZE*). Consumers may prefer beers with certain alcohol content (e.g. low or high) and thus competition among beers of similar alcohol content might be stronger. Product coverage measures the fraction of the market in a given city that is covered by a brand. Beers with low coverage may be interpreted as specialty brands that are targeted to a particular segment of the population. Beer is sold in a variety of sizes (e.g., six and twelve packs), and the variable *SIZE* measures the average package “size” of a brand. Higher volume brands (e.g., typical sales of twelve packs and cases) may compete less strongly with brands that are sold in smaller packages (e.g., six packs). The distance measures are computed in one-and two-dimensional Euclidean space.

The distance measure is an inverse expression of the distance between brands j and k is defined as: $1/[1+2 \times (\textit{Euclidean distance between } j \textit{ and } k)]$. For a one-dimensional product space, Euclidean distance is the absolute difference in the value of the characteristic between brands j and k . In n -dimensional space, Euclidean distance between two brands, j and k , is equal to $\sqrt{(j_1 - k_1)^2 + \dots + (j_n - k_n)^2}$, where the n subscript represents the brand’s coordinate value in each of the n -dimensions. The one-dimensional matrices are denoted *WALC*, *WCOV*, and *WSIZE* and the two-dimensional matrices are denoted

¹³The calculation of all distances and weighing matrices in this dissertation was performed with algorithms programmed in Matlab. See appendix B for details.

WAC , WAS , and WCS , where A , C , and S stand for alcohol content, product coverage and container size, respectively.

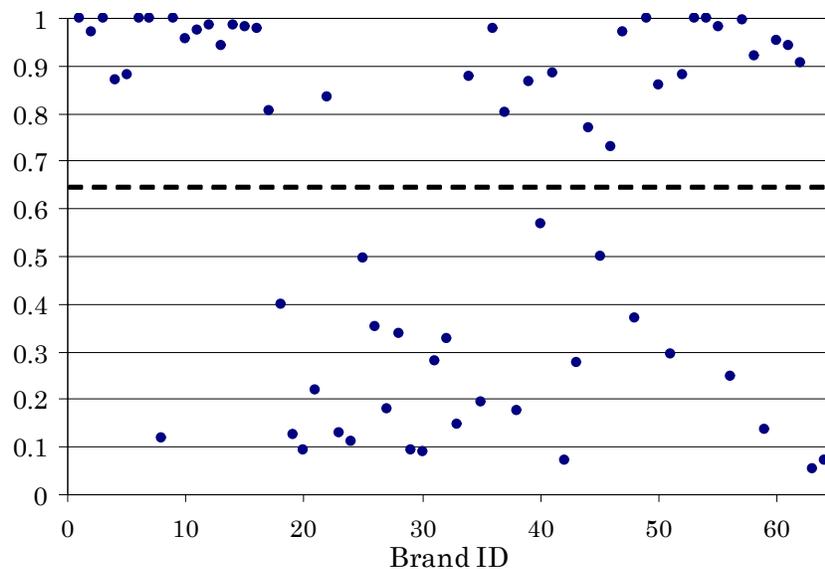
3.2.2. Discrete Distance Measures

Three different types of discrete distance measures are utilized. The first type focuses on various product groupings including product segment, brewer identity, and national brand identity. Previous studies on beer have considered several different product segment classifications. With no clear consensus on product segment classifications five different classifications are considered: (1) budget, light, premium, super-premium, and imports, (2) light and regular, (3) budget, light, and premium, (4) domestic and import, and (5) budget, premium, super-premium, and imports. Beers could also be classified by lagers, ales, porters and stouts, as in the studies of Pinkse and Slade and Slade (2004) on UK brewing. However, lagers account for 90% of sales in the U.S. The weighing matrices for the product segment classifications, denoted $WPROD1$ through $WPROD5$, are constructed such that element (j, k) is equal to one if brands j and k belong to the same product segment and zero otherwise.

A discrete distance measure for brewer identity is utilized to allow the model to determine if consumers are more apt to substitute between brands of the same firm when there are price changes, and if there are rival, or spillover, effects in advertising among beers produced by the same brewer. The weighting matrix $WBREW$ is constructed such that element (j, k) is equal to one if brands j and k are produced by the same brewer and zero otherwise.

As noted in the data section, some brands in the sample are only produced regionally. Figure 3.1 displays the average, over 20 quarters, of the fraction of cities in which a brand appears. Brands are clustered in two groups (divided by the dotted line): one group

Figure 3.1. Fraction of Cities in which a Brand appears (mean of 20 quarters)



that is closer to 1 (i.e. present in nearly all cities) and another group that is closer to zero (present in few cities). This product grouping classifies brands by whether they are regional (below dotted line) or national brands (above the dotted line). The distance measure constructed from this product grouping is used to test whether brands that are national (regional) compete more strongly with each other. The weighting matrix $WREG$ takes a value of one if brands j and k are either both regional or both national, and zero otherwise. According to this regional-national criteria, 15.2% of all observed brands are regional.¹⁴

All weighing matrices constructed from product groupings are normalized so that the sum of each row is equal to one. This normalization allows the weighted prices and

¹⁴Figure 3 suggests that about half of the brands are regional and the rest national and one would expect the mean of REG to be roughly equal to 50%. However, the fact that national brands appear more often (in nearly all cities) drives the mean of REG down to only 15.2%. The mean over 20 quarters is used as there are no important variations in this geographic pattern over time.

advertising expenditures of rival brands that are in the same grouping to equal their average.

Following PSB, two other types of discrete measures are constructed based on the nearest neighbor concept and if products share a common boundary in product space. A (j, k) element of a nearest neighbor matrix is equal to one if brands j and k are nearest neighbors (mutual or not) and zero otherwise. Brands j and k share a common boundary if there is a set of consumers that would be indifferent between both brands and prefer these two brands over any other brand in product space (assuming consumers have a preferred bundle of product characteristics).

A (j, k) element of a common boundary matrix is equal to one if brands j and k share a common boundary and zero otherwise. The nearest neighbor (NN) and common boundary (CB) measures are computed for all brands based on their location in alcohol content and coverage space (weighing matrices $WNNAC$ and $WCBAC$) and coverage and container size space (weighing matrices $WNNCS$ and $WCBCS$).¹⁵

To illustrate how common boundaries are computed, consider the common boundary between brands j and k in Coverage-Size space. Define the coordinates of j and k as (j_{cov}, j_{size}) and (k_{cov}, k_{size}) , respectively. A common boundary between j and k is defined as the set of points COV and $SIZE$ that satisfy the following equality:

$$(3.1) \quad \sqrt{(j_{cov} - COV)^2 + (j_{size} - SIZE)^2} = \sqrt{(k_{cov} - COV)^2 + (k_{size} - SIZE)^2}$$

Solving for $SIZE$ in (3.1) gives the common boundary equation:

¹⁵A discussion of the choice of the relevant product space for constructing these matrices is contained in Chapter 4.

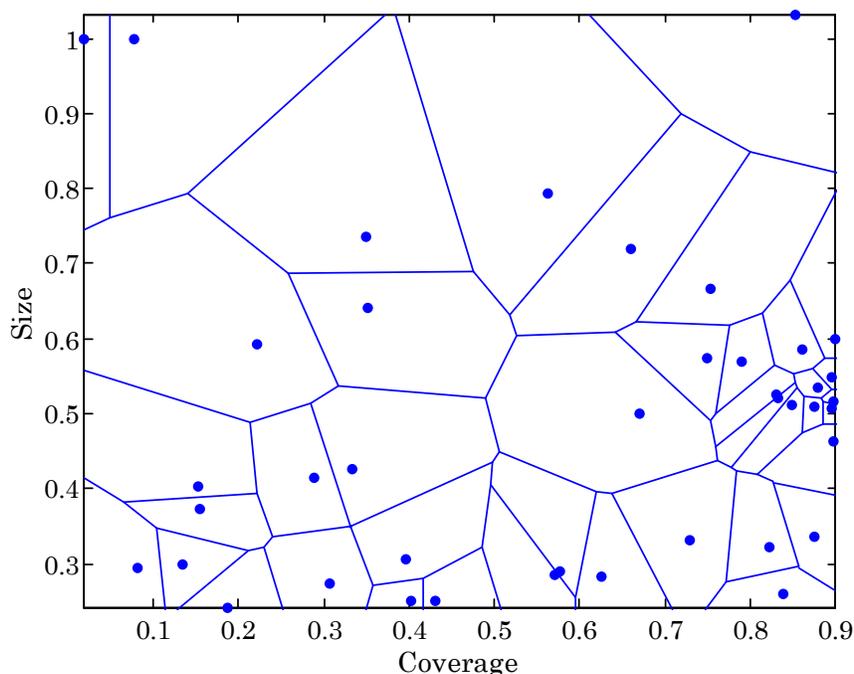
$$(3.2) \quad SIZE = COV \frac{(k_{cov} - j_{cov})}{(j_{size} - k_{size})} + \frac{j_{cov}^2 + j_{size}^2 - k_{cov}^2 - k_{size}^2}{2(j_{size} - k_{size})}$$

which is linear in COV . Once all pair-wise equations are solved for, it is necessary to determine the intersection points (if any) between the equations and establish which portion of the lines determined by these equations are actual common boundaries (see figure 3.2).

Because the continuous product characteristics alcohol content (ALC), product coverage (COV), and container size ($SIZE$) have different units of measurement, their values are rescaled before computing the weighing matrices. To restrict the product space for each of these characteristics to values between 0 and 1, each continuous product characteristic is divided by its maximum value. Restricting the product space in this manner eased the calculation of the common boundaries. Without this restriction, common boundaries of brands located on the periphery of the product space are difficult to define.

To illustrate the location of brands in product space and their common boundaries, figure 3.2 depicts the location of 41 brands in the coverage-size (CS) space in Chicago for the fourth quarter of 1992. The figure also shows the equidistant common boundaries that separate brands. There is a particular clustering of brands that are of medium size, around 0.5 or 12 packs, and that are carried by most of the stores (coverage between 0.8 and 0.9). These brands have a greater number of neighbors and hence face more local competition. If consumers are uniformly distributed over the product space, the area around a brand that is delimited by common boundaries can be regarded as the fraction of consumers who prefer that brand to any other, as in a Hotelling-type model.

Figure 3.2. Location of Common Boundaries of Brands in Coverage-Size Space (Chicago, 4th Quarter of 1992)



In addition to using product characteristics, a second set of nearest neighbor and common boundary measures are computed using product characteristics and price. Including price to calculate the nearest neighbor and common boundary measures allows consumers' brand choices to be influenced by both the distance in characteristics space and in price. For this case, nearest neighbors and common boundaries are identified based on the square of the Euclidean distance between brands and a price differential. The square of the Euclidean distance is employed to reduce complex computations.

To illustrate the complexity of computing common boundaries when price is included, consider the previous example of a common boundary between brands j and k in Coverage-Size space. If the price of j (p_j) is added to the left hand side of (3.1) and

the price of k (p_k) is added to the right hand side of (3.1), equation (3.2) becomes non-linear in COV and the complexity of computing intersection points becomes too costly for the dimension of the data set. For this reason, when price is included in the computation of common boundaries, the square of the Euclidean distance is applied to (3.1):

$$(j_{cov} - COV)^2 + (j_{size} - SIZE)^2 + p_j = (k_{cov} - COV)^2 + (k_{size} - SIZE)^2 + p_k$$

3.3. Own-Price and Own-Advertising Interactions

Two product characteristics are interacted with own-price and own-advertising in the model: the inverse of product coverage ($1/COV$) and the number of common boundary neighbors (NCB). These interactions are included because they had the greatest explanatory power in various model specifications. The inclusion of COV allows own-price and own-advertising coefficients to differ between brands with larger coverage (i.e. popular beers) and brands with low coverage (i.e. specialty beers). The number of common boundary neighbors is a measure of local competition that determines the number of competitors that are closely located to a brand in product space. NCB is computed in product coverage-container size space and alcohol content-coverage space.

3.4. Estimation Details

Given the strategic nature of price and advertising, all terms in equation (2.11) that contain these two variables are treated as endogenous and thus correlated with the unobserved demand shock. To avoid simultaneity bias, an instrumental variables approach is used to consistently estimate the model parameters.

Let n_z be the number of instruments, Z the $(T \times J) \times n_z$ matrix of instruments, S the collection of right hand side variables in equation (2.11) and θ the vector of parameters to be estimated. The generalized method of moments (GMM) estimator is used:

$$\hat{\theta}_{GMM} = (S' P_z S)^{-1} S' P_z \underline{w}$$

where \underline{w} are sales shares in vector form. The consistent estimator for the asymptotic variance is given by:

$$Avar(\hat{\theta}_{GMM}) = (S' P_z S)^{-1}$$

where, $P_z = Z(Z'\hat{\Omega}Z')^{-1}Z$, and $\hat{\Omega}$ is a $(T \times J) \times (T \times J)$ diagonal matrix, with diagonal element ε_j^2 equal to the squared residual obtained from a ‘first step’ 2-stage least squares regression.

3.4.1. Instruments

As in previous work, the instruments employed in this paper rely on the identification assumption that after controlling for brand, city, and time specific effects, the demand shocks are independent across cities. Because beer is produced in large-scale plants and then distributed to various states, prices of a brand across different markets share a common marginal cost component, implying that prices of a given brand are correlated across markets. If the identifying assumption is true, prices will not be correlated with demand shocks in other markets and can hence be used as instruments for other markets. In particular, the average price of a brand in other cities is used as its instrument.

The data employed in this study is based on broadly defined city/regional markets. These broad market definitions, which are similar to those used by the Bureau of Labor

Statistics, reduce the possibility of potential correlation between the unobserved shocks that affect two markets. The reason for a smaller potential correlation with these broad market definitions is that firms (for which these data are gathered) are likely to use the data to design an individual strategy for every market. Also, by including national advertising expenditures in the demand equation, we are attempting to control for advertising related demand shocks that may be correlated across markets. In general, any unobserved regional or national shock, like an interest rate shock will affect demand in various markets and will violate the independence assumption. To further control for such unobserved national shocks, quarterly time dummies are included in the specification.

Although a similar instrument could be constructed for advertising, brand-level advertising expenditures are only observed at the national level in each quarter and are thus invariant across markets. Alternatively, lagged advertising expenditures are used as instruments for advertising. This can be done if the identifying assumption is extended to independence of demand shocks over time, in addition to across cities, and there is correlation of advertising expenditures over time. Since expenditures, x_t , are constructed with price and quantity variables, this term is also treated as endogenous and instrumented with median income (*INC*). A final identification assumption, which is common practice in the literature, is that product characteristics are assumed to be mean independent of the error term.¹⁶

Whereas the identifying assumption of independence of demand shocks across markets may be problematic and difficult to assess, it has been widely used in the literature: Hausman, Leonard and Zona, Slade (1995), Hausman, Nevo (2000b, 2001), Pinkse and Slade, and Slade (2004). Nevo assumes independence of the demand shock over markets

¹⁶Unlike Berry, Levinsohn and Pakes; Pinkse and Slade; and Slade (2004), product characteristics are not used as instruments. This makes estimation more consistent with a broader model in which product characteristics may be endogenous.

and over time, as is assumed here. Despite its wide acceptance, the validity of the proposed instruments is assessed by conducting a formal test. Following Nevo (2000b, 2001), additional instruments for price are created as proxies for city-specific marginal costs and an over-identifying restrictions test is performed. The proxies utilized are city density (DEN) for the cost of shelf space and average wage in the retail sector ($WAGES$) for supermarket labor costs.

As observed by Berry, an additional source of endogeneity may be present in differentiated products industries. Unobserved product characteristics (included in the error term), which can be interpreted as product quality, style, durability, status, or brand valuation, may be correlated with price and produce a bias in the estimated price coefficient. Following Nevo (2001), this source of endogeneity is controlled for by exploiting the panel structure of the data with the inclusion of brand-specific fixed effects. These fixed effects control for the unobserved product characteristics that are invariant across markets, reducing the bias and improving the fit of the model. While brand fixed effects do not control for unobserved product characteristics that are city specific, the instruments discussed at the beginning of this section address this issue.

3.4.2. Brand Fixed Effects

One final detail on demand estimation is that by including brand fixed effects, it is not possible to identify the coefficients for the market-invariant product characteristics directly. These coefficients are recovered using a minimum distance procedure, as suggested by Nevo (2000b, 2001). The estimated coefficients on the brand dummies from the demand equation (in which the market-invariant characteristics and the constant are omitted) are used as the dependent variable in a GLS regression, while the market-invariant product characteristics and a constant are used as the explanatory variables.

CHAPTER 4

Results

4.1. Preliminary OLS Regressions

Given the large number of possible distance measures and high levels of collinearity between these measures, several preliminary OLS regressions are used to determine the most relevant continuous and discrete product spaces for cross-price and cross-advertising terms. Each OLS regression is a restricted version of equation (2.11) in which either one cross-price term or one cross-advertising term is specified. Table (4.1) reports the estimated coefficients and t-statistics on the weighted cross-term using each of the distance measures. Each of these coefficients was estimated in a separate OLS regression.

First, one- and two-dimensional continuous distance measures constructed from alcohol content, product coverage, and container size were used to weigh rival prices and rival advertising. Results for the one-dimensional distance measures indicate that cross-price coefficients appear to depend on closeness in alcohol content and product coverage. The cross-advertising coefficients depend on closeness in product coverage and container size. Results for the two-dimensional distance measures indicate that closeness in alcohol content-product coverage and product coverage-container size space is important for both rival prices and advertising. Because using the same product space for both rival prices and advertising causes the weighted rival prices and advertising to be highly collinear when pooled in one regression, alcohol content-product coverage is assumed to be the relevant product space for weighing rival prices and product coverage-container size is assumed to be the relevant product space for advertising.

Table 4.1. Estimated Coefficient on Weighted Prices and Weighted Advertising from separate OLS Regressions

Distance Measure (Weighing Matrix Acronym)	Cross-Price ^a		Cross-Advertising ^a	
	Coeff ^b	(t-stat)	Coeff ^b	(t-stat)
Continuous Distance Measures				
<i>One-Dimensional</i>				
Alcohol Content (WALC)	1.42**	(2.4)	0.02	(0.44)
Product Coverage (WCOV)	7.65*	(41.0)	0.45*	(57.6)
Container Size (WSIZE)	0.23	(0.74)	0.22*	(11.4)
<i>Two-Dimensional</i>				
Alcohol Content-product coverage (WAC)	10.90*	(34.8)	0.78*	(43.2)
Alcohol Content-container size (WAS)	1.28**	(2.42)	0.17*	(4.93)
Product coverage-container size (WCS)	8.28*	(30.5)	0.58*	(49.8)
Discrete Distance Measures				
<i>Common Boundary (CB)</i>				
Alcohol content-product coverage (WCBAC)	0.83**	(2.08)		
Alcohol content-product coverage-price (WCBACP)	5.20*	(12.1)		
Product coverage-container size (WBCBS)			0.38*	(32.7)
Product coverage-container size-price (WBCBSP)			0.53*	(25.1)
<i>Nearest Neighbor (NN)</i>				
Alcohol Content-product coverage (WNNAC)	11.2*	(20.7)		
Alcohol Content-product coverage-price (WNNACP)	2.60*	(4.73)		
Product coverage-container size (WNNCS)			0.50*	(25.2)
Product coverage-container size-price (WNNCSP)			0.39*	(14.8)
<i>Product Groupings</i>				
National Identity (WREG)	62.6*	(4.49)	0.92 [‡]	(1.70)
Brewer Identity (WBREW)	29.3*	(6.41)	-0.29**	(-2.12)
Product classification 1 (WPROD1) ^c	-2.07	(-0.17)	-1.08*	(-7.84)
Product classification 2 (WPROD2) ^c	116.7*	(7.51)	-1.89*	(-7.46)
Product classification 3 (WPROD3) ^c	19.4	(0.56)	-2.88*	(-13.7)
Product classification 4 (WPROD4) ^c	-82.8*	(-4.87)	-2.03*	(-4.42)
Product classification 5 (WPROD5) ^c	-42.9**	(-2.35)	-0.31**	(-1.96)

^aEach coefficient (and its t-statistic) is obtained from a separate OLS regression in which the coefficient displayed in each cell above corresponds to the only weighted rival term included (i.e. either weighted rival price or weighted rival advertising). All regressions include city, brand, and time binary variables.

^bCoefficients have been multiplied by 10,000 for readability.

^cProduct classifications are: (1) budget, light, premium, super-premium, and imports; (2) light and regular;

(3) budget, light, and premium; (4) domestic and import; and (5) budget, premium, super-premium, and imports.

*Significant at 1%, ** Significant at 5%, [‡] Significant at 10%

Using these relevant continuous product spaces, similar OLS regressions with common boundary and nearest neighbor distance measures are considered. For rival prices, the common boundary measure that includes price and the nearest neighbor measure without price perform better than their counterparts. For rival advertising, the distinction between including or not including price in common boundary and nearest neighbor distance measures is not clear. The t-statistics for the measures without price are slightly larger than their counterparts.

The last set of regressions focus on discrete measures constructed from product groupings. The positive coefficient on rival prices weighted by brewer identity indicates that consumers are more apt to substitute between brands of the same firm. This notion of substitution among a brewer's brands is reinforced by the negative coefficient on this measure for rival advertising which indicates that advertising of a particular brand leads to a reduction in the sales of other brands produced by the brewer. The positive coefficient on rival prices weighted by *WREG* indicates that national brands are closer rivals to each other than to regional brands and vice versa; however, the positive coefficient on rival advertising weighted by *WREG* suggests the existence of spillover effects of advertising among regional brands as well as among national beers.

For weighted rival prices, coefficients using different product segments take positive and negative values. Since brands that belong to the same product segment should be substitutes, the negative coefficients for product segments 4 and 5 have the "wrong sign" and indicate these product classifications are not appropriate. The only positive coefficient that is significant for rival prices is that of product classification 2 (*WPROD2*). This indicates that cross-price effects are larger for same-segment beers (either light or regular). For rival advertising, on the other hand, the coefficients on all product segments are

negative. The largest and most significant coefficient is that of product classification 3 (*WPROD3*). This classification is similar to 2 except that it includes the “budget” category in addition to light and regular.

4.2. Brand Share Equation

Because of the endogeneity of price and advertising, the brand share equation is estimated using the GMM estimator described in section 3.4. Results from OLS regressions presented in the previous section were used to guide the choice of variables in the final specification. However, not all variables were significant while others produced collinearities when pooled in a single regression. For example, there was a high-level of collinearity between the cross-prices weighted by the alcohol content-product coverage (*WAC*) distance measure and cross-advertising expenditures weighted by the product coverage-container size (*WCS*) distance measure. Weighting cross-advertising expenditures by container size only (*WSIZE*) reduced the collinearity problem while not affecting the other parameter estimates. In addition, to allow for variation by brand of the own-price and own-advertising terms, different interactions of price (and advertising) with different product characteristics were tried; the final specification includes those that had greatest explanatory power.

Table 4.2 reports the GMM regression results for two different models. The difference between models 1 and 2 is the inclusion of brand dummies. The two models contain time and city binary variables (estimated coefficients not reported). Because alcohol content, brewer dummies and product segment variables for a brand are constant across time and city, their coefficients can not be directly identified when brand dummies are included in model 2. A minimum distance (MD) procedure is utilized to recover these coefficients (see details in section 3.4.2). A second-stage regression is performed with the

estimated coefficients on brand dummies as the dependent variable and alcohol content (*ALC*), product segments (budget, light, premium, super-premium and import), brewer dummies, and a constant as explanatory variables.

The estimated coefficients from the MD procedure for model 2 are reported in the first set of variables of table 4.2. While the market-invariant product characteristics in the MD procedure explain only 12% of the variation in the coefficients of the brand dummies, all coefficients recovered from the MD procedure, except for the constant, are significantly different from zero at the 1% level. The positive coefficients on the product segment binary variables indicate that these product segments have larger budget shares than the light (or base) product segment. An increase in alcohol content is associated with a larger budget share.

The only product-specific variable that does vary by market is the number of common boundaries in alcohol content-product coverage space (*NCBAC*). The negative coefficient on *NCBAC* shows that brands that share a common boundary with more neighbors in alcohol content-coverage space have a lower sales share than those with fewer common boundaries. Thus, the higher number of close neighbors, the greater the competition between brands. Conversely, the demographic variable *OVER50K* has a negative sign which implies that sales shares tend to be smaller in cities where the fraction of high income families is larger. This finding is consistent with the fact that more than half of beer is consumed by households with an annual income of \$45,000 or less (Beer Institute).

The estimated coefficients for own-price, own-advertising, and their interactions with product characteristics are reported in the second group of variables in table 4.2. Because price and advertising are highly correlated with their corresponding interactions with product coverage, the inverse of this latter variable ($1/COV$) is used to avoid collinearity.

Table 4.2. Results of GMM Estimation of Demand Model

Dependent Variable: Sales Share (w_{jt}) ^a	Model 1		Model 2	
	Coeff	(t-stat) ^b	Coeff	(t-stat) ^b
Constant, a_{jt}				
Brand Dummies		no		yes
Constant ^c	-	-	-15.51	(-0.96)
ALC ^c	-	-	5.95	(3.24)
POPULAR ^c	-	-	49.98	(14.84)
PREMIUM ^c	-	-	63.52	(13.95)
SUPER-PREMIUM ^c	-	-	131.81	(23.85)
IMPORT ^c	-	-	211.18	(22.55)
NCBAC; # common boundary neighbors, Alcohol content-Coverage space	-1.15	(-0.85)	-3.91	(-3.66)
OVER50K	-94.84	(-0.57)	-240.0	(-1.90)
Own Price (b_{jj}) and Own-Advertising (c_{jj})				
$\log P$	-122.4	(-9.82)	-252.9	(-5.71)
$\log P \times (1/COV)$	-0.56	(-2.38)	-1.09	(-3.46)
$\log P \times NCBCSP$; NCBCSP= # of CB neighbors in CS - price space	-4.82	(-7.28)	-7.14	(-11.35)
A^γ	8.48	(31.15)	1.32	(4.39)
$A^\gamma \times (1/COV)$	-0.68	(-5.58)	-0.19	(-3.47)
$A^\gamma \times NCBCS$; NCBCS=# common boundary neighbors, CS space	-1.65	(-3.57)	-0.16	(-4.53)
Weighted Cross-Price and Cross-Advertising Terms (λ_l and τ_m)				
<i>Distance Measures for Price (Weighing Matrix Acronym)</i>				
Alcohol content-product coverage, two-dimensional product space, (WAC)	2.10	(13.66)	5.32	(11.0)
Nearest Neighbors in Alcohol Content-Product Coverage space (WNNAC)	-0.21	(-0.30)	8.87	(15.62)
Brewer Identity (WBREW)	-12.18	(-5.38)	17.30	(5.31)
Product Classification 2: Regular-Light (WPROD2)	52.39	(6.62)	93.56	(3.99)
National Identity (WREG)	40.83	(5.85)	49.61	(5.39)
<i>Distance Measures for Advertising (Weighing Matrix Acronym)</i>				
Container size, one-dimensional product space (WSIZE)	0.17	(7.83)	0.16	(8.64)
Common boundary in product coverage-container size-price space (WCBCSP)	0.85	(15.5)	0.71	(15.23)
Nearest neighbors in product coverage-container size space (WNNCS)	0.61	(14.7)	0.40	(12.24)
Product Classification 3: Budget, light, premium (WPROD3)	-2.78	(14.58)	-3.22	(-9.10)
National Identity (WREG)	-3.02	(-21.79)	5.30	(2.65)
Price Index (d_j)				
$\log(x_t/P_t^L)$	28.15	(1.08)	27.35	(1.38)
R^2 (Centered, uncentered)	0.40, 0.58		0.66, 0.76	
J -Statistic (p-value)	0.90		0.50	

^aBased on 33,892 observations. Coefficients have been multiplied by 10,000 for readability. All specifications include, time and city dummies (not reported). ^bAsymptotic t-statistics in parenthesis.

^cEstimates from minimum distance (MD) procedure. The MD regression includes brewer dummies (not reported).

The own-price and own-advertising coefficients are significantly different from zero at the 1% level and have the expected negative and positive signs. The negative coefficients on the interaction of price and advertising with the inverse of product coverage indicates that as the inverse of brand coverage decreases, the own-price effect for that brand decreases (becomes less negative) while the own-advertising effect increases (becomes more positive). Thus, the sales of brands that are widely sold within a city are less sensitive to a change in price than are brands that are less widely available. Also, advertising is more effective for brands that are more widely sold. Finally, as the number of common boundaries increases the own-price effect increases (becomes more negative) and the own-advertising effect decreases. This shows that higher brand competition is associated with more price responsive demand and less effective advertising.

Comparing models 1 and 2, the estimated own-price coefficient is nearly twice as large in absolute terms when brand dummies are included. Conversely, the own-advertising coefficient decreased by approximately 80 percent in model 2 compared to model 1. The better goodness-of-fit of model 2 and the magnitude of change on both price and advertising coefficients highlight the importance of accounting for endogeneity (resulting from unobserved product characteristics) with the inclusion of brand dummies. Furthermore, the overidentification test in model 2 (p-value=0.50) suggests that the choice of instruments is valid. Discussion of results is henceforth based on model 2.

In model 2, the estimated coefficients on the weighted cross-price terms are all positive. Thus, brands that are closer in the alcohol content-product coverage space (both in terms of Euclidean distance and nearest neighbor), produced by the same brewer, belong to the same product segment, or have similar geographic coverage are stronger substitutes than other brands. Intuitively, consumers will more likely switch to a brand located nearby in

product space and/or produced by the same brewer than to more distant brands. Based on the magnitude of the estimated coefficients, the strongest substitution effects are for brands in the same product segment and with similar geographic coverage.

With the exception of product segment, the estimated coefficients on weighted cross-advertising terms are positive. This suggests that there are spillover effects in advertising across brands that are located more closely in the product space and with the same geographic coverage. However, the negative coefficient for product segment indicates that there are rival cross-advertising effects for brands in the same product segment, thereby potentially offsetting some of the spillover effects. In general, positive and negative cross-advertising effects have the same order of magnitude. As shown in table 4.4, there are more positive cross-advertising elasticities than negative cross-advertising elasticities, indicating that spill-over effects dominate rival effects.

The estimated coefficient on real expenditures, $\log(x_t/P_t^L)$, is not statistically different from zero. Various specifications were tried that interacted product or market characteristics with real expenditures, but none of these specifications yielded statistically significant coefficients. This result implies that the brand-level income elasticities are not statistically different from one.

4.3. Elasticities

Price and advertising elasticities are calculated for each city-quarter pair using the estimated coefficients from the GMM estimation of model 2 in table 4.2. The median own-price elasticity across all brands is -3.34 while the median own-advertising elasticity is 0.024. All own-price elasticities are negative while approximately 85% of own-advertising elasticities are positive. All cross-price elasticities are positive and have a median value of 0.0593 whereas cross-advertising elasticities have a median of 0.021. In general, median

own-price elasticities are slightly lower to those reported in Hausman, Leonard and Zona (-4.98), and Slade (2004) (-4.1). Cross-price elasticities are similar to those in Slade but an order of magnitude smaller than those reported by Hausman, Leonard and Zona.

Tables 4.3 and 4.4 contain a sample of the median values of the price and advertising elasticities for selected brands. To facilitate comparison of the cross-price and cross-advertising patterns, these tables also contain information on the distance measures used in the estimated demand function. Table 4.3 divides brands into light and regular. Brands that are located closer in product space have, in general, higher cross-elasticities. For example, Budweiser, Michelob, Coors, Miller Genuine Draft, and Miller High Life are located close to one another in the product space. The cross-price elasticities between these brands are generally larger than the cross-price elasticities with Keystone, Old Style, Olympia, Pabst, and all light beers. Estimated confidence intervals (not shown in table 4.3) indicate that all price elasticities are significantly different than zero at the 5% level. 95% confidence intervals were computed with 5,000 draws from the asymptotic distribution of the estimated coefficients.

As shown in table 4.4, the median advertising elasticities vary considerably across brands. While all of the own-advertising elasticities in the table and most of the cross-advertising elasticities are positive, there are several negative cross-advertising elasticities. These negative cross-advertising elasticities occur between brands in the same product segment. This is due to the negative coefficient on the cross-advertising term that is weighted by product segment (table 4.2). In these cases, the rival cross-advertising effects for brands in the same product segment outweigh the positive advertising spillover effects from closely located brands. Not all of the advertising elasticity estimates are statistically different than zero. Approximately 85% of negative advertising elasticities and 86% of

positive elasticities are significant at the 5% level (for both own- and cross-advertising elasticities).

Table 4.4. Sample of Median Own- and Cross-Advertising Elasticities

	Bud	Bud	Busch	Michb	Michb	Coors	Coors	Kstone	Kstone	Old	Old St	Olymp	Pabst	MGD	MGD	High	Miller
	Light	Light	Light	Style	Light	Light	Light	Light	Light	Life	Lite						
Coverage	0.96	0.95	0.82	0.94	0.92	0.93	0.95	0.72	0.76	0.54	0.52	0.59	0.72	0.95	0.87	0.95	0.95
SIZE	0.41	0.42	0.47	0.29	0.29	0.40	0.42	0.44	0.45	0.41	0.48	0.46	0.43	0.36	0.36	0.39	0.43
AB BUD	0.0177	0.0138	0.0007	0.0006	0.0022	0.0013	0.0141	0.0053	0.0050	-0.0003	0.0006	0.0002	0.0023	0.0012	0.0082	0.0018	0.0166
AB BUD LT	0.0424	0.0326	-0.0002	0.0067	-0.0015	0.0112	-0.0046	0.0119	-0.002	0.0009	-0.0018	0.0003	0.0051	0.0102	-0.0018	0.0165	-0.0034
AB BUSCH LT	0.1115	-0.0124	0.0037	0.0136	-0.0032	0.0211	-0.0131	0.0415	-0.0040	0.0032	-0.0044	0.0008	0.0178	0.0412	-0.0066	0.0414	-0.0140
AB MICHELOB	0.0267	0.1027	0.0068	0.0521	0.0399	0.0087	0.1075	0.0521	0.0473	-0.0028	0.0049	0.0016	0.0218	0.0161	0.0992	0.0140	0.1267
AB MICHB LT	0.1461	-0.0418	-0.0019	0.0642	0.0316	0.0469	-0.0444	0.0580	-0.0131	0.0035	-0.0111	0.0019	0.0207	0.0543	-0.0128	0.0800	-0.0477
AC COORS	0.0334	0.0963	0.0070	0.0054	0.0157	0.0266	0.1216	0.0499	0.0404	-0.0027	0.0055	0.0012	0.0158	0.0096	0.0815	0.0138	0.1239
AC COORS LT	0.0386	-0.0048	-0.0002	0.0070	-0.0014	0.0128	0.0356	0.0122	-0.0021	0.0012	-0.0031	0.0004	0.0051	0.0099	-0.0021	0.0157	-0.0056
AC KEYSTONE	0.4538	0.3569	0.0248	0.0629	0.0510	0.1047	0.4052	0.1026	0.2052	0.0192	0.0218	0.0007	0.0087	0.1767	0.2552	0.1774	0.4526
AC KEYST LT	0.2709	-0.0330	-0.0012	0.0371	-0.0068	0.0540	-0.0380	0.1408	0.0515	0.0092	-0.0142	0.0025	0.0392	0.1060	-0.0148	0.0981	-0.0426
B OLD STYLE	-0.3792	0.4093	0.0396	-0.0964	0.0830	-0.1047	0.4136	0.2923	0.2645	0.0041	0.1451	0.0075	0.0579	-0.0908	0.2122	-0.1790	0.4387
B OLD STYLE L	0.5477	-0.9079	-0.0417	0.1573	-0.1773	0.1716	-0.8905	0.2678	-0.2785	0.2674	0.0109	0.0116	0.0778	0.1431	-0.3332	0.2451	-1.0173
P OLYMPIA	0.3203	0.2663	0.0184	0.0697	0.0421	0.0947	0.2776	0.0333	0.1247	0.0147	0.0331	0.0011	0.0157	0.1026	0.1523	0.1508	0.3258
P PABST	0.2501	0.2065	0.0136	0.0635	0.0380	0.0766	0.2300	0.0103	0.0765	0.0047	0.0053	0.0007	0.0163	0.0813	0.1128	0.1207	0.2594
PM MGD	0.0189	0.0605	0.0030	0.0044	0.0119	0.0050	0.0653	0.0253	0.0230	-0.0015	0.0026	0.0007	0.0104	0.0230	0.0392	0.0094	0.0746
PM MGD LT	0.0933	-0.0103	-0.0005	0.0143	-0.0016	0.0133	-0.0111	0.0335	-0.0043	0.0034	-0.0048	0.0006	0.0171	0.0566	0.0407	0.0377	-0.0130
PM HIGH LIFE	0.0208	0.0687	0.0036	0.0037	0.0119	0.0062	0.0697	0.0286	0.0257	-0.0013	0.0027	0.0009	0.0120	0.0059	0.0554	0.0403	0.0877
PM MILLER L	0.0283	-0.0038	-0.0002	0.0057	-0.0012	0.0082	-0.0041	0.0092	-0.0015	0.0006	-0.0013	0.0003	0.0040	0.0081	-0.0017	0.0131	0.0291

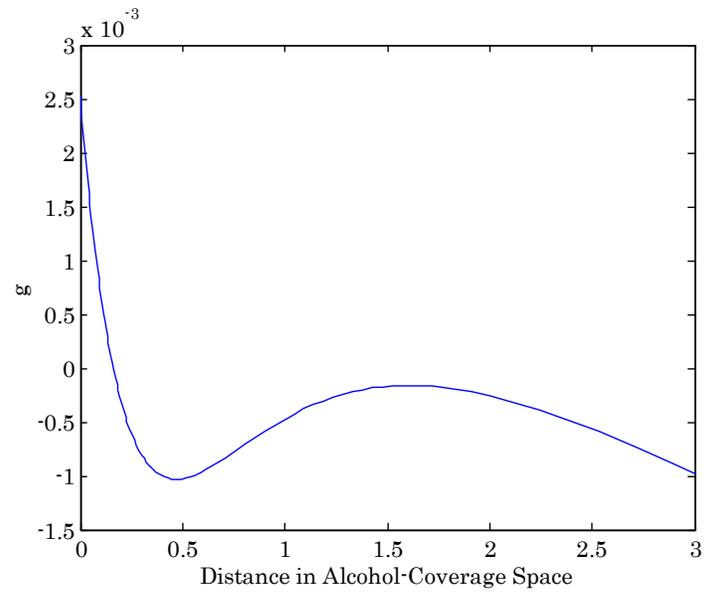
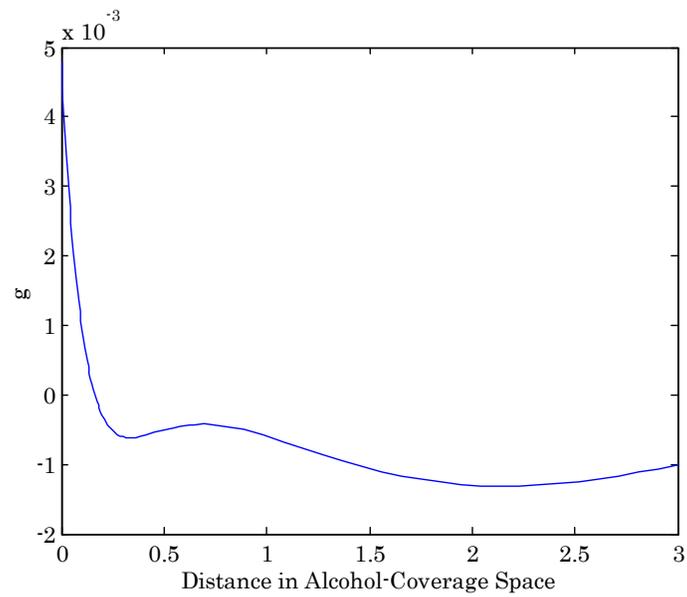
Each entry i, j represents the median i, j advertising elasticity over all markets (i.e. all city-quarter pairs)

4.4. Semi-Parametric Results

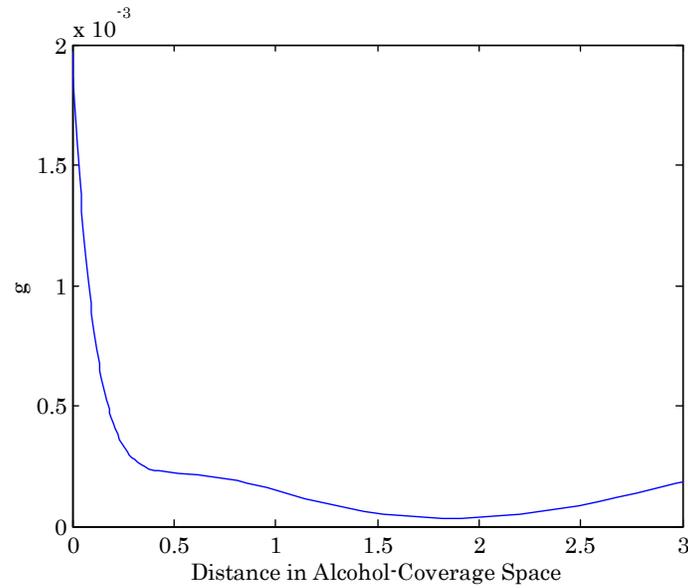
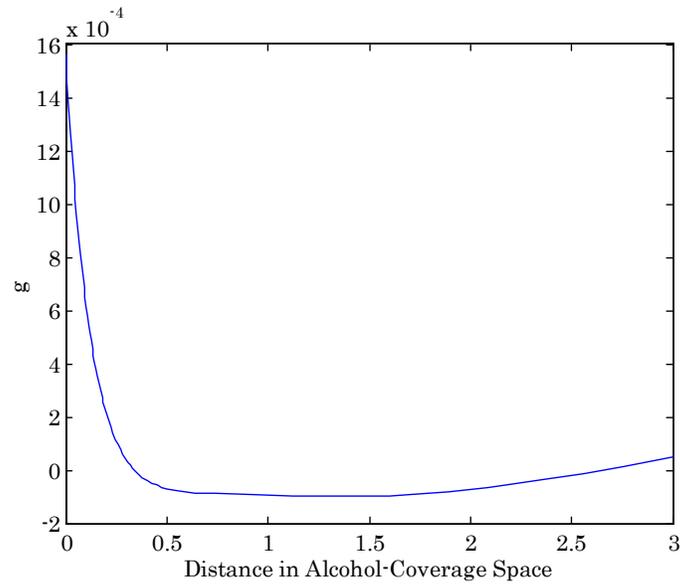
Parametric results are inconsistent if the functional form imposed on g and h is incorrect (see Pinkse, Slade and Brett; footnote 57). When the distance measure is discrete, the parametric estimator is only consistent if the magnitude of the sales share (w_{jt}) response to a change in the price of another brand does not depend on the Euclidean distance. Pinkse, Slade and Brett proposed a consistent semi-parametric estimator for the cross-price and cross-advertising weighing functions g and h that specifies one series expansion of the continuous distance measure for each discrete measure (see section 2.2). Several alternative semi-parametric specifications were estimated as a means of providing some evidence that the parametric specification is not a restrictive version of the functions g and h . Each of the semi-parametric regressions contains either cross-price terms or cross-advertising terms, but not both (due to collinearity problems), and all other variables in table 4.2. In each specification, only one of the discrete measures of the cross-term in table 4.2 (e.g. *BREW* for cross-price or *PROD3* for advertising) and a polynomial series expansion of order 4 is specified for the corresponding continuous measure (i.e. Alcohol-Coverage for price and Size for advertising).

Figures 4.1 through 4.4 plot the function $g(\delta_{jk}^{AC})$, with alcohol-coverage distance on the x -axis, for each of the four different discrete measures relevant for cross-price terms: if brands j and k are nearest neighbors (*NNAC*), produced by the same brewer (*BREW*), in the same product group (*PROD2*), and have similar geographic coverage (*REG*). Each function g was estimated in a separate GMM regression.

Overall, the results show that the value of the g function decays rapidly with Euclidean distance, suggesting that competition among beers is mainly local. All graphs show that g not only decays rapidly but it also stabilizes (around zero). This suggests that

Figure 4.1. Function g for Nearest Neighbors (*NNAC*)Figure 4.2. Function g for Same-Brewer Beers (*BREW*)

each discrete measure does not depend on the Euclidean distance, as required for the consistency of the parametric specification.

Figure 4.3. Function g for Same-Product Segment Beers ($PROD2$)Figure 4.4. Function g for Beers with Similar Geographic Presence (REG)

For the function h the expansion terms were rarely significant and hence it is not plotted here. This indicates not only that non-linear terms are unimportant, but also

that discrete measures are not dependent on the Euclidean distance. Therefore, these results suggest that the parametric specification is not a restrictive version of h as well.

4.5. Summary of Demand Results

In Part I of this dissertation, the Distance Metric method proposed by Pinkse, Slade and Brett is extended to allow the computation of cross-advertising elasticities and to incorporate the flexible Almost Ideal Demand System. These extensions are applied to estimate a brand-level demand system for 64 brands of beer, produced by 13 different brewers, sold in the United States. The U.S. brewing industry is chosen because of the interest it has generated from researchers in the past and for the important role that traditional advertising (television, radio, and press) plays in this industry (Elzinga; Greer; Tremblay and Tremblay).

Much of the previous research on U.S. brewing has utilized aggregated industry data. Accounting for heterogeneity at the firm and brand level may shed additional light on issues raised by previous studies. For example, the results show that the majority of the brand-level cross-advertising elasticities are positive, suggesting that traditional advertising stimulates the overall demand for beer (all else equal). This is contrary to an extensive literature that supports the view that advertising does not stimulate the demand for beer; some examples are Nelson; Nelson and Moran; Lee and Tremblay, and references cited therein. This argument was used by the Federal Trade Commission in a case that dealt with a petition from the Center for Science in the Public Interest (CSPI) in 1983 to ban broadcast advertising of alcohol (including beer). The FTC dismissed the petition on the grounds that advertising does not increase the consumption of alcoholic beverages. In addition to this issue, other policy questions that relate alcohol consumption with health and taxes can be analyzed in more detail with brand level estimates.

In general, the results indicate that brands that are closer in product space have larger substitution coefficients. In particular, light and regular beers (and to a lesser extent budget or popular beers) appear to compete more aggressively with same-segment beers. This is supported by the fact that negative cross-advertising elasticities emerge only between brands that belong to the same segment.

The rising concentration in the U.S. brewing industry, where the sales of the top three brewers account for more than 90% of domestic consumption, and the emergence of Anheuser-Busch as the sole industry leader raise concerns about deviations from competitive behavior (Tremblay and Tremblay: 283). Part II of this dissertation, uses the results of this section to test alternative hypotheses of brand pricing behavior and market power.

Part 2

Price Competition

CHAPTER 5

Background

“A price increase is needed, but it will take Anheuser-Busch to do it”

-Robert Uihlein, Chairman of the Schlitz Brewery, *Fortune* (November, 1975: 92).

“I think the industry is stupid...If they only had price leadership...which they don't have...that isn't violative of anything”

-Allan T. Demaree, legal counsel for the Plumbing Fixture Manufacturers Association, after learning that members had been illegally colluding. *Fortune* (December, 1969: 97-98).

5.1. Introduction

A large part of the empirical work on differentiated products markets has focused on the pricing behavior of firms because it is an important input for addressing issues such as merger simulation, market power and welfare implications of new product introductions. Using an accurate model of pricing behavior in these applications is therefore crucial for making more precise policy inferences. For the most part, however, previous applications have assumed or focused on static Bertrand-Nash and collusive models of pricing behavior.

Anheuser-Busch, the U.S. leading beer producer has a 50% market share and has been identified as a price leader especially through its heavily marketed “King of Beers”

brand Budweiser (Greer; Tremblay and Tremblay; and references therein). For example, in 1954 Anheuser-Busch raised the price of Budweiser after an increase in costs due to a new union wage agreement. Some regional brewers in St. Louis did not follow suit. After this, Anheuser-Busch decided to aggressively reduce the price of Budweiser in St. Louis, which elevated its market share in the region from 12.5% to 39.3%. A few months later, Anheuser-Busch increased the price of Budweiser and this time the regional brewers learned their costly lesson and followed. This and other evidence supports the conclusion that by the 1990's Anheuser-Busch, assisted by this successful punishing strategy, had become the clear price leader (Tremblay and Tremblay: 171; Greer: 49-51).

This chapter empirically assesses different models of price competition in the U.S. brewing industry. Two types of leadership models are considered. The first is a “collusive price leadership” model in which followers match Budweiser’s price changes. The second is a Stackelberg model. In one variant of the Stackelberg model Budweiser acts as the price leader while in the other Anheuser-Busch leads with all its brands. These leadership models are compared to alternative hypothesis of Bertrand-Nash and collusion.

An exogenous variation in the data or “natural experiment” is used to evaluate the various models of pricing behavior. In this paper, the natural experiment is the 1991 100% increase of the federal excise tax on beer.¹ After controlling for other factors, model assessment is carried out by comparing how well the alternative models of firm behavior predict observed prices after the tax change.

To perform the assessment of pricing behavior, estimates of the structural demand system in the previous chapter are used to compute the implied marginal costs for the different models during the pre-increase period. Then, the exogenous increase in the

¹Hausman and Leonard use a similar strategy in which the introduction of a new brand is exploited to evaluate different models of competition.

federal excise tax from \$9 to \$18 per barrel levied on all brewers, together with the pre-increase marginal costs and the demand estimates, are used to compute price changes for all brands that would have prevailed under each model of pricing behavior. These predicted price changes are compared to actual price changes after the tax introduction to assess the predictive power of each model.

5.2. An Overview of Strategies for Assessing Pricing Behavior

Broadly speaking, there are two strategies that may be followed: the menu approach and the conjectural variations approach (CV). With the menu approach, several well-educated guesses of firm conduct are specified and ranked according to model fit measures. In some cases, supply is modeled explicitly for each of the alternatives considered and model comparisons are carried out through pair-wise tests such as non-nested statistics. Some examples include Gasmi, Laffont and Vong; Kadiyali, Vilcassim and Chintagunta; Villas-Boas; and Goldberg and Verboven. Gasmi, Laffont and Vong; and Kadiyali, Vilcassim and Chintagunta estimate demand and supply simultaneously, while Villas-Boas; and Goldberg and Verboven proceed in a two-step fashion by first estimating the demand parameters and then supply relations implied by different game-theoretic scenarios. In other cases, ranking the models involves comparison of observed price-cost margins to price-cost margins implied by the different models, without explicitly estimating a supply equation (e.g. Nevo, 2001).

The CV approach allows for estimation of a continuous conduct parameter that determines the strength (or lack) of competition among firms. The CV approach, however, has been found to be an inaccurate measure of firm conduct under certain circumstances (Corts, 1999). Also, it is not straightforward to attach the value of the conduct parameter to a particular model, especially when many models and many products are being

considered. Finally, the requirements for identification of the conduct parameters in the presence of numerous differentiated products are unlikely to be met (Nevo, 1998).

Given the interest in testing prior beliefs about the price conduct of beer producers, a menu approach is chosen. Model comparisons via non-nested tests is a non-trivial task in this application given the difficulty of obtaining brand level data to model supply explicitly. Alternatively, a natural experiment in the data is used to compare actual prices to those implied by the equilibrium of different models. This procedure is described in the next two chapters.

5.3. The Federal Excise Tax Increase

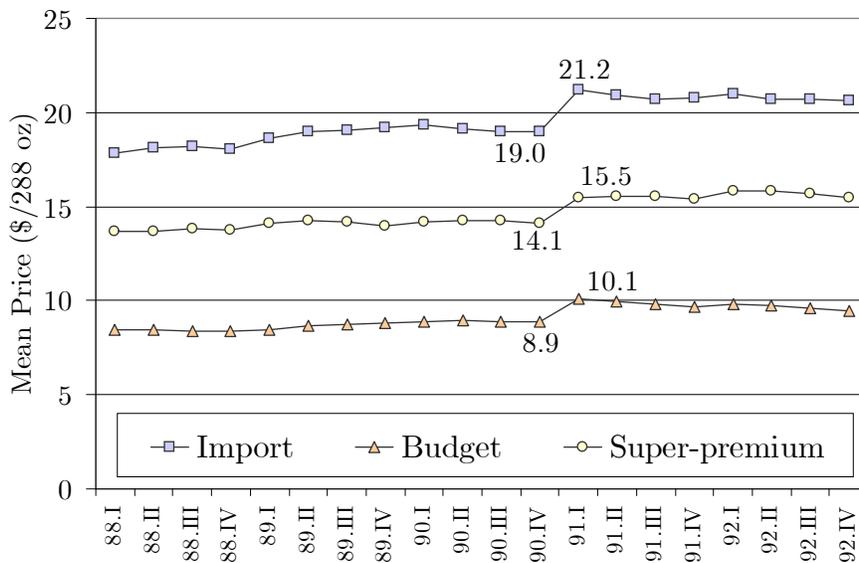
On October 18, 1990, U.S. Congress approved an increase in the federal excise tax on beer from \$9 to \$18 per barrel as part of the Omnibus Budget Reconciliation Act of 1990.² Domestic brewers producing more than 60,000 barrels of beer per year and all importers were required to pay this tax on all units as of January of 1991.³ This increase, which was equivalent to an additional 64 cents in federal taxes per 24-pack (288 ounces), represented the largest federal tax hike for beer in U.S. history.

Figure 5.1 shows mean quarterly prices (over all cities) for three beer segments from the data set described in section 3.1. There is a clear shift in the mean price of all three categories. While the mean price increase in imports is larger than the other two series, all mean increases are higher than the actual tax hike of 64 cents per 288 ounces: \$2.20 for imports, \$1.40 for super-premium beers and \$1.20 for budget beers. These increases were,

²This increase is stipulated in Title XIII, Section 13201 of Public Law 101-508 (also known as HR 5835). This act modified the section 5051 of the Internal Revenue Code regarding the federal excise tax on beer sales.

³Domestic brewers with a production of less than 60,000 barrels per year were exempt of the tax increase on their first 30,000 barrels. These brewers were only required to pay (on the additional units) the \$7/barrel tax stipulated for small domestic producers prior to the approval of this Act.

Figure 5.1. Quarterly Mean Prices, Various Segments (1988-1992)



Source: IRI Database, University of Connecticut

on average, 237%, 114%, and 84%, respectively, larger than the tax increase of 64 cents per case. As shown by Anderson, de Palma and Kreider, in oligopolies with differentiated products an excise tax can be passed on to consumers by more than 100%.

CHAPTER 6

The Supply Model

6.1. Equilibrium Conditions

Formally, let $\Lambda = \{1, \dots, N\}$ denote the set of firms and F_n the set of brands produced by firm $n \in \Lambda$. Assuming constant marginal costs and linear additivity of advertising, the profit of firm n in each market can be expressed as:

$$(6.1) \quad \pi_n = \sum_{j \in F_n} (p_j - c_j) q_j(p, A) - \sum_{j \in F_n} A_j,$$

where c_j denotes brand j 's marginal cost, p_j its price and A_j firm n 's advertising expenditures on brand j . From (6.1), firm n 's first order conditions can be expressed as:

$$(6.2) \quad q_j(p, A) + \sum_{k \in F_n} (p_k - c_k) \left[\frac{\partial q_k}{\partial p_j} + \sum_{m \notin F_n} \frac{\partial q_k}{\partial p_m} \frac{dp_m}{dp_j} \right] = 0, \text{ with respect to } p_j$$

$$(6.3) \quad \sum_{k \in F_n} (p_k - c_k) \frac{\partial q_k}{\partial A_j} - 1 = 0, \text{ with respect to } A_j$$

While derivatives are obtained directly from demand estimates, the term $\frac{dp_m}{dp_j}$ takes different values depending on the model of interest. For example, Bertrand-Nash competition implies that $\frac{dp_m}{dp_j}$ is equal to zero as firms take competitors' prices as given in this model. In principle, different games in advertising may also be modeled by setting $\partial q_k / \partial A_j$ to an expression similar to the one in brackets in equation (6.2). However, given the small magnitude of advertising coefficients obtained from the demand equation, advertising games (collusion, Bertrand-Nash and Stackelberg) produced equilibrium conditions

that were essentially indistinguishable from each other. As a consequence, price is treated as the main strategic variable of interest while it is assumed that firms compete in a Bertrand-Nash fashion in advertising.

Stackelberg Leadership

For the Stackelberg leadership game, the term $\frac{dp_m}{dp_j}$ takes a value of zero if j is a follower brand. If j is a leading brand the term $\frac{dp_m}{dp_j}$ is computed from the first order conditions of all followers by applying the implicit function theorem.¹ More specifically, define a partition of the product set as $\Psi = (\Psi_f, \Psi_l)$, where Ψ_f is the set of follower brands and Ψ_l is the set of leading brands. For each leader, a system of J^F equations (where J^F is the number of followers) is constructed. The l^{th} system of equations is then used to compute the vector that contains all the $\frac{dp_m}{dp_l}$ terms for leader l . Each equation in system l corresponds to a follower brand. An equation in system l is obtained by totally differentiating the price first order condition of a follower brand (6.2)² with respect to all followers' prices (p_f , for all $f \in \Psi_f$) and the price of the l^{th} leader, p_l ($l \in \Psi_l$):

$$\begin{aligned}
 & \sum_{f \in \Psi_f} \underbrace{\left[\frac{\partial q_j}{\partial p_f} + \sum_{k \in \Psi_f} \left(\Delta_{kj}^* (p_k - c_k) \frac{\partial^2 q_k}{\partial p_j \partial p_f} \right) + \Delta_{fj}^* \frac{\partial q_f}{\partial p_j} \right]}_{g(j, m)} dp_f + \\
 & + \underbrace{\left[\frac{\partial q_j}{\partial p_l} + \sum_{k \in \Psi_f} \left(\Delta_{kj}^* (p_k - c_k) \frac{\partial^2 q_k}{\partial p_j \partial p_l} \right) \right]}_{h(j, l)} dp_l = 0; \quad j, k, f \in \Psi_f
 \end{aligned}
 \tag{6.4}$$

¹Villas-Boas uses a similar strategy and notation in her study of models of vertical competition.

²It is assumed that the first order condition with respect to advertising (6.3) does not play a role in deriving $\frac{dp_m}{dp_j}$. Without this assumption, inversion of matrix G below is not possible since it is not a square matrix. Results are unlikely to be sensitive to this assumption given the estimated small impact advertising has on demand.

where Δ_{jk}^* takes the value of one if brands j and k are produced by the same firm and zero otherwise. Therefore, for a given leader l there are J^F equations like (6.4). Let G be the $(J^F \times J^F)$ matrix that contains all g elements above and define the $(J^F \times 1)$ vectors D_s and H_l as:

$$D_s = \begin{bmatrix} dp_1 \\ \cdot \\ \cdot \\ \cdot \\ dp_{J^F} \end{bmatrix}; \quad H_l = \begin{bmatrix} -h(1, l) \\ \cdot \\ \cdot \\ \cdot \\ -h(J^F, l) \end{bmatrix}$$

For a given p_l , and using the above definitions, (6.4) can be written in matrix notation as:

$$GD_s - H_l dp_l = 0$$

where dp_l is treated as a scalar (for matrix operations). The J^F derivatives of the followers' prices with respect to a given p_l can then be computed as:

$$(6.5) \quad \frac{D_s}{dp_l} = G^{-1}H_l$$

Concatenating the $(J - J^F)$ vectors of dimension $(J^F \times 1)$ given in (6.5) (i.e. one vector for each p_l) gives $D = G^{-1}H$. The $(J^F) \times (J - J^F)$ matrix D has all the derivatives of the followers' prices with respect to all of the leaders' prices. Specifically, D has typical

element $\frac{dp_f}{dp_l}$, for $f \in \Psi_f$ and $l \in \Psi_l$. Note that to obtain D , all terms in (6.4) must be known. In particular, this procedure is feasible if marginal costs (c_k) can be computed before D is calculated. If only a subset (i.e. not the complete product line) of a firm's brands are chosen as followers, marginal costs can not be recovered (see section 6.2).

Collusive Price Leadership

Based on accounts of the industry, it may be assumed that followers exactly match Budweiser's price changes. In this case, called "collusive price leadership" (see Scherer and Ross: 248), only the first order conditions of the firm producing the leading brand (i.e. Anheuser-Busch) are relevant, since followers do not price via profit-maximization but by imitating the leader. The term $\frac{dp_m}{dp_l}$ in (6.2) is set to 1 in Budweiser's first order condition and it is set to zero for Anheuser-Busch's remaining first order conditions. In the sample used, Anheuser-Busch produces 10 brands other than Budweiser. With collusive price leadership, these 10 brands' first order conditions are the same as in the Bertrand-Nash case (i.e. $\frac{dp_m}{dp_j} = 0$) while Budweiser's first order condition has $\frac{dp_m}{dp_l} = 1$.

6.2. Implied Marginal Costs

In each market, there are J unknown marginal costs (c_k) to be computed but $2 \times J$ first order conditions that contain marginal cost. One solution is to add up (6.2) and (6.3) for each brand j :³

$$(6.6) \quad q_j(p, A) - 1 + \sum_{k \in F_n} (p_k - c_k) \left[\frac{\partial q_k}{\partial p_j} + \sum_{m \notin F_n} \frac{\partial q_k}{\partial p_m} \frac{dp_m}{dp_j} + \frac{\partial q_k}{\partial A_j} \right] = 0,$$

³If c_k is the same in both equations, the same solution is obtained by subtracting both equations, or by adding both equations, or by setting one equation equal to the other.

and obtain a solution for c_k in this new system. Since this is a linear problem, the solution to (6.6) is unique. Moreover, if c_k is the same in both (6.2) and (6.3), which in this case it is by assumption, the solution to (6.6) will also be a solution to (6.2) and (6.3) individually.⁴

In vector notation, (6.6) can be expressed as:

$$(6.7) \quad Q^o - \Delta(p - c) = 0,$$

Q^o and $(p-c)$ are $J \times 1$ vectors with elements $(q_j(p, A) - 1)$ and $(p_j - c_j)$, respectively; Δ is a $J \times J$ matrix with element $\Delta_{jk} = -\Delta_{jk}^* \left[\frac{\partial q_k}{\partial p_j} + \sum_{m \notin F_n} \frac{\partial q_k}{\partial p_m} \frac{dp_m}{dp_j} + \frac{\partial q_k}{\partial A_j} \right]$, where Δ_{jk}^* takes a value of 1 if brands j and k are produced by same firm and zero otherwise. Marginal costs are then given by:

$$(6.8) \quad c = p - \Delta^{-1}Q^o$$

Bertrand-Nash Model

Using the demand estimates, (6.8) can be computed in each market (i.e. in each city-quarter) by simple inversion. For the Bertrand-Nash model, this is done by setting $\frac{dp_m}{dp_j}$ equal to zero and constructing the Δ matrix with the corresponding price and advertising derivatives.

To illustrate, consider the following example. Suppose there are two firms in the market: Anheuser-Busch produces brands 1 and 2 and Miller produces brands 3 and 4. The four equations given by (6.6) are:

⁴If, on the other hand, two different c_k 's solve (6.2) and (6.3), the solution to (6.6) will give a linear combination of the two c_k 's.

$$\begin{aligned}
q_1 - 1 + (p_1 - c_1) \left[\frac{\partial q_1}{\partial p_1} + \frac{\partial q_1}{\partial A_1} \right] + (p_2 - c_2) \left[\frac{\partial q_2}{\partial p_1} + \frac{\partial q_2}{\partial A_1} \right] &= 0, \\
q_2 - 1 + (p_1 - c_1) \left[\frac{\partial q_1}{\partial p_2} + \frac{\partial q_1}{\partial A_2} \right] + (p_2 - c_2) \left[\frac{\partial q_2}{\partial p_2} + \frac{\partial q_2}{\partial A_2} \right] &= 0,
\end{aligned}$$

for Anheuser-Busch and:

$$\begin{aligned}
q_3 - 1 + (p_3 - c_3) \left[\frac{\partial q_3}{\partial p_3} + \frac{\partial q_3}{\partial A_3} \right] + (p_4 - c_4) \left[\frac{\partial q_4}{\partial p_3} + \frac{\partial q_4}{\partial A_3} \right] &= 0, \\
q_4 - 1 + (p_3 - c_3) \left[\frac{\partial q_3}{\partial p_4} + \frac{\partial q_3}{\partial A_4} \right] + (p_4 - c_4) \left[\frac{\partial q_4}{\partial p_4} + \frac{\partial q_4}{\partial A_4} \right] &= 0,
\end{aligned}$$

for Miller.

The 4×4 Δ matrix is thus:

$$\begin{bmatrix}
\frac{\partial q_1}{\partial p_1} + \frac{\partial q_1}{\partial A_1}; & \frac{\partial q_2}{\partial p_1} + \frac{\partial q_2}{\partial A_1}; & 0 & 0 \\
\frac{\partial q_1}{\partial p_2} + \frac{\partial q_1}{\partial A_2}; & \frac{\partial q_2}{\partial p_2} + \frac{\partial q_2}{\partial A_2}; & 0 & 0 \\
0 & 0 & \frac{\partial q_3}{\partial p_3} + \frac{\partial q_3}{\partial A_3}; & \frac{\partial q_4}{\partial p_3} + \frac{\partial q_4}{\partial A_3} \\
0 & 0 & \frac{\partial q_3}{\partial p_4} + \frac{\partial q_3}{\partial A_4}; & \frac{\partial q_4}{\partial p_4} + \frac{\partial q_4}{\partial A_4}
\end{bmatrix}$$

Inverting this matrix and replacing it into (6.8) gives the desired marginal costs. While marginal costs for Stackelberg, collusion, and collusive price leadership models are carried out in the same fashion, the matrix Δ is defined differently for each model. Details and considerations for computing the marginal costs in these models are presented next.

Stackelberg Model

Operationally, marginal costs are also obtained by applying (6.8). However, some elements of the matrix Δ contain the derivative $\frac{dp_m}{dp_l}$ (for $l \in \Psi_l$). This derivative is computed

via (6.5). Several technical difficulties arise in this model. First, there is an immense number of possible Stackelberg scenarios. Given the motivation in this dissertation, only the case in which Anheuser-Busch acts as a leader, both with all its brands as well as with Budweiser only, are considered.

Second, since the term $\frac{dp_m}{dp_l}$ in the leaders' first order conditions is a function of followers' marginal costs (see equations 6.4 and 6.5), marginal costs of followers need to be computed before $\frac{dp_m}{dp_l}$ is calculated. When Anheuser-Busch acts as a leader with all its brands, followers' marginal costs can be found by applying (6.8) to the smaller system of dimension J^F made up of followers' first order conditions. Once marginal costs of followers are computed, these are used to calculate $\frac{dp_m}{dp_l}$ and then to estimate the marginal costs of the leaders. When Budweiser is a sole brand leader, the term $\frac{dp_m}{dp_l}$ is set to zero if m is produced by Anheuser-Busch, except when the brand is Budweiser. Thus, it is assumed that Budweiser only leads brands produced by rival firms (i.e. not by Anheuser-Busch). The reason for this assumption is that marginal costs (c_k) for Anheuser-Busch's remaining 10 brands that are needed to compute $\frac{dp_m}{dp_l}$ can not be calculated via (6.8) as this system has 11 unknown marginal costs but only 10 equations.

Again, consider the 2-firm, 4-brand example given above to illustrate how marginal costs are recovered for the two types of Stackelberg scenarios. When Anheuser-Busch leads with all its brands, the Δ matrix is defined as:

$$(6.9) \quad \left[\begin{array}{cc|cc} \frac{\partial q_1}{\partial p_1} + \frac{\partial q_1}{\partial p_3} \frac{dp_3}{dp_1} + \frac{\partial q_1}{\partial p_4} \frac{dp_4}{dp_1} + \frac{\partial q_1}{\partial A_1}; & \frac{\partial q_2}{\partial p_1} + \frac{\partial q_2}{\partial p_3} \frac{dp_3}{dp_1} + \frac{\partial q_2}{\partial p_4} \frac{dp_4}{dp_1} + \frac{\partial q_2}{\partial A_1}; & 0 & 0 \\ \frac{\partial q_1}{\partial p_2} + \frac{\partial q_1}{\partial p_3} \frac{dp_3}{dp_1} + \frac{\partial q_1}{\partial p_4} \frac{dp_4}{dp_1} + \frac{\partial q_1}{\partial A_2}; & \frac{\partial q_2}{\partial p_2} + \frac{\partial q_2}{\partial p_3} \frac{dp_3}{dp_1} + \frac{\partial q_2}{\partial p_4} \frac{dp_4}{dp_1} + \frac{\partial q_2}{\partial A_2}; & 0 & 0 \\ 0 & 0 & \frac{\partial q_3}{\partial p_3} + \frac{\partial q_3}{\partial A_3}; & \frac{\partial q_4}{\partial p_3} + \frac{\partial q_4}{\partial A_3} \\ 0 & 0 & \frac{\partial q_3}{\partial p_4} + \frac{\partial q_3}{\partial A_4}; & \frac{\partial q_4}{\partial p_4} + \frac{\partial q_4}{\partial A_4} \end{array} \right]$$

Marginal costs in the Stackelberg case are obtained in several steps:

- (1) Marginal costs of followers are obtained. This is done by applying (6.8) to the smaller system comprised of followers equations. In this example, the smaller Δ matrix has the four non-zero elements displayed in the lower right portion of matrix (6.9). Inversion is applied to this smaller matrix and the marginal costs of Miller brands are recovered via (6.8).
- (2) Terms dp_m/dp_j in leaders' equations are computed. This is done via (6.5) which employs demand estimates and the marginal costs of followers (obtained in step 1). In this example, the terms dp_m/dp_j appear in the non-zero elements displayed in the upper left portion of matrix (6.9).
- (3) Marginal costs of leaders are calculated. This is done by applying (6.8) to the smaller system comprised of leaders equations. In this example, the smaller Δ matrix has the four non-zero elements displayed in the upper left portion of matrix (6.9). These four elements contain the dp_m/dp_j terms computed in the previous step. Inversion is then applied to this smaller matrix and the marginal costs of Anheuser-Busch brands are recovered via (6.8).

When Anheuser-Busch leads with Budweiser only, and assuming Budweiser is brand number 1, then matrix Δ is defined as:

$$(6.10) \quad \begin{bmatrix} \frac{\partial q_1}{\partial p_1} + \frac{\partial q_1}{\partial p_3} \frac{dp_3}{dp_1} + \frac{\partial q_1}{\partial p_4} \frac{dp_4}{dp_1} + \frac{\partial q_1}{\partial A_1}; & \frac{\partial q_2}{\partial p_1} + \frac{\partial q_2}{\partial p_3} \frac{dp_3}{dp_1} + \frac{\partial q_2}{\partial p_4} \frac{dp_4}{dp_1} + \frac{\partial q_2}{\partial A_1}; & 0 & 0 \\ \frac{\partial q_1}{\partial p_2} + \frac{\partial q_1}{\partial A_2}; & \frac{\partial q_2}{\partial p_2} + \frac{\partial q_2}{\partial A_2}; & 0 & 0 \\ 0 & 0 & \frac{\partial q_3}{\partial p_3} + \frac{\partial q_3}{\partial A_3}; & \frac{\partial q_4}{\partial p_3} + \frac{\partial q_4}{\partial A_3} \\ 0 & 0 & \frac{\partial q_3}{\partial p_4} + \frac{\partial q_3}{\partial A_4}; & \frac{\partial q_4}{\partial p_4} + \frac{\partial q_4}{\partial A_4} \end{bmatrix}$$

Computation of marginal costs follow the same three steps explained above. The only difference is in the way the second row of the matrix Δ is defined.

Collusive Price Leadership Model

In this case, only Anheuser-Busch's marginal costs need to be obtained since first order conditions of other firms are not relevant. These marginal costs are also recovered by applying (6.8) to a system of dimension J^L (where J^L is the number of brands produced by Anheuser-Busch) and by setting $\frac{dp_m}{dp_i}$ to 1 in Budweiser's first order condition and zero for the remaining first order conditions of Anheuser-Busch. Using the previous 2-firm, 4-brand example, Anheuser-Busch's Δ matrix would be defined as:

$$(6.11) \quad \begin{bmatrix} \frac{\partial q_1}{\partial p_1} + \frac{\partial q_1}{\partial p_3} + \frac{\partial q_1}{\partial p_4} + \frac{\partial q_1}{\partial A_1}; & \frac{\partial q_2}{\partial p_1} + \frac{\partial q_2}{\partial p_3} + \frac{\partial q_2}{\partial p_4} + \frac{\partial q_2}{\partial A_1} \\ \frac{\partial q_1}{\partial p_2} + \frac{\partial q_1}{\partial A_2}; & \frac{\partial q_2}{\partial p_2} + \frac{\partial q_2}{\partial A_2} \end{bmatrix}$$

Collusive Model

Any collusive possibilities (e.g. between specific products or firms) can be investigated by appropriately modifying the ownership elements Δ_{jk}^* . For example, full collusion, or joint profit maximization, can be represented by setting $\Delta_{jk}^* = 1$ for all j, k in the product

set. Similarly, if two firms are colluding, the appropriate elements Δ_{jk}^* are replaced with a 1 to reflect the new structure.⁵ For full collusion between Anheuser-Busch and Miller in the previous example, the Δ matrix is modified to equal:

$$\begin{bmatrix} \frac{\partial q_1}{\partial p_1} + \frac{\partial q_1}{\partial A_1}; & \frac{\partial q_2}{\partial p_1} + \frac{\partial q_2}{\partial A_1}; & \frac{\partial q_3}{\partial p_1} + \frac{\partial q_3}{\partial A_1}; & \frac{\partial q_4}{\partial p_1} + \frac{\partial q_4}{\partial A_1} \\ \frac{\partial q_1}{\partial p_2} + \frac{\partial q_1}{\partial A_2}; & \frac{\partial q_2}{\partial p_2} + \frac{\partial q_2}{\partial A_2}; & \frac{\partial q_3}{\partial p_2} + \frac{\partial q_3}{\partial A_2}; & \frac{\partial q_4}{\partial p_2} + \frac{\partial q_4}{\partial A_2} \\ \frac{\partial q_1}{\partial p_3} + \frac{\partial q_1}{\partial A_3}; & \frac{\partial q_2}{\partial p_3} + \frac{\partial q_2}{\partial A_3}; & \frac{\partial q_3}{\partial p_3} + \frac{\partial q_3}{\partial A_3}; & \frac{\partial q_4}{\partial p_3} + \frac{\partial q_4}{\partial A_3} \\ \frac{\partial q_1}{\partial p_4} + \frac{\partial q_1}{\partial A_4}; & \frac{\partial q_2}{\partial p_4} + \frac{\partial q_2}{\partial A_4}; & \frac{\partial q_3}{\partial p_4} + \frac{\partial q_3}{\partial A_4}; & \frac{\partial q_4}{\partial p_4} + \frac{\partial q_4}{\partial A_4} \end{bmatrix}$$

6.3. Predicted Prices with Higher Excise Taxes

The tax increase is combined with the pre-tax increase marginal cost (6.8) and the demand estimates to compute each model's predicted equilibrium prices in the post-increase period. Since excise taxes were increased for all beers at a uniform rate of E per unit, predicted prices in the first quarter of 1991 (91.1) can be computed in each city by solving for $p_j^{91.1}$ ($j = 1, \dots, J$) in the following system of non-linear equations:

$$q_j(p_j^{91.1}, A) - 1 + \sum_{k \in F_f} (p_k^{91.1} - c_k^{90.4} - E) \left[\frac{\partial q_k}{\partial p_j} + \sum_{m \in F_n} \frac{\partial q_k}{\partial p_m} \frac{dp_m}{dp_j} + \frac{\partial q_k}{\partial A_j} \right] = 0, \text{ for } j = 1, \dots, J$$

where the superscript 90.4 denotes the fourth quarter of 1990 (the quarter prior to the tax increase).⁶ Both q_j and the derivatives in brackets are non-linear functions of price ($p_j^{91.1}$) so the search includes these terms as well. Other variables (i.e. advertising expenditures, distance measures, product characteristics and total expenditures x_t) are

⁵In practice, the ownership structure when two firms are colluding is the same to when two firms are merging.

⁶To avoid sensitivity to potential outliers in quarter 90.4, the median city-specific marginal cost of brand k over the period 1988-1990 is used for $c_k^{90.4}$.

held constant at time 90.4 values and parameters are those obtained from demand estimation.⁷ Because the functional form of demand constitutes only a local approximation to any unknown demand function, demand parameters can potentially differ between the two regimes (pre- and post-tax increase). However, aside from slightly larger standard errors, demand estimates with pre-increase data produced results that were essentially the same as those obtained with the full sample. Demand estimates, therefore, appear to be robust to the price shift due to the tax introduction.

In some cities, a few brands (1 or 2) exited or entered the market between the fourth quarter of 1990 and the first quarter of 1991. In these cities, the search was performed with the subset of brands that were present in both quarters. While this simplification may be problematic since price movements due to changes in the product set are ignored, the potential bias is likely to be small as these are marginal brands in terms of sales.

The predicted prices are computed in every city and for each brand. This system is solved by using the iterative Newton algorithm for large-scale problems provided by *Matlab*. While convergence is quickly achieved for the Bertrand-Nash and collusive models, it takes several hours to solve the system in all cities for the leadership cases.

6.4. Actual Price Increases

A mean estimate of the observed price increase for each brand is calculated by estimating a separate regression of the following form for each brand⁸:

$$(6.12) \quad p_{yz} = \theta_z + \eta' I + \xi_{yz}$$

⁷Because advertising expenditures play a small role in demand and do not change significantly after the tax increase, results are invariable to whether pre- or post-increase advertising is used in the simulation.

⁸Hausman and Leonard follow a similar strategy.

where p_{yz} is price in quarter y and city z (i.e. each city-quarter pair y, z corresponds to a market t), θ_z are city fixed effects, I is a vector quarter dummy variables and η its corresponding vector of coefficients. If the dummy on the fourth quarter of 1990 is omitted (i.e. this is the reference quarter), the coefficient on the dummy for the first quarter of 1991 can be interpreted as the absolute mean price increase for that brand due to the tax increase. This coefficient, however, captures the mean effect on price of all city-invariant factors present in the first quarter of 1991 (i.e. other national shocks besides the tax increase). A dummy variable that takes a value of 1 in the first quarter of each year was included in (6.12) to control for a possible seasonality effect.

The main reason for estimating the mean of the actual price increase via (6.12) is to obtain confidence intervals that allow comparison with the mean of the predicted price increase while holding city-specific effects constant. In addition, model comparison below also considers the actual price change for each brand in each city defined as: $(p_{actual}^{91.1} - p_{actual}^{90.4})$, where $p_{actual}^{91.1}$ is the price of the brand in the first quarter of 1991 and $p_{actual}^{90.4}$ is the price of the brand in the fourth quarter of 1990.

CHAPTER 7

Results of Model Comparisons

7.1. Implied Price-Cost Margins

Implied marginal costs in the pre-tax increase period are calculated for each model according to details presented in section 6.2. Comparing these marginal costs across models is informative about differences in the equilibrium predictions of the models. Because price-cost margins as a fraction of price ($[p - c]/p$) are more readily interpretable than marginal costs, pre-tax increase summary statistics of this measure (in percentage format) are presented in table 7.1. Six different models are considered: Bertrand-Nash; two Stackelberg scenarios: firm leadership by Anheuser-Busch and brand leadership by Budweiser; collusive leadership by Budweiser; and two collusive scenarios: collusion of the three leading firms (Anheuser-Busch, Coors and Miller) and collusion of three brands, Budweiser, Coors and Miller High Life (the main regular brands of leading firms).¹

The mean price-cost margins for Bertrand-Nash competition, Budweiser as a Stackelberg leader, Anheuser-Busch as a Stackelberg leader, and Budweiser-Coors-Miller High Life collusion do not differ substantially from one another. Both Anheuser-Busch as a Stackelberg leader and collusion among three brands predict similar mean price-cost margins that are slightly higher than Bertrand-Nash. As it will become apparent in the next section, the similarity between Bertrand-Nash and the Stackelberg models is due to the almost negligible slope of the reaction functions of followers to changes in the price of

¹The full collusion case produced unlikely price-cost margins (over 100%). Hence, these two plausible collusive scenarios were explored.

Table 7.1. Summary Statistics of Price Cost Margins for Different Scenarios, (1988-1990)*

	Mean	Median	St. Dev
Bertrand-Nash	39.50	36.68	25.96
Anheuser-Busch Stackelberg Leadership	40.08	37.58	26.25
Budweiser Stackelberg Leadership	39.52	36.68	26.00
Collusive Leadership (Budweiser)**	70.51	60.88	45.09
Collusion 3 firms [§]	46.02	46.95	29.14
Collusion 3 brands [±]	40.20	37.17	26.50

* Margins are defined as $(p - c)/p$ (in percentage). Presented are the summary statistics of the 18,369 (brand-city-quarter) observations in the pre-increase period (1988-1990).

** Price-cost margins for Anheuser-Busch brands only

[§] Anheuser-Busch, Adolph Coors, Miller (Philip Morris)

[±] Budweiser, Coors, Miller Genuine Draft

the leader(s) $\left(\frac{dp_m}{dp_j}\right)$. While Budweiser, Coors and Miller High Life have larger price-cost margins in the three-brand collusion model, the mean price-cost margin only increases by a small amount thereby making it similar to that of Bertrand-Nash.

Collusion among the 3 largest firms predicts the largest mean price-cost margin. Since in the collusive price leadership scenario price-cost margins are only computed for Anheuser-Busch brands, summary statistics for this case are not directly comparable with those of other models. However, price-cost margins are unreasonably large for Budweiser (mean price-cost margin of 164% vs. 82% in Bertrand-Nash, not shown) and similar to Bertrand-Nash price-cost margins for other brands (mean price-cost margin 56% vs 54% in Bertrand-Nash, not shown).

Medians and standard deviations are similar across models, except for the 3-firm collusion and the collusive price leadership scenarios. In all models price-cost margins vary considerably across brands. When calculating each model's predicted prices, this heterogeneity plays an important role.

7.2. Predicted vs. Actual Price Increases

In this section, predicted price increases, $(p_{predicted}^{91.1} - p_{predicted}^{90.4})$ are compared to actual price increases $(p_{actual}^{91.1} - p_{actual}^{90.4})$. The reason for using an absolute measure of price increases rather than a measure relative to the pre-increase price $\left(e.g. \frac{[p^{91.1} - p^{90.4}]}{p^{90.4}} \right)$ is that the excise tax is based on quantity sold and hence tax pass-through should be unrelated to the pre-increase price $(p_j^{90.4})$.²

The estimated demand parameters of model 2 in table 4 together with the pre-tax increase marginal costs and the tax increase are used to compute the predicted prices $(p_j^{91.1})$ that would have prevailed under each model in the post-tax increase period (see details in section 6.3).

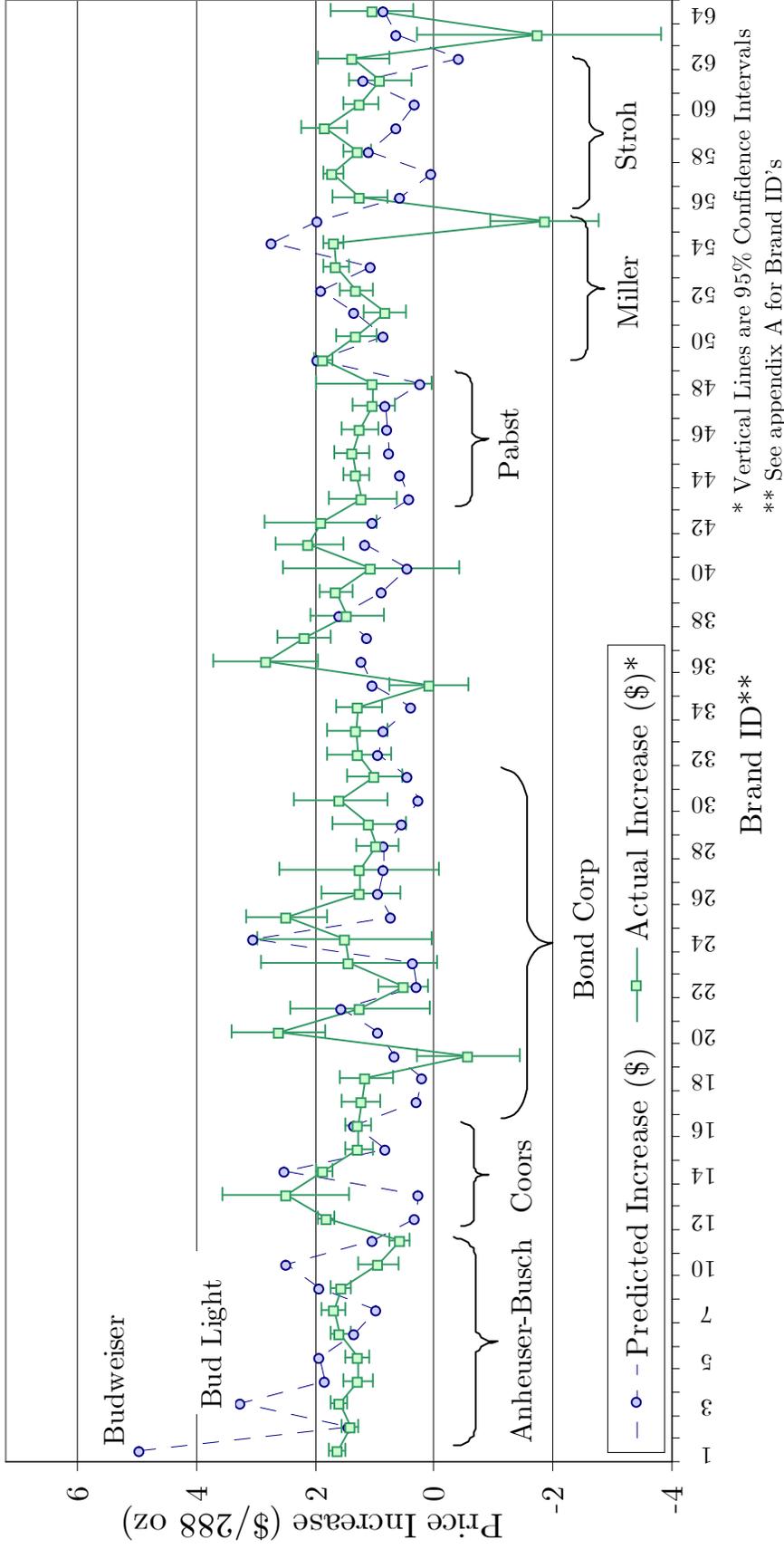
The next figures plot the **mean** of predicted prices (one figure for each model) and the **mean** of actual price increases for each brand. The mean of the actual price increases is computed according to details presented in section 6.4. The mean of predicted price increases is calculated by averaging each brand's predicted increases over the 46 cities. Each figure corresponds to one of the scenarios discussed in the previous section. 95% confidence intervals are displayed for the means of actual price increases.³ It should be noted that many brands have tight 95% confidence intervals around actual mean increases (around 15¢ and 20¢), indicating that actual price increases do not vary substantially across cities. This pattern can particularly be observed for brewers that tend to produce nationally: Anheuser-Busch, Coors, Pabst, Miller and Stroh. Table 7.2 displays statistics of actual increases for a sample of brands. Corona, Molson and Old Style show greater variation

²Price increase with ad-valorem taxes, on the other hand, are directly related to price.

³The non-linear systems for predicted price increases require 12 hours of computing time. Calculating confidence intervals for the mean of predicted price increases with a bootstrapping technique are hence extremely costly even with a modest number of draws.

in the price increase than beers produced by brewers with more national presence: Budweiser, Budweiser Light (Anheuser-Busch), Coors Light (Coors), Miller Lite (Miller) and Schaefer (Stroh).

Figure 7.1. Predicted Price Increases by Bertrand-Nash behavior vs. Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)



* Vertical Lines are 95% Confidence Intervals
 ** See appendix A for Brand ID's

Figure 7.2. Predicted Price Increases by Leadership of Anheuser-Busch and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities).

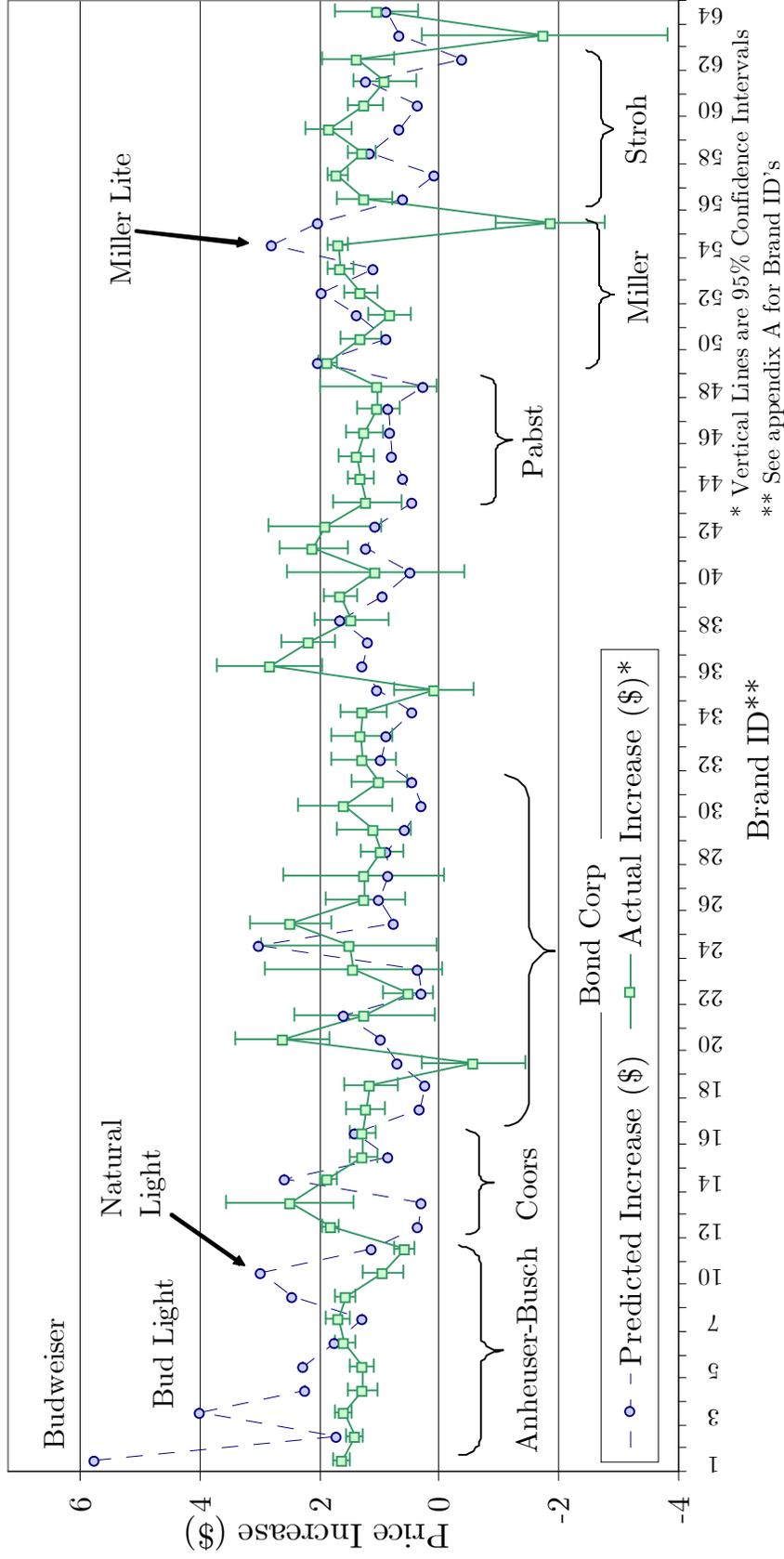


Figure 7.3. Predicted Price Increases by Leadership of Budweiser and Actual Price Increases and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)

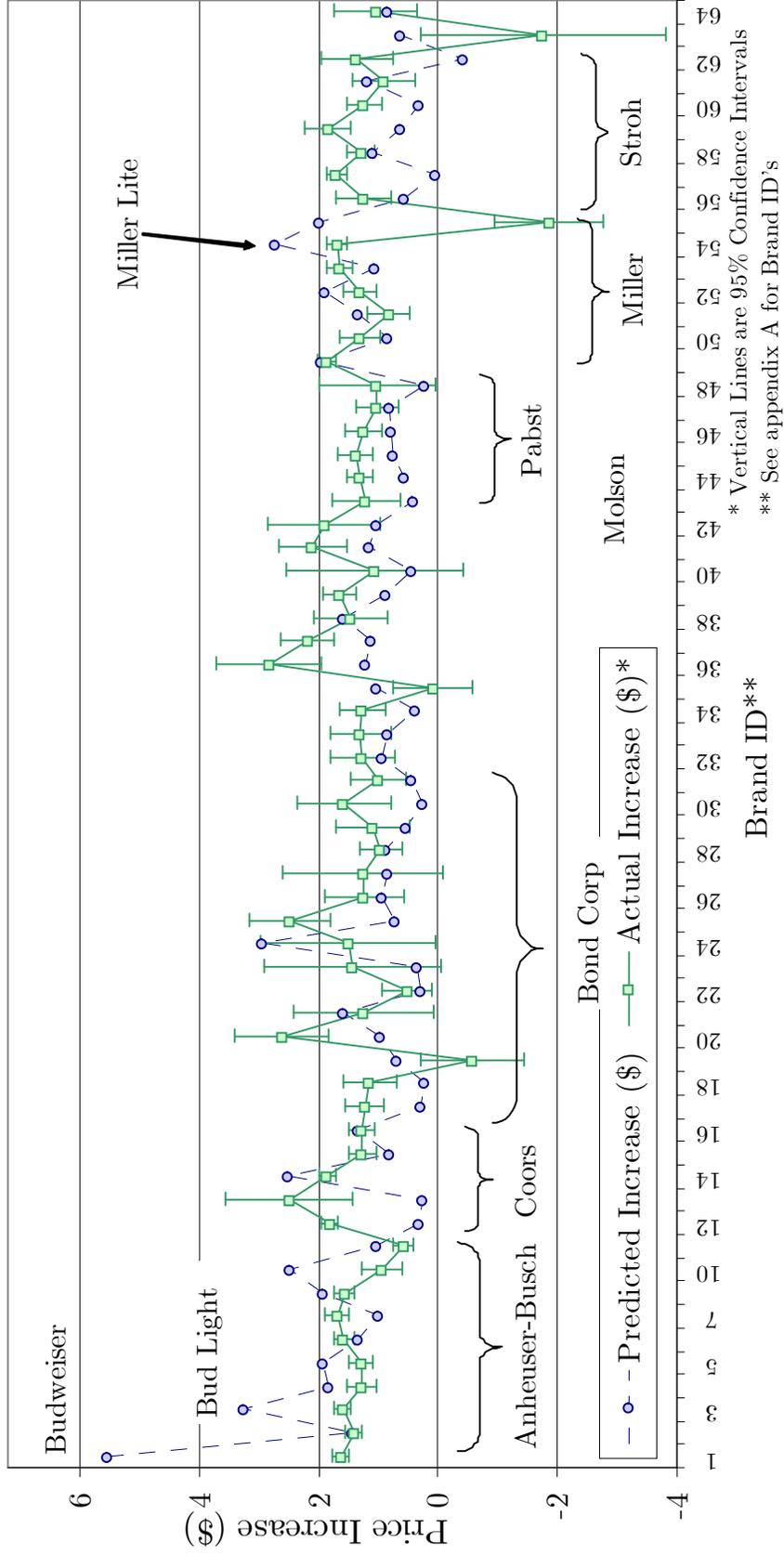


Figure 7.4. Predicted Price Increases of Collusive Price Leadership by Budweiser and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)

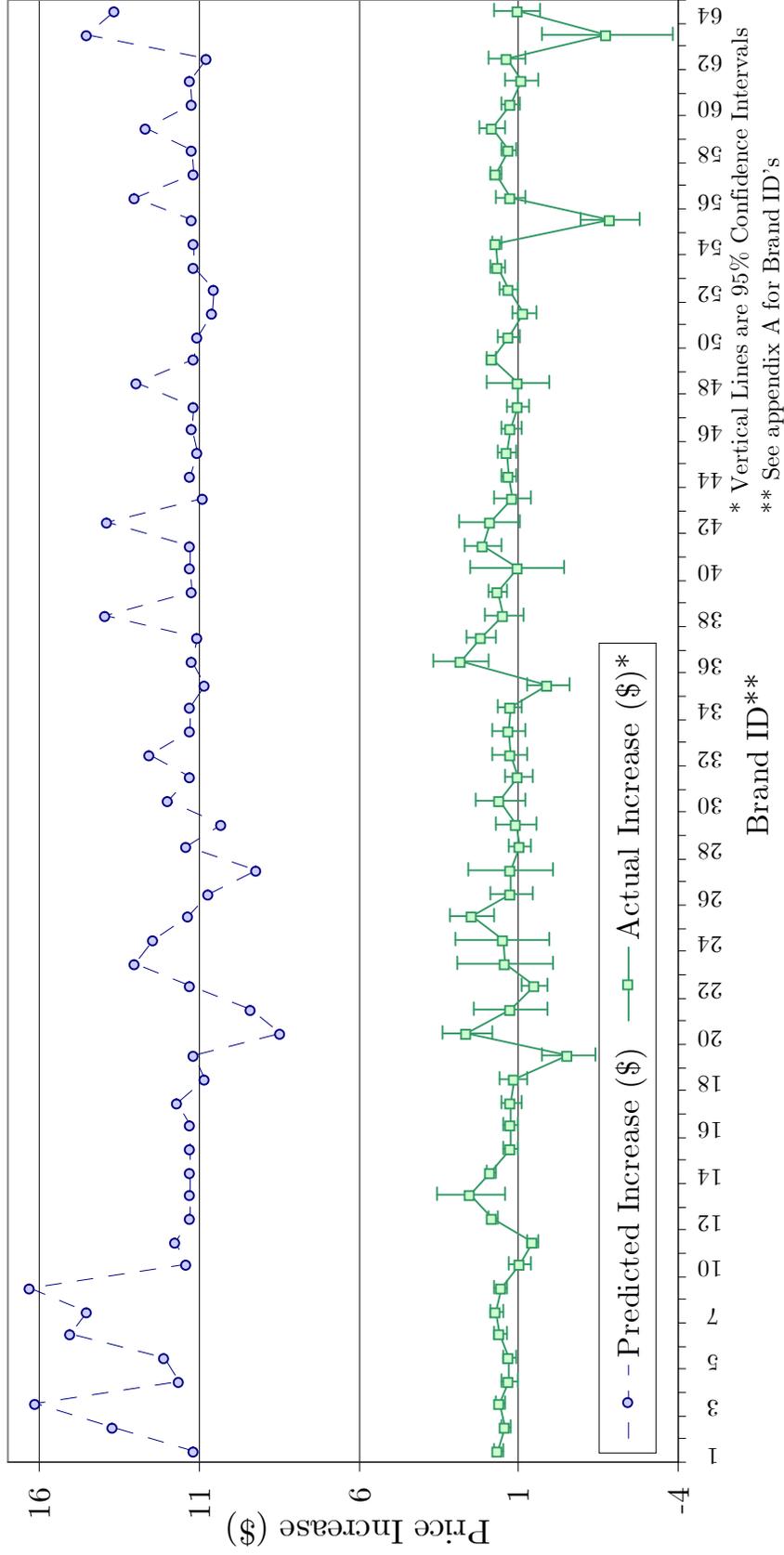


Figure 7.5. Predicted Price Increases by Collusive behavior between 3 largest firms and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)

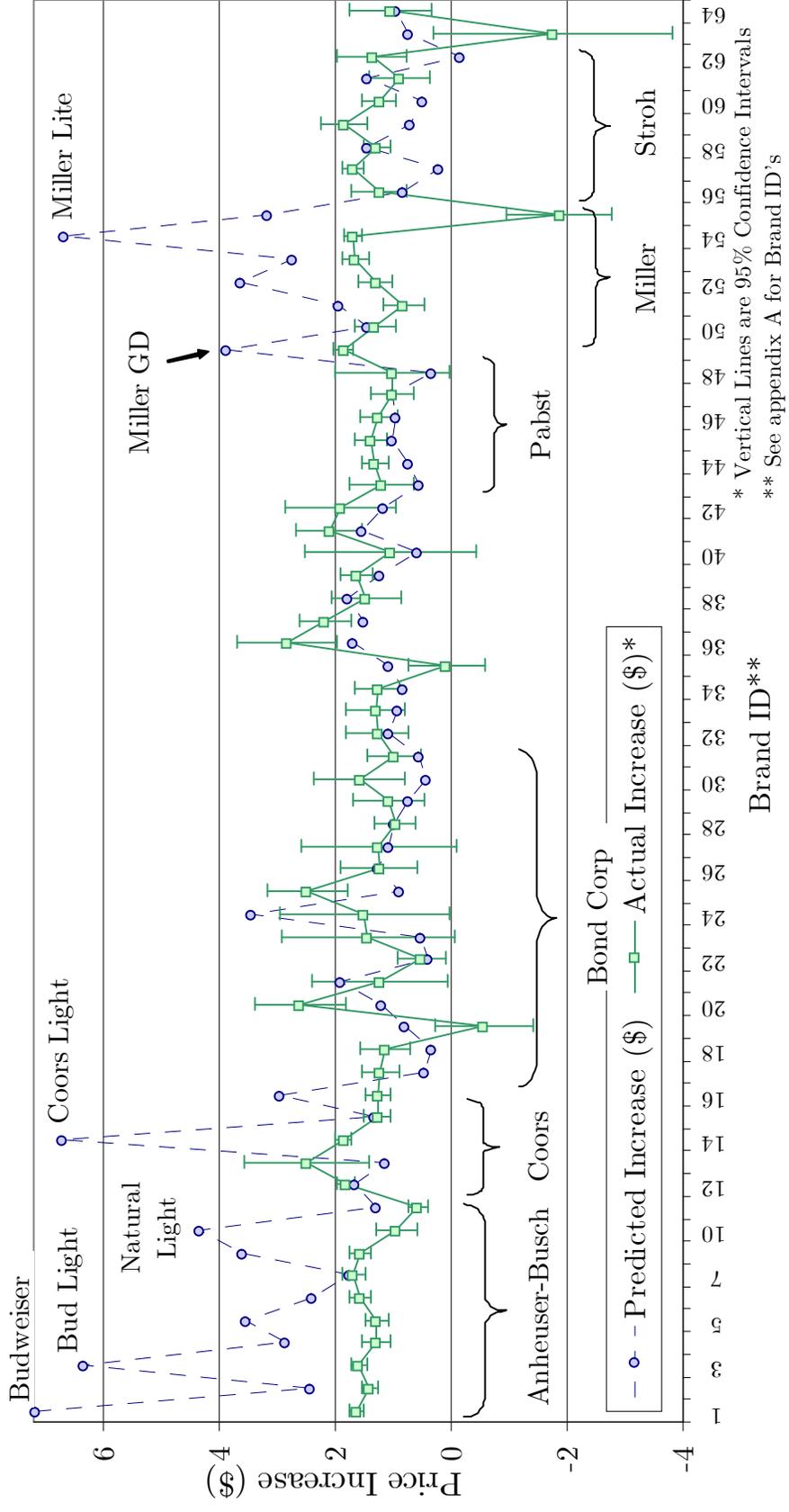


Figure 7.6. Predicted Price Increases by Collusive behavior between leading brands of 3 largest firms and Actual Price Increases per brand after 100% Hike in the Federal Excise Tax (Mean over 46 cities)

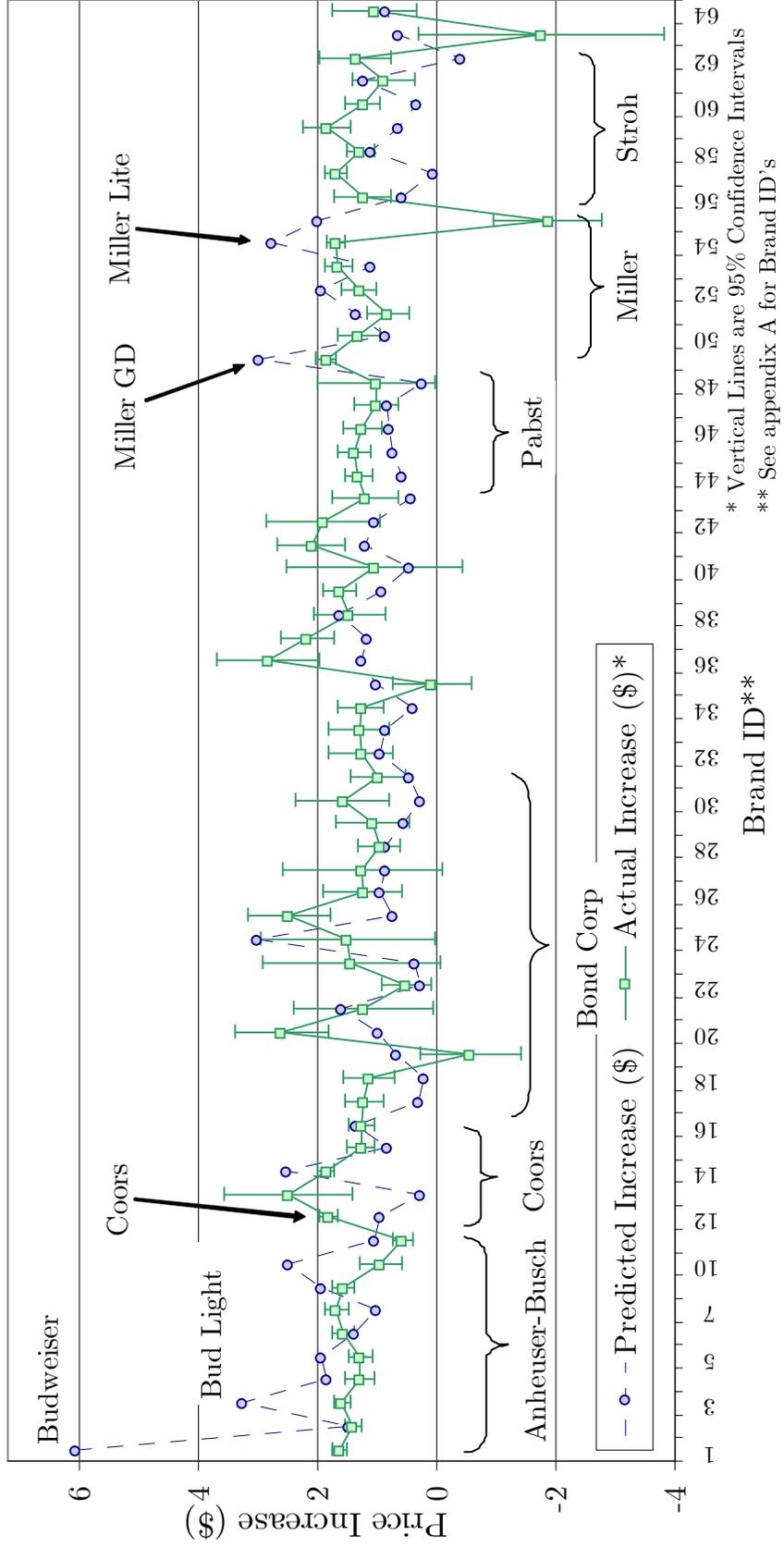


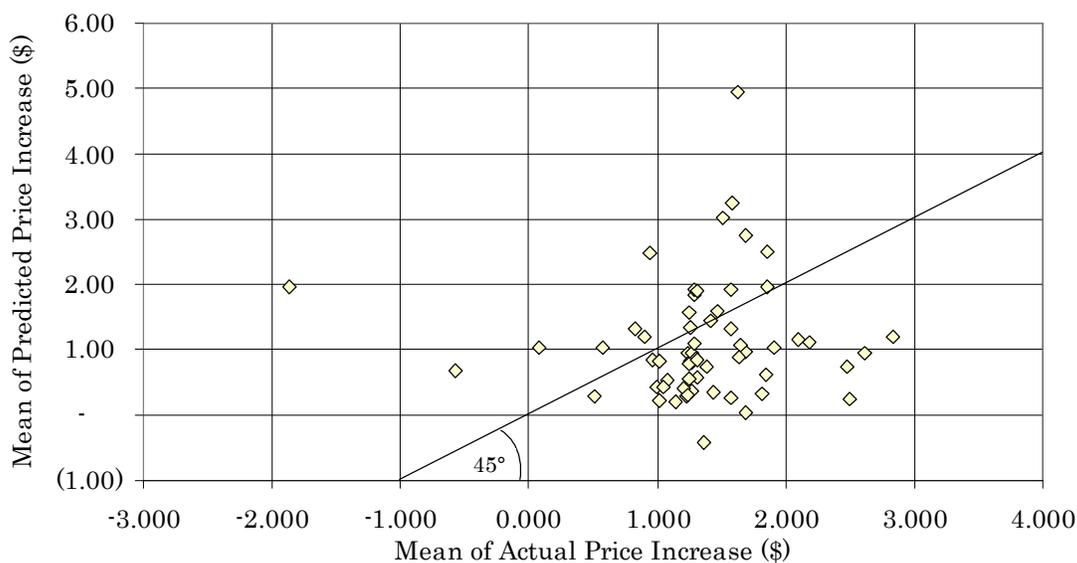
Table 7.2. Summary Statistics of Actual Price Increases for Selected Brands

Brand	Mean	St. Dev	Min (city)	Max (city)
Budweiser	1.37	0.31	0.87 (Seattle-Tacoma)	2.60 (Hartford-Springfield)
Bud Light	1.39	0.30	0.86 (Seattle-Tacoma)	2.58 (Hartford-Springfield)
Coors Light	1.52	0.30	0.72 (St. Louis)	2.34 (Detroit)
Old Style	1.35	0.53	0.50 (Phoenix-Tucson)	2.91 (Birmingham)
Corona	1.37	1.20	-0.95 (Jacksonville)	3.91 (San Francisco/Oakland)
Molson	1.05	0.68	0.39 (Grand Rapids)	1.88 (Albany)
Miller Lite	1.40	0.30	0.36 (Seattle/Tacoma)	2.12 (Albany)
Shaefer	1.28	0.32	0.41 (Knoxville)	1.92 (Seattle/Tacoma)

While the Bertrand-Nash model (figure 7.1) appears to be a closer predictor of actual firm behavior, there are several patterns in the data. Bertrand-Nash behavior tends to under-predict price increases: 41 out of 63 are “under-predicted” brands. This can be seen in figure 7.7, which plots mean predicted price increases (x-axis) against mean actual price increases (y-axis): most brands are under the 45° degree line. Also, over-predicted brands appear to be more frequent among the two largest beer producers: Anheuser-Busch (8 out of 10) and Miller (4 out of 7). Since, in general, more inelastic brands are associated with higher tax pass-through rates, the two largest mean predicted increases correspond to the first and third most inelastic brands in the sample: Budweiser (predicted: \$4.94, actual \$1.63) and Bud Light (predicted: \$3.24, actual \$1.59).

The Stackelberg model in which Anheuser-Busch acts as the price leader with all its brands (figure 7.2) has a pattern similar to the Bertrand-Nash case. Although it can not be discerned from the figure, in this model predicted increases are higher than in the Bertrand-Nash case for all but one brand. For Anheuser-Busch’s brands, especially Budweiser, Bud Light and Natural Light, this difference is larger and hence discernible from the figures.

Figure 7.7. Mean of Actual Price Increases vs. Mean of Predicted Price Increases, Bertrand-Nash Competition



Aside from a larger over-prediction for Budweiser (an additional 60 ¢), the Budweiser Stackelberg model (figure 7.3), yields predicted mean price increases that are essentially the same to the Bertrand-Nash case. Overall, the two Stackelberg models considered do not differ substantially from Bertrand-Nash behavior. The reason is that, because reaction functions of followers depend mainly on very small cross-price coefficients, the term $\frac{dp_m}{dp_j}$ in (6.2) takes small positive values making the first order conditions of the leader not substantially different from the Bertrand-Nash case (for which the term $\frac{dp_m}{dp_j}$ is equal to zero).

To illustrate, consider the first order condition of Budweiser in Albany in the first quarter of 1988. For the case in which Budweiser is a Stackelberg leader, the first order condition is given by:

$$(7.1) \quad q_{Bud} - 1 + \sum_{k \in AB} (p_k - c_k) \left[\frac{\partial q_k}{\partial p_{Bud}} + \sum_{m \notin F_n} \frac{\partial q_k}{\partial p_m} \frac{dp_m}{dp_{Bud}} + \frac{\partial q_k}{\partial A_{Bud}} \right] = 0,$$

where AB denotes the product set of Anheuser-Busch. For the Bertrand-Nash case, the term $\frac{dp_m}{dp_{Bud}}$ takes a value of zero, so the first order condition is:

$$(7.2) \quad q_{Bud} - 1 + \sum_{k \in AB} (p_k - c_k) \left[\frac{\partial q_k}{\partial p_{Bud}} + \frac{\partial q_k}{\partial A_{Bud}} \right] = 0,$$

There are 5 of Anheuser-Busch brands present in this city-quarter: $AB = \{\text{Budweiser, Bud Light, Busch, Michelob and Michelob Light}\}$. The value of the bracketed expression in (7.1) for each of these $k \in AB$ are: -1613, 9.67, 14.14, 10.49 and 8.37, respectively. Conversely, for the Bertrand-Nash case, (7.2), the values are -1615, 7.60, 11.30, 8.47 and 6.42, respectively. Although there are important differences in the 4 positive terms between the two models, when terms are added, the difference between (7.1) and (7.2) is minimal. The reason is that the own-price derivative $\frac{\partial q_{Bud}}{\partial p_{Bud}}$ is three orders of magnitude larger than the cross-price derivatives. Therefore, the first-order condition is mostly determined by the bracketed expression for $k = \text{Budweiser}$.

Collusive price leadership by Budweiser (figure 7.4) predicts unlikely predicted price increases that are, on average, almost 7 times larger than actual mean price increases, with some of Anheuser-Busch's brands in the vicinity of \$15-\$16. This extreme case can therefore be rejected.

The 3-firm collusion scenario (figure 7.5) over-predicts the price increases of the best selling brands of the colluding firms (e.g. Budweiser, Bud Light, Coors Light, Miller Genuine Draft and Miller Lite) by a large amount. The 3-brand collusion scenario (figure 7.6)

differs less strikingly with Bertrand-Nash: there is a higher over-prediction for Budweiser and, less noticeably, for the other two colluding brands: Coors and Miller Genuine Draft.

The left part of table 7.3 presents summary statistics of price increases (i.e. the difference in prices between the two quarters). The mean of predicted increases between Bertrand-Nash, the two Stackelberg models and collusion among three brands are similar and also close to the mean of actual increases, suggesting that these are superior models. Although the Anheuser-Busch Stackelberg model predicts the mean of actual increases more accurately than Bertrand-Nash, closer inspection of graph 7.2 indicates that this is due mainly to larger over-predictions for Anheuser-Busch's brands, and not by smaller under-predictions of other brands. A similar argument can be made for the model of collusion among three brands.

Medians and standard deviations indicate that predicted price increases are more heterogeneous in the models than suggested by actual increases. While the median of actual increases is similar to its mean, means of predicted increases are higher than their respective medians (except in collusive price leadership) as a consequence of the outliers given by over-predicted brands (e.g. Budweiser, Bud Light). Compared to actual price increases standard deviations of predicted increases are larger, which corroborates this heterogeneity.

One metric of assessing the different models is the number of brands whose predicted mean price increases fall within the confidence intervals of actual mean price increases shown in the graphs. The right part of table 7.3 presents this number (# Non-Rejections) for the models considered. According to this metric, collusion among 3 firms explains firm behavior better than the other models.

Two more rigorous metrics are considered. The first is the market share-weighted price increase. With this metric, accuracy in prediction is more important for more popular brands. Using this criterion, the Bertrand-Nash outperforms the other models (see table 7.3), though it is more than twice the value for weighted actual increases (3.13 vs. 1.33). The large difference between the market share-weighted predicted increases and the market share-weighted actual increases is due to the over-prediction of more popular brands: the combined share of Budweiser (19%), Bud Light (6%), Coors Light (7%) and Miller Lite (9%) is 41%. In all models, there is an over-prediction for these brands, the largest of which is for Budweiser. The similarity between the weighted actual increases and its non-weighted counterpart indicates homogeneous actual price increases across brands. The second metric is the sum of squared deviations, where a deviation is defined as the difference between the predicted increase and the actual increase.⁴ This criterion also suggests that Bertrand-Nash is a more accurate model for explaining firms' actual pricing behavior.

⁴The same conclusion is reached if each deviation is weighted by market shares.

Table 7.3. Summary Statistics of Actual and Predicted Price Increases and Performance Metrics of Models

	Summary Statistics ^a			Performance Metrics		
	Mean	Median	St. Dev	# No- Reject ^b	Weighted Mean ^c	SSD ^d
<i>Actual Increases</i>	1.38	1.37	0.65	N/A	1.33	N/A
<i>Predicted Increases:</i>						
Bertrand-Nash	1.24	0.97	1.41	20	3.13	3993
A-B Stackelberg Leader ^e	1.39	1.04	1.54	20	3.44	4663
Budweiser Stackelberg Leader	1.26	0.98	1.46	21	3.26	4232
Collusive Leadership (Bud)	11.83	11.03	2.93	0	11.44	20629
Collusion 3 firms ^f	2.21	1.50	2.40	22	5.20	11671
Collusion 3 brands ^g	1.34	1.03	1.52	19	3.52	4545

^a Computed with absolute price increases for each brand: the absolute price difference between the first quarter of 1991 and the fourth quarter of 1990, over 46 cities (1748 observations)

^b Number of brands for which the mean of predicted increases falls within the confidence intervals of the mean of actual increases (graphs 7.1 to 7.6)

^c Weighted average of absolute price increases (weight=volume of brand sold in city/ total volume of all brands in all cities in the first quarter of 1991)

^d Sum of squared deviations over all brands and all cities; deviation=predicted-actual (1748 obs.)

^e A-B = Anheuser-Busch.

^f Anheuser-Busch, Adolph Coors, Miller (Philip Morris).

^g Budweiser, Coors, Miller Genuine Draft

CHAPTER 8

Conclusion

This dissertation analyzes pricing behavior in the U.S. brewing industry where there is some evidence of price leadership by Anheuser-Busch and its Budweiser brand. Bertrand-Nash, leadership and collusive models are considered as possible candidates of pricing behavior. To choose the model that is best supported by the data, an exogenous variation in the data, namely the 100% increase in the federal excise tax, is used to compare actual price increases with price increases predicted by the different models of pricing behavior.

Using several metrics of closeness between predicted price increases and actual price increases revealed that, from the models considered, Bertrand-Nash appears to be more consistent with the actual behavior of firms. However, because reaction of followers to the leader's price changes are almost negligible, Stackelberg models predict prices that are not substantially different from competitive behavior.

Overall, actual price increases tend to be more similar across brands than any of the models predict. In particular, the data fails to support the inverse relationship between own-price elasticity and excise tax pass-through rates predicted by all the profit-maximizing models considered. As a consequence, Bertrand-Nash (and other models) tend to over-predict tax pass-through rates of more price-inelastic brands, especially Budweiser, and to under-predict price increases of more price-elastic brands. While dynamics have been ignored in this dissertation (because data available was aggregated by quarters), it is unclear whether a dynamic setting can yield better predictions of firms' behavior. Because

tax pass-through rates are driven by elasticities, further robustness checks with alternative demand estimation models, such as the random coefficients logit, may be conducted as an extension to this work.

An interpretation of this evidence in a static setting is that, although price-cost margins are large in this industry, Anheuser-Busch could exert larger market power than it actually does even under competitive Bertrand-Nash behavior. One policy implication of this interpretation is that antitrust concern should be low in terms of the use of market power by the leading beer producer to obtain larger price-cost margins. The results in this dissertation are consistent with the findings of Nevo (2001) in the ready-to-eat breakfast cereals who suggests that large price-cost margins in Bertrand-Nash competition can be the result of product differentiation and the portfolio effect of firms carrying more than one brand rather than the results of actual non-competitive behavior.

The models considered here, as most models of pricing behavior, are built upon the assumption of profit maximization. However, the results of this research may be consistent with simpler, yet plausible pricing strategies. For instance, the fact that actual price increases for large brewers' brands (Anheuser-Busch, Coors, Miller) as a result of the tax increase have minimal variation across cities and are similar to actual price increases of smaller brewers' brands may be interpreted as leading brewers setting a common cost mark-up for all brands, regardless of where they are sold (and possibly of how elastic they are), and smaller brewers matching these mark-ups. This conjecture is strengthened by the fact that price increases for elastic brands, which are produced mainly by smaller brewers and are generally more limited in their ability to increase prices, appear much higher than what Bertrand-Nash and other models suggest, with values close to actual price increases of the more inelastic brands produced by larger brewers. While this conjecture

is consistent with the informal observations in the industry, the models considered are not able to capture it. Moreover, there is no clear way to test it within the profit maximization framework used in this paper. Deriving price changes from models based on objectives other than profit maximization remains as an area for future research.

Scherer and Ross (p. 261-265) explain that a type of “rule-of-thumb” pricing of the sort suggested above is common in many industries and is used as a way to cope with “uncertainties in the estimation of demand function shapes and elasticities” (p. 262). Furthermore, this type of pricing behavior can be used as a coordinating device, especially when there are changes in costs and firms in the industry share similar production technologies. To the extent that this type of behavior is facilitating coordination in the brewing industry, it appears that, as a consequence of the tax increase, it benefited smaller brewers rather than large firms.

The availability of more detailed data would allow to capture aspects that are not addressed in this dissertation. Dynamic adjustment to the observed prices can be better assessed with less aggregated data on the time dimension. Detailed cost data at the manufacturer and retailer level can allow to extend the analysis to manufacturer-retailer aspects and also more rigorous econometric tests of the competing pricing models considered in this work.

References

- [1] Anderson, S., A. de Palma and B. Kreider, 2001, "Tax Incidence in Differentiated Product Oligopoly" *Journal of Public Economics*, 81, 173-192.
- [2] *Anheuser-Busch 2003 Annual report*, available at <http://www.anheuser-busch.com/annual/2003/Domestic.pdf>
- [3] Baker, J. and T. Bresnahan (1985): "The Gains from Merger or Collusion in Product-Differentiated Industries," *Journal of Industrial Economics*, 33, 427-444.
- [4] Beer Intsitute, available at: <http://www.beerintstitute.org/>
- [5] Berry, S. (1994): "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25, 242-262.
- [6] Berry, S., J. Levinsohn and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841-890.
- [7] Bresnahan, T. (1989): "Empirical Methods for Industries with Market Power," in *Handbook of Industrial Organization*, Vol. II, ed. by R. Schmalensee and R. Willig, Amsterdam: North-Holland.
- [8] Brewers Almanac, various issues.
- [9] Bureau of Labor Statistics, Current Population Survey Basic Monthly Files, Various issues, available at the National Bureau of Economic Research: http://www.nber.org/data/cps_index.html
- [10] Case, G., A. Distefano and B. K. Logan (2000): "Tabulation of Alcohol Content on Beer and Malt Beverages," *Journal of Analytical Toxicology*, 24, 202-210.
- [11] Center for Science in the Public Interest (1983): "Omnibus petition for regulation of unfair and deceptive alcoholic beverage advertising and marketing practices," Docket No. 209-49 before the Federal Trade Commission, Center for Science in the Public Interest, Washington, DC.

- [12] Christensen, R., D. Jorgenson and L. Lau (1975): "Trascendental Logarithmic Utility Functions," *American Economic Review*, 65, 367-383.
- [13] Corts, K.S. (1999): "Conduct Parameters and the Measurement of Market Power," *Journal of Econometrics*, 88, 227-250.
- [14] Deaton, A. and J. Muellbauer (1980): "An Almost Ideal Demand System," *American Economic Review*, 70, 312-326.
- [15] Demographia. www.demographia.com
- [16] Dhar, T., J. Chavas and B. Gould (2003): "An Empirical Assessment of Endogeneity Issues in Demand Analysis of Differentiated Products," *American Journal of Agricultural Economics*, 85, 605-617.
- [17] Elzinga, K. (1973): "The Restructuring of the U.S. Brewing Industry," *Industrial Organization Review*, 1, 104-114.
- [18] ——— (2000): "The Beer Industry," in *The Structure of American Industry*, ed. by W. Adams, New York: Macmillan.
- [19] Gasmi, F., J. J. Laffont and Q. Vuong (1992): "Econometric Analysis of Collusive Behavior in a Soft-Drink Market," *Journal of Economics and Management Strategy*, 1, 277-311.
- [20] Goldberg, P. and F. Verboven (2001): "The Evolution of Price Dispersion in European Car Markets," *Review of Economic Studies*, 811-848
- [21] Greer, D. (1998): "Beer: Causes of Structural Change," in *Industry Studies*, ed. by L. Duetsch, NJ: Prentice Hall.
- [22] Hausman, J., and G. Leonard (2002): "The Competitive Effects of a New Product Introduction: A Case Study," *Journal of Industrial Economics*, 50, 237-263.
- [23] Hausman, J., G. Leonard and D. Zona (1994): "Competitive Analysis with Differentiated Products," *Annales d'Économie et de Statistique*, 34, 159-180.
- [24] Internal revenue code available at: <http://www.fourmilab.ch/ustax/ustax.html> (accessed 5/17/05).
- [25] Kadiyali, V., N. Vilcassim and P. Chintagunta (1996): "Empirical Analysis of Competitive Product Line Pricing Decisions: Lead, Follow, or Move Together?," *Journal of Business*, 69, 459-487.
- [26] Langan, G. (1997): "Brand Level Demand and Oligopolistic Interaction Among Domestic and Foreign Beer Brands," Ph.D. Dissertation, University of Connecticut.

- [27] Lee, B. and V. Tremblay (1992): "Advertising and the U.S. Market Demand for Beer," *Applied Economics*, 24, 69-76.
- [28] McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, ed. by P. Zarembka, New York: Academic Press.
- [29] ——— (1978): "Modeling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, ed. by A. Karlqvist et. al., Amsterdam: North Holland.
- [30] ——— (1981): "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data*, ed. by C. Manski and D. McFadden, Cambridge: MIT Press, 198-272.
- [31] ——— (1984): "Econometric Analysis of Qualitative Response Models," in *Handbook of Econometrics*, Vol III, ed. by Z. Griliches and M. Intrilligator, Amsterdam: North-Holland.
- [32] McFadden, D. and K. Train (2000): "Mixed MNL Models of Discrete Response," *Journal of Applied Econometrics*, 15, 447-470.
- [33] Moschini, G. (1995): "Units of Measurement and the Stone Index in Demand System Estimation," *American Journal of Agricultural Economics*, 63-68.
- [34] Modern Brewery Age, Blue Book, various issues.
- [35] Nelson, J. (1999): "Broadcast Advertising and U.S. Demand for Alcoholic Beverages," *Southern Economic Journal*, 65, 774-784.
- [36] Nelson, J. and J. Moran (1995): "Advertising and U.S. Alcohol Demand: System-wide estimates," *Applied Economics*, 27, 1225-36.
- [37] Nevo, A. (1998): "Identification of the Oligopoly Solution Concept in a Differentiated Products Industry," *Economics Letters*, 59(3), 391-395.
- [38] ——— (2000a): "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 9, 513-548.
- [39] ——— (2000b): "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, 31, 395-421.
- [40] ——— (2001): "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69, 307-342.
- [41] Omnibus Act available at: <http://thomas.loc.gov/> (accessed 5/17/05)

- [42] Pagan, A. and A. Ullah (1999): *Nonparametric Econometrics*, Cambridge: Cambridge University Press.
- [43] Pinkse, J., M. Slade and C. Brett (2002): "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, 70, 1111-1155.
- [44] Pinkse, J. and M. Slade (2004): "Mergers, Brand Competition, and the Price of a Pint," *European Economic Review*, 48.
- [45] Reiss, P. and F. Wolak (2005): "Structural Econometric Modeling: Rationales and Examples from Industrial Organization" in *Handbook of Econometrics*, forthcoming.
- [46] Rotemberg, J. and G. Saloner, 1990, "Collusive Price Leadership," *Journal of Industrial Economics*, 39, 1, 93-111.
- [47] Scherer, F. and D. Ross (1990): *Industrial Market Structure and Economic Performance*, Houghton Mifflin: Boston.
- [48] Slade, M. (1995): "Product Rivalry with Multiple Strategic Weapons: An Analysis of Price and Advertising Competition," *Journal of Economics and Management Strategy*, 4, 445-476.
- [49] ——— (2004): "Market Power and Joint Dominance in UK Brewing," *Journal of Industrial Economics*, 52, 133-163.
- [50] Sutton, J. (1991): *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration*, Cambridge: MIT Press.
- [51] Tremblay, V. and C. Tremblay (2005): *The U.S. Brewing Industry: Data and Economic Analysis*, MIT Press: Cambridge.
- [52] Villas-Boas, S. B. (2005): "Vertical Contracts Between Manufacturers and Retailers: Inference with Limited Data," *Working Paper*, University of California, Berkeley.
- [53] U.S. Census Bureau, *Concentration Ratios in Manufacturing*, 1997 Economic Census. <http://www.census.gov/prod/ec97/m31s-cr.pdf>

APPENDIX A

Derivation of Elasticity Formulas

A.1. Price Elasticities

Differentiating (2.6) with respect to own- and cross-prices (p_{jt} and p_{kt}) gives:

$$(A.1) \quad \frac{\partial w_{jt}}{\partial p_{kt}} = \begin{cases} \frac{\partial(p_{jt}q_{jt}/x_t^*)}{\partial p_{jt}} = \frac{q_{jt}}{x_t^*} + \frac{p_{jt}}{x_t^*} \frac{\partial q_{jt}}{\partial p_{jt}} = \frac{q_{jt}}{x_t^*} + \frac{p_{jt}}{x_t^*} \frac{\partial q_{jt}}{\partial p_{jt}} \frac{q_{jt}}{q_{jt}} = \frac{q_{jt}}{x_t^*} + \frac{q_{jt}}{x_t^*} \eta_{jj}, & \text{for } j = k \\ \frac{\partial(p_{jt}q_{jt}/x_t^*)}{\partial p_{kt}} = \frac{p_{jt}}{x_t^*} \frac{\partial q_{jt}}{\partial p_{kt}} = \frac{p_{jt}}{x_t^*} \frac{\partial q_{jt}}{\partial p_{kt}} \frac{p_{kt}}{q_{jt}} \frac{q_{jt}}{p_{kt}} = \frac{w_{jt}}{p_{kt}} \eta_{jk}, & \text{for } j \neq k \end{cases}$$

Solving for η_{jj} and η_{jk} in (A.1) gives:

$$(A.2) \quad \eta_{jk} = \begin{cases} -1 + \frac{x_t^*}{q_{jt}} \frac{\partial w_{jt}}{\partial p_{jt}}, & \text{for } j = k \\ \frac{\partial w_{jt}}{\partial p_{kt}} \frac{p_{kt}}{w_{jt}}, & \text{for } j \neq k \end{cases}$$

where the own- and cross-price derivatives, $\frac{\partial w_{jt}}{\partial p_{jt}}$ and $\frac{\partial w_{jt}}{\partial p_{kt}}$, are:

$$(A.3) \quad \frac{\partial w_{jt}}{\partial p_{kt}} = \begin{cases} \frac{\partial w_{jt}}{\partial \log p_{jt}} \frac{\partial \log p_{jt}}{\partial p_{jt}} = \frac{1}{p_{jt}} \frac{\partial w_{jt}}{\partial \log p_{jt}} = \frac{1}{p_{jt}} \left[b_{jj} - d_j \frac{\partial \log P_t^L}{\partial \log p_{jt}} \right], & \text{for } j = k \\ \frac{\partial w_{jt}}{\partial \log p_{kt}} \frac{\partial \log p_{kt}}{\partial p_{kt}} = \frac{1}{p_{kt}} \frac{\partial w_{jt}}{\partial \log p_{kt}} = \frac{1}{p_{kt}} \left[b_{jk} - d_j \frac{\partial \log P_t^L}{\partial \log p_{kt}} \right], & \text{for } j \neq k \end{cases}$$

Price elasticities are obtained by replacing (A.3) into (A.2):

$$(A.4) \quad \eta_{jk} = \begin{cases} -1 + \frac{1}{w_{jt}} \left[b_{jj} - d_j \frac{\partial \log P_t^L}{\partial \log p_{jt}} \right], & \text{for } j = k \\ \frac{1}{w_{jt}} \left[b_{jk} - d_j \frac{\partial \log P_t^L}{\partial \log p_{kt}} \right], & \text{for } j \neq k \end{cases}$$

In this application, the term d_j is not statistically different from zero. Hence, for computational purposes equation (A.4) is reduced to:

$$(A.5) \quad \eta_{jk} = \begin{cases} -1 + \frac{b_{jj}}{w_{jt}}, & \text{for } j = k \\ \frac{b_{jk}}{w_{jt}}, & \text{for } j \neq k \end{cases}$$

A.2. Advertising Elasticities

Advertising elasticities are estimated in as similar way:

$$(A.6) \quad \frac{\partial w_{jt}}{\partial A_{kt}} = \begin{cases} \frac{\partial(p_{jt}q_{jt}/x_t^*)}{\partial A_{jt}} = \frac{p_{jt}}{x_t^*} \frac{\partial q_{jt}}{\partial A_{jt}} = \frac{p_{jt}}{x_t^*} \frac{\partial q_{jt}}{\partial A_{jt}} \frac{A_{jt}}{q_{jt}} \frac{q_{jt}}{A_{jt}} = \frac{w_{jt}}{A_{jt}} \rho_{jj}, & \text{for } j = k \\ \frac{\partial(p_{jt}q_{jt}/x_t^*)}{\partial A_{kt}} = \frac{p_{jt}}{x_t^*} \frac{\partial q_{jt}}{\partial A_{kt}} = \frac{p_{jt}}{x_t^*} \frac{\partial q_{jt}}{\partial A_{kt}} \frac{A_{kt}}{q_{jt}} \frac{q_{jt}}{A_{kt}} = \frac{w_{jt}}{A_{kt}} \rho_{jk}, & \text{for } j \neq k \end{cases}$$

Solving for the own- and cross-advertising terms (ρ_{jj} and ρ_{jk}):

$$(A.7) \quad \rho_{jk} = \begin{cases} \frac{\partial w_{jt}}{\partial A_{jt}} \frac{A_{jt}}{w_{jt}}, & \text{for } j = k \\ \frac{\partial w_{jt}}{\partial A_{kt}} \frac{A_{kt}}{w_{jt}}, & \text{for } j \neq k \end{cases}$$

Now, from (2.6), the own- and cross-advertising derivatives ($\frac{\partial w_{jt}}{\partial A_{jt}}$ and $\frac{\partial w_{jt}}{\partial A_{kt}}$), are:

$$(A.8) \quad \frac{\partial w_{jt}}{\partial A_{kt}} = \begin{cases} \gamma c_{jj} A_{jt}^{\gamma-1}, & \text{for } j = k \\ \gamma c_{jk} A_{kt}^{\gamma-1}, & \text{for } j \neq k \end{cases}$$

Finally, substituting (A.8) into (A.7) gives the advertising elasticities:

$$(A.9) \quad \rho_{jk} = \begin{cases} \frac{\gamma^{c_{jj}A_{jt}^\gamma}}{w_{jt}}, & \text{for } j = k \\ \frac{\gamma^{c_{jk}A_{kt}^\gamma}}{w_{jt}}, & \text{for } j \neq k \end{cases}$$

APPENDIX B

Data and Programming Details

The two programs used in this dissertation are Stata 8.0 and Matlab release 12.1. Matlab is used to program routines for several requirements of the model. Stata is used for most of the statistical analysis and also for reading and preparing the data set. This appendix presents details on data preparation and programming of routines.

B.1. Demand

Data

IRI data is in ASCII format and comes in 20 files, one for each quarter. A dictionary that assigns each column of the file to a variable is created. This dictionary is then used to read the data into Stata. After the data selection procedure (see section 3.1) the data is transformed to a balanced format by creating additional observations so as to ensure the presence of 64 observations per city-quarter. These additional observations have missing values and are just created to simplify the computation of distance measures and common boundaries in Matlab.

Weighing Matrices

To illustrate how weighing matrices are created, consider the discrete distance measure of whether two brands are produced by the same brewer. The weighing matrix *WBREW* is created with the following program in Matlab:

```
i=1;
```

```

T=59392; % Number of observations in balanced format
N=928; % this is the number of markets (city-quarters)
B=64; % this is the number of brands being analyzed

%Brewer Type weighing matrix
while i<T+1
    WBREW(i:B+i-1,:)=(brew(i:B+i-1,:)*(brew(i:B+i-1,:)'));
    i=i+B;
end

```

where, “brew” is a matrix containing 13 dummy variables, one for each brewer. Normalization is done with:

```

A=sum(WBREW,2)-1;
for i=1:T
    if A(i,1)~=0
        WBREW(i,:)=(1/(A(i,:))*WBREW(i,:));
    end
end

```

For continuous measures, the weighing matrix for the distance in alcohol content is calculated with:

```

i=1;
WALC=zeros(T,B); % this is the matrix in which square BxB weighing matrices
                  % for each market (city-quarter) will be stored
while i<T+1
    alc2(i:B+i-1,:)=kron(ones(B,1),(alc(i:B+i-1,1))');

```

```

for j=1:B
    WALC(i:B+i-1,j)=1./(1+2*(abs(alc2(i:B+i-1,j)-alc(i:B+i-1,:))));
end
i=i+B;
end

```

The weighing matrix for the nearest neighbor in Alcohol-Coverage space is calculated with:

```

i=1;
while i<T+1
    alc2(i:B+i-1,:)=kron(ones(B,1),(alc(i:B+i-1,1))');
    cov2(i:B+i-1,:)=kron(ones(B,1),(cov(i:B+i-1,1))');
    i=i+B;
end
WNNACX=zeros(T,B);
for i=1:T; %these loops assign the euclidean distance
    for j=1:B;
        WNNACX(i,j)=sqrt((alc2(i,i-floor((i-1)/B)*B)-alc2(i,j))^2+
            +(cov2(i,i-floor((i-1)/B)*B)-cov2(i,j))^2);
    end
end
%Assigning NaN to diagonal elements
for i=1:T
    for j=1:B
        if j==(i-floor((i-1)/B)*B) & WNNACX(i,j)==0

```

```

        WNNACX(i,j)=NaN;
    end
end
end
%Assigning one's to nearest neighbor
for i=1:T
    for j=1:B
        if WNNACX(i,j)==min(WNNACX(i,:), [], 2)
            WNNACX(i,j)=1;
        end
    end
end
end
%Assigning zeros to non-nearest neighbors
for i=1:T
    for j=1:B
        if WNNACX(i,j)~=1
            WNNACX(i,j)=0;
        end
    end
end
end

```

Weighing matrices for common boundaries in Alcohol-Coverage space are computed with the following programs:

```

i=1;
while i<=T

```

```

CBACX((1+floor((i-1)/B)*B):(B+floor((i-1)/B)*B),:)=
    =cbx([alc((1+floor((i-1)/B)*B):(B+floor((i-1)/B)*B),1)
        cov((1+floor((i-1)/B)*B):(B+floor((i-1)/B)*B),1)]);
i=i+B;
end

```

where "cbx" is a program given by:

```

function CBX=cbx(A);
AX=A(:,1);
AY=A(:,2);
% This fixes the problem of having more than three points
% with the same x (or y) coordinate
AX=AX+(rand(size(AX))*(max(AX)/100)*2-max(AX)/100);
AY=AY+(rand(size(AY))*(max(AY)/100)*2-max(AY)/100);
n=length(A);
%this loop assigns the x-coordinates of the
% intersection points
x=zeros(n-2,(n-1)*(n-2)/2);
i=1;
while i<=(n-2)
    m=((n-2)+(n-i))*(i-1)/2+1;
    while m<=(n-1)*(n-2)/2
        k=i+1;
        while k<=(n-1)
            l=k+1;

```

```

while l<=n
    x(i,m)=0.5*((AX(k)^2-AX(i)^2)*(AY(1)-AY(i))+
        +(AY(k)-AY(1))*(AY(k)-AY(i))*(AY(1)-AY(i))-
        -(AX(1)^2-AX(i)^2)*(AY(k)-AY(i)))/((AX(k)-
        -AX(i))*(AY(1)-AY(i))-(AX(1)-AX(i))*(AY(k)-
        -AY(i)));
    l=l+1;
    m=m+1;
end
k=k+1;
end
end
i=i+1;
end
%this loop assigns the y-coordinates of the intersection points
y=zeros(n-2,(n-1)*(n-2)/2);
i=1;
while i<=(n-2)
    m=((n-2)+(n-i))*(i-1)/2+1;
    while m<=(n-1)*(n-2)/2
        k=i+1;
        while k<=(n-1)
            l=k+1;
            while l<=n

```

```

        y(i,m)=0.5*((AX(1)^2-AX(i)^2)/(AY(1)-AY(i))+
            +(AY(1)+AY(i)))-((AX(1)-AX(i))/(AY(1)-
            -AY(i)))*x(i,m);

        l=l+1;

        m=m+1;

    end

    k=k+1;

end

end

i=i+1;

end

%Creating a Matrix with the distance from i(which is
% equivalent to k,l also!) to each intersection point
D1=zeros(n-2,(n-1)*(n-2)/2);
for i=1:(n-2)
    for j=(1+(n-2+(n-i))*(i-1)/2):(n-1)*(n-2)/2
        D1(i,j)=sqrt((AX(i)-x(i,j))^2+(AY(i)-y(i,j))^2);
    end
end

end

%Creating a matrix with the minimum of all distances from m~=i
%to the each intersection point
D2=zeros(n-2,(n-1)*(n-2)/2);
for i=1:(n-2)
    for j=1+(n+2-(n-i))*(i-1)/2:(n-1)*(n-2)/2

```

```

    for m=1:n
        D3(m,1)=sqrt((AX(m)-x(i,j))^2+(AY(m)-y(i,j))^2);
    end
    D2(i,j)=min(D3);
end
end
clear D3
%Checking whether the intersection points in D1 are the
%closest ones to i,k,l; recording coordinates
x1=zeros(n-2,(n-1)*(n-2)/2);
y1=zeros(n-2,(n-1)*(n-2)/2);
for i=1:(n-2)
    for j=1+(n+2-(n-i))*(i-1)/2:(n-1)*(n-2)/2
        if D1(i,j)-D2(i,j)<0.001
            x1(i,j)=x(i,j);
            y1(i,j)=y(i,j);
        end
    end
end
end
clear x y D1 D2
%Assigning ones to those intersection points which are
% within bounds and 1/2 to the ones off bounds
cm=max(AX);
sm=max(AY);

```

```

x2=0.5*(x1<0|x1>=cm);
x3=(x1>0&x1<cm);
x4=x2+x3;
clear x1 x2 x3
y2=0.5*(y1<0|y1>=sm);
y3=(y1>0&y1<sm);
y4=y2+y3;
clear y1 y2 y3
D4=(y4==1&x4==1);
D5=0.5*(y4==0.5|x4==0.5);
D6=D4+D5; %matrix with 1/2 or 1 given criteria above
clear x4 y4 D4 D5
%Lastly, assigning ones to rival brands with which i
%shares a boundary
%===First, creating a matrix that eases the process
j=1;
while j<=(n-1)*(n-2)/2
    for k=2:(n-1)
        i=n-k;
        while i>0
            M(1,j)=k; %creating the 1st column of new matrix
            i=i-1;
            j=j+1;
        end
    end
end

```

```

        end
    end
    j=1;
    while j<=(n-1)*(n-2)/2
        for k=3:n
            for i=0:(n-k)
                N(1,j)=k+i; %creating the 2nd column of new matrix
                j=j+1;
            end
        end
    end
    end
    D6=[D6; zeros(2,(n-1)*(n-2)/2)];
    CBX=zeros(n,n);
    for i=1:n
        pi=[M(1+(n-2+(n-i))*(i-1)/2:(n-1)*(n-2)/2); N(1+
            +(n-2+(n-i))*(i-1)/2:(n-1)*(n-2)/2); D6(i,(1+
            +(n-2+(n-i))*(i-1)/2):((n-1)*(n-2)/2))];
        for j=1:((n-1)*(n-2)/2-(n-2+(n-i))*(i-1)/2)
            if pi(3,j)==1
                CBX(i,pi(1,j))=1;
                CBX(pi(1,j),i)=1;
                CBX(i,pi(2,j))=1;
                CBX(pi(2,j),i)=1;
                CBX(pi(1,j),pi(2,j))=1;
            end
        end
    end
end

```

```

    CBX(pi(2,j),pi(1,j))=1;
elseif pi(3,j)==0.5 % this case deals with the fact
                    % that some intersection points are
                    % outside the bounds, hence not
                    % all three points share a boundary,
                    % only the closest one to the other
                    % two shares a boundary with both,
                    % but not the other two
                    % with each other
d12=sqrt((AX(i)-AX(pi(1,j)))^2+(AY(i)-AY(pi(1,j)))^2);
d13=sqrt((AX(i)-AX(pi(2,j)))^2+(AY(i)-AY(pi(2,j)))^2);
d23=sqrt((AX(pi(1,j))-AX(pi(2,j)))^2+(AY(pi(1,j))-
        -AY(pi(2,j)))^2);
if (d12+d13)<(d12+d23)&(d12+d13)<(d13+d23)
    CBX(i,pi(1,j))=1;
    CBX(pi(1,j),i)=1;
    CBX(i,pi(2,j))=1;
    CBX(pi(2,j),i)=1;
elseif (d12+d23)<(d12+d13)&(d12+d23)<(d13+d23)
    CBX(i,pi(1,j))=1;
    CBX(pi(1,j),i)=1;
    CBX(pi(1,j),pi(2,j))=1;
    CBX(pi(2,j),pi(1,j))=1;
elseif (d13+d23)<(d12+d23)&(d13+d23)<(d12+d13)

```

```
        CBX(i,pi(2,j))=1;
        CBX(pi(2,j),i)=1;
        CBX(pi(1,j),pi(2,j))=1;
        CBX(pi(2,j),pi(1,j))=1;
    end
end
end
end
clear D6
% llll
for i=1:i
    for j=1:n
        d(1,j)=sqrt((AX(i)-AX(j))^2+(AY(i)-AY(j))^2);
    end
    for j=1:n
        if j==i
            d(j)=NaN;
        end
    end
end
for j=1:n
    if d(j)==min(d)
        CBX(i,j)=0;
    end
end
end
```

end

Once weighing matrices are created, they are interacted with price (and advertising) to create the appropriate weighted vector. For example, cross-prices weighted by the *WBREW* matrix are calculated with:

```
for i=1:T;
    RPBREW(i,1)=WBREW(i,:)*p((1+floor((i-1)/B)*B):(B+floor((i-1)/B)*
        *B),i-floor((i-1)/B)*B);
end
```

where "p" is price. Weighted price and advertising vectors are then transported to Stata where the regression analysis is performed.

B.2. Implied Marginal Costs

Price cost margins ($p - c$) are calculated first. Then price is subtracted to obtain the marginal cost, c . For the Bertrand-Nash case, ($p - c$) is calculated with:

```
%Creating the delta matrix
delta=WBREW.*dau+WBREW.*dpu;
i=1;
j=1;
while i<T+1
    id=nonzeros(cdid2(i:i+B-1)); % this vector contains
                                % non-missing entries
    k=length(id);
    q=vol(i:i+B-1)-1;
    temp1=q(id);
```

```

temp2=delta(i:i+B-1,:)*(-1);
temp2=inv((temp2(id,id))');
mc(j:j+k-1,1)=temp2*temp1; % marginal costs
                                % in unbalanced format

i=i+B;
j=j+k;
end

```

where "dau" and "dpu" are matrices of advertising and price derivatives. *WBREW* is the non-normalized ownership matrix. "cdid2" is a vector that allows conversion of data for each market from balanced to unbalanced format. To illustrate the calculations for leadership scenarios, for the firm leadership case ($p - c$) is calculated with:

```

gammal=brew(:,1); % a 0/1 vector, with ones for leading brands
gammaf=sum(brew,2)-brew(:,1); % a 0/1 vector, with 1's for followers
L=1; % initiating loop for every market
m=1; % index that will be used for storing computed pcm's
while L<T+1
    %*****
    %Defining variables to be used in each market
    id=nonzeros(cdid2(L:L+B-1)); %choosing indices that identify
                                % existing brands in each mkt
    n=length(id); %# of brands in each mkt
    q=vol(L:L+B-1);
    xm=x(L:L+B-1); %price index
    A=kron(ones(B,1),ad(L:L+B-1)'); %Creating a matrix for advertising

```

```

P=kron(ones(B,1),price(L:L+B-1)'); %Creating a matrix for price
cm=cjk2(L:L+B-1,:); %advertising coefficients
bm=bjk2(L:L+B-1,:); %price coefficients
dpm=dpu(L:L+B-1,:); %price derivatives
dam=dau(L:L+B-1,:); %advertising derivatives
gammalm=gammal(L:L+B-1);
gammafm=gammaf(L:L+B-1);
WBREWm=WBREW(L:L+B-1,:);
deltam=delta(L:L+B-1,:);

%*****

%Choosing only brands present in mkt
q=q(id);
xm=xm(id);

A=A(id,id); %advertising enters to the power of .5
P=P(id,id);
bm=bm(id,id);
cm=cm(id,id);
dpm=dpm(id,id);
dam=dam(id,id);

gammalm=gammalm(id);
gammafm=gammafm(id);
WBREWm=WBREWm(id,id);
deltam=deltam(id,id);

%*****

```

```

%Calculating the 3-D derivatives for mkt
%i=row; j=column and k=dimension
for i=1:n
    for j=1:n
        for k=1:n

            if i==j & j==k %3-D diagonal elements
                dqpdp(i,j,k)=(-2)*bm(i,i)*(xm(i)/(P(i,i)^3))+
                    +q(i)/(P(i,i)^2)-dpm(i,i)/P(i,i);
            elseif i==j & i~=k % 2-D Diagonal elements (j=m in theory)
                dqpdp(i,j,k)=(-1)*bm(i,k)*xm(i)/((P(i,i))^2*P(k,k));

            elseif i~=j & j==k % column = dimension (k=m in theory)
                dqpdp(i,j,k)=(-1)*bm(i,k)*xm(i)/((P(k,k))^2*P(i,i));

            elseif i~=j & i==k % row = dimension
                dqpdp(i,j,k)=(-1)*dpm(i,j)/P(i,i);
            else
                dqpdp(i,j,k)=0;
            end

        end

    end

end

end

%*****

```

```

%Calculating PCM of FOLLOWERS

for r=1:n

    count(r,1)=r; % Counting # of brands in mkt

end

countl=nonzeros(gammalm.*count); % entries for leading brands
countf=nonzeros(gammafm.*count); % entries for follower brands
WBREWf=WBREWm(countf,countf); %ownership entries for followers
WBREWl=WBREWm(countl,countl); %ownership entries for leaders
tempdelta=deltam(countf,countf); % delta of followers
tempq=q(countf)-1; % q of followers
temppcm=inv((tempdelta*(-1))')*tempq;

%*****

% Deriving Response Matrix D
dpfol2=dpm(countf,countl); % price derivatives for H matrix
dqdpdp1=dqdpdp(countf,countf,countf); % 3-D matrices for G
dpf=dpm(countf,countf); % price derivatives of follower
pcmfol=(kron(ones(length(temppcm),1),temppcm')).*WBREWf;

                                % PCM followers

f=length(nonzeros(gammafm)); % # of followers in mkt

for w=1:f

    for y=1:f

        %*****

        % This loop calculates the G matrix

```

```

g(w,y)=dpf(w,y)+pcmfol(w,:)*dqdpdp1(:,y,w)+
      +dpf(y,w)*WBREWf(w,y); %dimension y fixed
end
end

dqdpdp2=dqdpdp(countf,countl,countf); % 3-D matrices for H
f2=length(nonzeros(gammalm)); % # of leaders in mkt
for w=1:f
  for y=1:f2
    %*****
    % This loop calculates the H matrix
    h(w,y)=dpfol2(w,y)+pcmfol(w,:)*dqdpdp2(:,y,w);
  end
end
end
d=inv(g)*(h*(-1)); % response
dpreaction=d(1:f,:); % D matrix (price derivatives)
%*****
% Calculating PCM of LEADERS
dpfol=dpm(countl,countf); % Derivatives of followers
tempP=dpm(countl,countl); % Derivatives of leaders
tempA=dam(countl,countl);
for t=1:f2
  for u=1:f2
    %*****

```

```

% This loop calculates the derivatives of leaders
if WBREW1(t,u)==1
    dpl(t,u)=tempP(t,u)+dpfol(t,:)*dpreaction(:,u);% +
            +dafol(t,:)*dareaction(:,u);
else
    dpl(t,u)=0;
end
end
end

pcmm(countl,1)=inv((dpl+tempA)'*(-1))*(q(countl)-1); % elements of
                %PCM corresponding to leaders
pcmm(countf,1)=temppcm; % elements of PCM corresponding to followers
pcm3(m:m+n-1,1)=pcmm;
L=L+B;
m=m+n;

```

B.3. Simulation

A data set that contains marginal costs and demand data in the pre-increase period (1988-1990) is prepared in Stata. Only brands that are present in both the fourth quarter of 1990 and the first quarter of 1991 are considered. For the brand leadership case, the following programs are used in Matlab:

```

global cov NCBCSN alc NCBCSX WAC WNNACX WBREW WPROD2 WREG WSIZE
        WCBCSN WNNCSX WPROD3 WREG vol xnom adnom adnom2 pn adv ad
        x brand ind adv_sim B T N c brew vol2 cdid2 pn2 xnom2
        pricenom2 pricenom demogdummies avar avar2 wn pricevarnom

```

```

xstar_sim_nom w_sim_nom wnom ad_sim_nom wnom2 premium
mcBNmedian mcCmedian mcFLmedian mcCOLL_3firmsmedian
mcCOLL_3brandsmedian pn_sim mcBN mcC mcFL mcCOLL_3firms
mcCOLL_3brands mcCOL_LEADmedian mcFULL_COLmedian
mcBUDleader_allmedian

tic;

load mboundariesNOMINAL WAC WNNACX WSIZE WNNCSX WPROD2 WPROD3 WBREW WREG
load mboundariesNOMINAL2 CBCSN NCBCSX NCBCSN
WCBCSN=CBCSN;

clear CBCSN

load dataequal.txt
W=dataequal;

brand=W(:,1); % brand id's
brew=W(:,9:21); % brewer dummies
alc=W(:,208); % alcohol content, values at 90.4
cov=W(:,209); % 1/cov, values at 90.4
vol=W(:,30); % volume sold
count=W(:,34); % subset of numbers 1-52,932 that allows choosing
                % entries from the matrices WAC, WBREW, etc.
cdid2=W(:,32); % repeating patterns of numbers 1-64, containing
                % missing when brand is not present
demogdummies=W(:,43:175); % demographic and other dummy variables
xindex=W(:,177);
%-----

```

```
%non-mean variables
ad_sim_nom=W(:,183); % nominal advertising
pricevarnom=W(:,182); %nominal price at 90.4
xstar_sim_nom=W(:,184); % nominal sales at 90.4
w_sim_nom=W(:,185); % nominal sales shares at 90.4
wonom=W(:,186); % nominal (modified) sales shares at 90.4

%median pcm's and mc's
mcBNmedian=W(:,187);
mcCmedian=W(:,188);
mcFLmedian=W(:,189);
mcCOLL_3firmsmedian=W(:,191);
mcCOLL_3brandsmedian=W(:,192);
mcBN=W(:,196);
mcC=W(:,197);
mcFL=W(:,198);
mcBUDleader=W(:,199);
mcCOLL_3firms=W(:,200);
mcCOLL_3brands=W(:,201);
mcCOL_LEAD=W(:,210);
mcCOL_LEADmedian=W(:,211);
mcFULL_COL=W(:,212);
mcFULL_COLmedian=W(:,213);
mcBUDleader_all=W(:,214);
mcBUDleader_allmedian=W(:,215);
```

```

c=nonzeros(W(:,33));
demogdummiesdj=[demogdummies xindex];
WAC=WAC(count,:); %selecting 904 data only
WBREW=WBREW(count,:);
WCBCSN=WCBCSN(count,:);
WNNACX=WNNACX(count,:);
WNNCSX=WNNCSX(count,:);
WPROD2=WPROD2(count,:);
WPROD3=WPROD3(count,:);
WREG=WREG(count,:);
WSIZE=WSIZE(count,:);
NCBCSN=NCBCSN(count,:);
NCBCSX=NCBCSX(count,:);
N=length(c);
B=64;
T=length(W);
clear W
%creating the matrices of rival prices and rival advertising
i=1;
pn_sim=zeros(T,B);
adv_sim=zeros(T,B);
wn=zeros(T,B);
while i<T+1
    pn_sim(i:B+i-1,:)=kron(ones(B,1),(pricevarnom(i:B+i-1,1))');

```

```

adv_sim(i:B+i-1,:)=kron(ones(B,1),(ad_sim_nom(i:B+i-1,1))');
wn(i:B+i-1,:)=kron(ones(B,1),(wonom(i:B+i-1,1))');
i=i+B;
end
ind=(pn_sim~=0);
avar=adv_sim;
% Weighing matrix, vectors unbalanced (eliminating zero rows)
WAC=WAC(c,:);
WBREW=WBREW(c,:);
WCBCSN=WCBCSN(c,:);
WNNACX=WNNACX(c,:);
WNNCSX=WNNCSX(c,:);
WPROD2=WPROD2(c,:);
WPROD3=WPROD3(c,:);
WREG=WREG(c,:);
WSIZE=WSIZE(c,:);
NCBCSN=NCBCSN(c,:);
NCBCSX=NCBCSX(c,:);
alc=alc(c,:);
acvdist=acvdist(c,:);
vol=vol(c,:);
[z1 z2 z3 z4 z5 z6]=meanpriceincreaseNOMINAL(pricevarnom);

```

where the program "meanpriceincreaseNOMINAL" is given by (only brand leadership case is reproduced here):

```

function [z1,z2,z3,z4,z5,z6]=meanpriceincreaseNOMINAL(pricevarnom)
global cov NCBCSN alc NCBCSX WAC WNNACX WBREW WPROD2 WREG WSIZE
    WCBCSN WNNCSX WPROD3 WREG vol xnom adnom adnom2 pn adv ad
    x brand ind adv_sim B T N c brew vol2 cdid2 pn2 xnom2
    pricenom2 pricenom demogdummies avar avar2 wn pricevarnom
    xstar_sim_nom w_sim_nom wnom ad_sim_nom wnom2 premium
    mcBNmedian mcCmedian mcFLmedian mcCOLL_3firmsmedian
    mcCOLL_3brandsmedian pn_sim mcBN mcC mcFL mcCOLL_3firms
    mcCOLL_3brands mcCOL_LEADmedian mcFULL_COLmedian
    mcBUDleader_allmedian%First I calculate the own- and cross-
        %coefficients
% The price index and own-price and own-adv coefficients
for i=1:N
    bjj(i,1)=-0.0252899-0.0001086*(acvdist(i))-0.0007144*NCBCSN(i);
    cjj(i,1)=0.000132-0.0000196*(acvdist(i))-0.0000159*NCBCSX(i);
end
%The cross-price and adv coefficients
%these are with normalized PROD, BREW, REG
for i=1:N
    for j=1:B
        bjk(i,j)=0.0005321*WAC(i,j)+0.0008868*WNNACX(i,j)+
            +(0.0017301)*WBREW(i,j)+(0.0093557)*WPROD2(i,j)+
            +(0.0049613)*WREG(i,j);
        cjk(i,j)=0.0000161*WSIZE(i,j)+0.0000707*WCBCSN(i,j)+

```

```
+0.0000398*WNNCSX(i,j)-(0.0003215)*WPROD3(i,j)+
+(0.00053)*WREG(i,j);

end

end

% Defining the constant and price index coefficient
alpha=[-0.0003906 -0.0000882 1.49E-07 -0.0239917 0.0099759
0.0354938 -0.0089721 0.0085787 0.0022645 0.000584
0.005973 -0.0020055 -0.0057535 -0.003341 0.010107
-0.0012724 -0.0044653 -0.0012754 0.0031536 -0.0085126
-0.0131908 -0.0054958 0.1490923 0.0053445 0.0485939
0.0197239 -0.0016011 0.0009127 -0.0010837 -0.000518
0.0011706 0.0184935 0.0049203 -0.0070905 0.0477655
-0.0265357 -0.0138707 -0.023962 -0.0004068 -0.0007255
0.0290781 0.023564 -0.0012706 0.0168251 0.0215111
0.0226133 0.0150868 0.0340253 0.0050704 0.0040284
-0.0020049 0.001964 0.0464521 0.0091258 0.0140341
0.0413293 0.015082 0.0017375 0.0284878 -0.0027244
0.025623 0.0062351 0.0299044 -0.0036716 -0.0221028
0.0072093 -0.0228231 -0.0188224 -0.0003919 0.0187113
-0.0225343 0.0003978 0.0041498 0.0124225 0.0762948
0.0026649 0.0068191 -0.0088122 -0.0080348 0.0098225
-0.0211067 -0.0196201 -0.016998 0.0088228 -0.0260071
-0.0151486 -0.0224111 -0.0172222 -0.0320171 -0.0184702
-0.0236895 -0.0266141 -0.0263237 -0.026684 -0.0303663
```

```

-0.013654 -0.0178036 -0.0204212 -0.0437732 -0.0215876
-0.0214243 -0.0252541 -0.0078813 -0.0186163 -0.0168508
-0.0253362 -0.0176489 -0.0334501 -0.0166824 -0.0196471
-0.0305869 -0.0218276 -0.0175367 -0.0189998 -0.0274318
-0.0113462 -0.0178381 -0.0271592 -0.0287044 -0.0257356
-0.0237532 -0.0254882 -0.022848 -0.0180537 -0.0150007
-0.0351286 -0.0248639 -0.0267993 -0.0227065 -0.0257153
-0.0258188 -0.029947 -0.0277399];

alpha=kron(ones(B,1),alpha);
dj=.0027346;
% Making coefficients balanced
bjku=zeros(T,B); %where cross-price coefficients are stored
ckju=zeros(T,B);
bjju=zeros(T,1); %where own-price coefficients are stored
cjju=zeros(T,1);
bjku(c,:)=bjk; %making these matrices balanced
ckju(c,:)=ckj;
bjju(c,:)=bjj;
cjju(c,:)=cjj;
for i=1:T % joining own and cross derivatives in a single matrix
    bjku(i,i-floor((i-1)/B)*B)=bjju(i);
    ckju(i,i-floor((i-1)/B)*B)=cjju(i);
end
bjku=bjku.*ind;

```

```

cjku=cjku.*ind;
%-----
%%%Creating the delta matrix and ind matrices
%+++++
%ind1
i=1;
while i<T+1
    WBREWA(i:B+i-1,:)=(brew(i:B+i-1,:)*(brew(i:B+i-1,:)'));
    i=i+B;
end
ind1=((pn_sim.*WBREWA)~=0);
%+++++
%ind2
brew2=brew;
brew2(:,1)=brew(:,1)+brew(:,11);
brew2=[brew2(:,1) brew2(:,2) brew2(:,3) brew2(:,4) brew2(:,5)
        brew2(:,6) brew2(:,7) brew2(:,8) brew2(:,9) brew2(:,10)
        brew2(:,12) brew2(:,13)];
i=1;
while i<T+1
    WBREW2(i:B+i-1,:)=(brew2(i:B+i-1,:)*(brew2(i:B+i-1,:)'));
    i=i+B;
end
ind2=((pn_sim.*WBREW2)~=0);

```

```

%+++++
%ind5
brew5=brew;
brew5(:,1)=brew(:,1)+brew(:,2)+brew(:,11);
brew5=[brew5(:,1) brew5(:,3) brew5(:,4) brew5(:,5) brew5(:,6)
       brew5(:,7) brew5(:,8) brew5(:,9) brew5(:,10) brew5(:,12)
       brew5(:,13)];
i=1;
while i<T+1
    WBREW5(i:B+i-1,:)=(brew5(i:B+i-1,:)*(brew5(i:B+i-1,:)'));
    i=i+B;
end
ind5=((pn_sim.*WBREW5)~=0);
%+++++
%ind6
B12=(1201==brand);
B1=(1101==brand);
B49=(2101==brand);
brew6=brew;
brew6(:,1)=brew(:,1)+B12+B49;
brew6(:,2)=brew(:,2)+B1+B49;
brew6(:,11)=brew(:,11)+B1+B12;;
i=1;
while i<T+1

```

```

    WBREW6(i:B+i-1,:)=(brew(i:B+i-1,:)*(brew6(i:B+i-1,:)'));
    i=i+B;
end
ind6=((pn_sim.*WBREW6)~=0);
%+++++
%ind9 (full collusion)
WBREW9=zeros(T,B);
ind9=((pn_sim.*WBREW9)~=0);
%+++++
%ind11
WBREW11=kron(ones(T/B,1),eye(B));
ind11=((pn_sim.*WBREW11)~=0);
%-----
% Simulating the prices after tax increase.
%-----
% Marginal Cost
mcuA=mcBNmedian;
mcuB=mcCmedian;
mcuC=mcFLmedian;
mcuE=mcCOLL_3firmsmedian;
mcuF=mcCOLL_3brandsmedian;
mcuG=mcCOL_LEADmedian;
mcuH=mcFULL_COLmedian;
mcuI=mcBUDleader_allmedian;

```

```

% the following is a special variation of the ad derivative
% needed for simulation. As opposed to derivatives, this matrix
% is computed with nominal current values (not mean values)
for i=1:T
    for j=1:B
        if adv_sim(i,j)>0
            nau(i,j)=(cjku(i,j))/(2*adv_sim(i,j));
        else
            nau(i,j)=0;
        end
    end
end
end
nau=nau.*WBREWA;
clear bju cju
%Scenario 3, BRAND LEADERSHIP (all non-AB brands)
i=1;
k=1;
gammal2=(brand==1101); % 1 if Budweiser
gammal=(brew(:,1)); % a 0/1 vector, with ones for brands
                    % produced by leader
gammalf=ones(T,1)-gammal; % a 0/1 vector, with ones for non-
                    % leading brands (All except A-B brands)
clear y
while i<T+1

```

```

global alpha1 alphadj1 demogdummies1 bjku1 cjku1 avar1 wn1
      mcu1 nau1 ownership1 ownership2 xstar1 y u gammal1
      gammal21 gammaf1 countl1 countl2 countf countf2 A1 w1
id=nonzeros(cdid2(i:i+B-1));
gammal1=gammal(i:i+B-1,:);
gammal21=gammal2(i:i+B-1,:);
gammaf1=gammaf(i:i+B-1,:);
A1=kron(ones(B,1),ad_sim_nom(i:i+B-1)');
demogdummies1=demogdummies(i:i+B-1,:);
bjku1=bjku(i:i+B-1,:); % I could use bjk_sim, but these matrices
                        % are essentially the same
cjku1=cjku(i:i+B-1,:);
avar1=avar(i:i+B-1,:);
ownership11=ind1(i:i+B-1,:); % THESE TWO OWNERSHIP MATRICES
                              % ARE REDEFINED FOR EACH SCENARIO
ownership22=ind1(i:i+B-1,:);
nau1=nau(i:i+B-1,:);
mcu1=mcuI(i:i+B-1,:); %VARIABLE NEEDS IS REDIFINED FOR E/SCENARIO
wn1=wn(i:i+B-1,:);
pricevar1=pricevarnom(i:i+B-1,:);
xstar1=xstar_sim_nom(i:i+B-1,:);
w1=w_sim_nom(i:i+B-1,:);
y=length(id);
gammal1=gammal1(id,:);

```

```
gammal21=gammal21(id,:);
gammaf1=gammaf1(id,:);
for r=1:y
    count(r,1)=r; % Counting # of brands in mkt
end
countl1=nonzeros(gammal1.*count); % entries for leading brands
countl2=nonzeros(gammal21.*count); % entries for leading brands
countf=nonzeros(gammaf1.*count); % entries for follower brands
countf2=countf; % I do this, so that other programs need not be
                % changed substantially (foc and dpreaction)
alpha1=alpha(id,:);
demogdummies1=demogdummies1(id,:);
A1=A1(id,id);
bjku1=bjku1(id,id);
cjku1=cjku1(id,id);
ownership1=ownership11(id,id);
ownership2=ownership22(id,id);
mcu1=mcu1(id,:);
nau1=nau1(id,id);
wn1=wn1(id,id);
avar1=avar1(id,id);
xstar1=xstar1(id,:);
pricevar1=pricevar1(id,:);
w1=w1(id,:);
```

```

xo=pricevar1;
for j=1:y
    u(j,1)=alpha1(j,:)*(demogdummies1(j,:)')+bjku1(j,:)*
        *log(pricevar1)+cjku1(j,:)*(avar1(j,:)')+dj*
        *(log(xstar1(j,:))-(wn1(j,:)*log(pricevar1)))-
        -w1(j,1);
end
z_temp=fsolve(@foc_brand_leadall,xo,optimset('fsolve'));
z_temp2=zeros(B,1);
z_temp2(id,1)=z_temp;
z(i:i+B-1,:)=z_temp2;
clear global alpha1 alphadj1 demogdummies1 bjku1 cjku1 avar1
    wn1 mcu1 nau1 ownership1 ownership2 xstar1 y u gammal1
    gammaf1 gammal21 countl1 countl2 countf countf2 A1 P1 w1
clear u z_temp2 count
i=i+B;
j=j+k;
end
z6=z;

```

where "foc_brand_leadall" is given by the following programs:

```

function F=foc_brand_leadall(z)
global alpha1 demogdummies1 bjku1 cjku1 avar1 dj wn1 mcu1 nau1
    ownership1 ownership2 xstar1 y u countl1 countl2 countf2 A1
    w1 y gammal1 gammal21 gammaf1 w_mod p_over_x2

```

```

for i=1:y
    zvector(i,1)=log(z(i));
end
bstar=bjku1(countl1,countf2);
AA=kron(ones(length(countf2),1),z(countl2)');
BB=kron(ones(1,length(countl2)),z(countf2));
C=AA./BB;
for i=1:y
    w(i,1)=alpha1(i,:)*(demogdummies1(i,:)')+bjku1(i,:)*
        *zvector+cjku1(i,:)*(avar1(i,:)')+dj*(log(xstar1(i,:))-
        -(wn1(i,:)*zvector))-u(i,1);
end
p_over_x=z./xstar1;
w_mod=w-p_over_x;
p_over_x2=kron(ones(y,1),p_over_x');
D=dpreaction_brandall(z,w); %another function outside this program
dstar=D.*C;
F=zeros(y,1);
for i=1:y
    for j=1:y
        if gammal21(i)==0
            if i==j
                F(i,1)=[F(i,1)+(1-(z(i)-0.653-mcu1(i,1))/z(i))*w(i,1)+
                    +(z(i)-0.653-mcu1(i,1))/z(i))*(bjku1(j,i)*

```

```

        *ownership1(j,i)+ownership2(j,i)*nau1(j,i)*
        *z(i))-z(i)/xstar1(i,:)];
%
    F(i,1)=[F(i,1)+(1-(z(i)-mcul(i,1))/z(i))*w(i,1)+
        +((z(i)-mcul(i,1))/z(i))*(bjku1(j,i)*
        *ownership1(j,i)+ownership2(j,i)*nau1(j,i)*
        *z(i))-z(i)/xstar1(i,:)]);
elseif i~=j & (ownership1(j,i)==1 | ownership2(j,i)==1)
%
    F(i,1)=[F(i,1)+(z(j)-mcul(j,1))/z(j)*(bjku1(j,i)*
        *ownership1(j,i)+ownership2(j,i)*nau1(j,i)*
        *z(j))];
    F(i,1)=[F(i,1)+(z(j)-0.653-mcul(j,1))/z(j)*(bjku1(j,i)*
        *ownership1(j,i)+ownership2(j,i)*nau1(j,i)*z(j))];
end
else
    if i==j
%
        F(i,1)=[F(i,1)+(1-(z(i)-mcul(i,1))/z(i))*w(i,1)+((z(i)-
            -mcul(i,1))/z(i))*(bjku1(j,i)*ownership1(j,i)+
            +bstar(j,:)*dstar(:,i)*ownership1(j,i)+
            +ownership2(j,i)*nau1(j,i)*z(i))-z(i)/
            /xstar1(i,:)]);
        F(i,1)=[F(i,1)+(1-(z(i)-0.653-mcul(i,1))/z(i))*w(i,1)+
            +((z(i)-0.653-mcul(i,1))/z(i))*(bjku1(j,i)*
            *ownership1(j,i)+bstar(j,:)*dstar(:,i)*
            *ownership1(j,i)+ownership2(j,i)*nau1(j,i)*

```

```

        *z(i))-z(i)/xstar1(i,:)];
elseif i~=j & (ownership1(j,i)==1)
%       F(i,1)=[F(i,1)+(z(j)-mcul(j,1))/z(j)*(bjku1(j,i)*
           *ownership1(j,i)+bstar(j,:)*dstar(:,i)*
           *ownership1(j,i)+ownership1(j,i)*nau1(j,i)*
           *z(j))];
       F(i,1)=[F(i,1)+(z(j)-0.653-mcul(j,1))/z(j)*
           *(bjku1(j,i)*ownership1(j,i)+bstar(j,:)*
           *dstar*ownership1(j,i)+ownership2(j,i)*
           *nau1(j,i)*z(i))];
        end
    end
end
end
end
end

```

where the program "dpreaction_brandall" is given by:

```

function G=dpreaction_brandall(z,w
global xstar1 A1 w1 bjku1 cjku1 y gammal1 gammal21 gammaf1
       ownership1 w_mod p_over_x2 count1 countl2 countf countf2
%Calculating the 3-D derivatives for mkt
Z=kron(ones(y,1),z');
%Derivatives
for i=1:y
    for j=1:y
        if i==j

```

```

        dpm(i,j)=(bjku1(i,j)-w(i))*xstar1(i)/(Z(i,i)^2);
    else
        dpm(i,j)=bjku1(i,j)*xstar1(i)/(Z(i,i)*Z(i,j));
    end
    if A1(i,j)>0
        dam(i,j)=(cjku1(i,j)*xstar1(i))/(2*A1(i,j)*Z(i,i));
    else
        dam(i,j)=0;
    end
end
end

end

% PCM followers
delta1=ownership1.*dpm+ownership1.*dam;
delta_mod=delta1.*p_over_x2;

%Response matrix
for i=1:y
    for j=1:y
        for k=1:y

            if i==j & j==k %3-D diagonal elements
                dqpdp(i,j,k)=(xstar1(i)/(Z(i,i)^3))*(2*w(i)-
                    -3*bjku1(i,i)); %diagonal elements
            elseif i==j & (j==k & A1(i,i)~=0)
                dqpda(i,j,k)=(-1)*cjku1(i,i)*xstar1(i)/(2*

```



```

WBREWf2=ownership1(countf2,countf2);
WBREWl=ownership1(countl1,countl1);
tempdelta=delta_mod(countf,countf);
tempq=w_mod(countf);
temppcm=inv((tempdelta*(-1))')*tempq;
temppcm2=zeros(y,1);
temppcm2(countf,1)=temppcm;
temppcm2=countf2,1);
pcmfol=(kron(ones(length(temppcm2),1),temppcm2')).*WBREWf2;
%*****
% Deriving Response Matrix D
dpfol2=dpm(countf2,countl2);
dqdpdp1=dqdpdp(countf2,countf2,countf2);
dpf=dpm(countf2,countf2);

f=length(nonzeros(gammaf1));
for a=1:f
    for b=1:f
        g(a,b)=dpf(a,b)+pcmfol(a,:)*dqdpdp1(:,b,a)+dpf(b,a)*
            *WBREWf2(a,b);
    end
end
dqdpdp2=dqdpdp(countf2,countl2,countf2);
f2=length(nonzeros(gammal21));

```

```
for a=1:f
    for b=1:f2
        %*****
        % This loop calculates the H matrix
        h(a,b)=dpfol2(a,b)+pcmfol(a,:)*dqdpdp2(:,b,a);
    end
end
G=inv(g)*(h*(-1)); % response
```

Vita

Christian Rojas was born in Quito, Ecuador, on March 22, 1976. Christian obtained his undergraduate degree in Economics at the Pontificia Universidad Catolica del Ecuador in 2000. Before his graduate studies, he worked in the Ecuadorian telecommunications industry. Christian entered Virginia Polytechnic Institute and State University in 2001 where he obtained the M.A. degree in Economics in 2003 and the Ph.D. in 2005. Christian will be a Visiting Assistant Professor at the University of Texas at Dallas starting in September of 2005.