

Essays on Applied Industrial Organisation

Wanwiphang Manachotphong

Keble College

University of Oxford

A THESIS submitted for the

Degree of

Doctor of Philosophy

At the University of Oxford

Trinity 2009

© 2009 by Wanwiphang Manachotphong

All rights reserved.

Essays on Applied Industrial Organisation

A D.PHIL. THESIS presented by

Wanwiphang Manachotphong

Keble College and Department of Economics, University of Oxford

Trinity 2009

Abstract

The first chapter uses the conduct parameter approach to evaluate the sources of market power enjoyed by UK supermarkets from 1994-2006. During this period, the households rapidly shifted their purchase of fresh liquid milk from the milkmen to the supermarkets. At the same time, the supermarkets' margin grew from negative to over 25 percent. The conduct parameter is used to investigate the sources of such market power. The analysis accounts for the change in consumers' demand pattern, and uses direct measures of the marginal costs. My results suggest that the increase in the supermarkets' margin was explained mainly by a significant decline in the price elasticity of demand for supermarket milk and, to a lesser extent, an increase in the market concentration. I do not find sufficient evident to support any significant improvement of collusion.

Chapter two further investigates collusion by accounting for different vertical arrangements. In 2007, the major dairy processors and retailers admitted to the UK competition commission that they were involved in anti-competitive practices, such as exchanging commercially-sensitive information. This chapter assesses the outcome of such anti-competitive behaviours. It generates a menu of outcomes under different vertical and horizontal competition scenarios. Then, it uses an encompassing test to find the scenario most consistent with the reality. The results suggest that the observed market prices are most consistent with two scenarios; 1) the perfectly competitive (Bertrand) manufacturers and noncooperative Cournot retailers under a linear vertical arrangement, and 2) the noncooperative Cournot retailers under a non-linear vertical contract with the manufacturers. Hence, I do not find sufficient evidence to support that the supermarkets achieved benefits beyond the static-noncoorporative level.

The final chapter investigates the agglomeration in quality-type among hotels in Texas. I adopt a discrete choice framework which allows for strategic interactions among firms. That is, a firm's decision could influence other firms' decisions, and vice-versa. After controlling for demand effects through area-specific variables and area-specific unobserved heterogeneity, the incentive to agglomerate in quality choice still exists. This suggests that hotels of the same quality type enjoy positive spillovers from the presence of one another.

Word count: 58,000 (based on 300 words per page count, excluding references, graphs, and tables)

Acknowledgements

The process of writing this thesis has been a long one. Doubtlessly, many people have contributed to its completion. I am truly grateful to all of them. First and foremost, I am indebted to Dr. Howard Smith, my supervisor, for his invaluable advices and encouragements throughout the course of my graduate study at the University of Oxford. This thesis would not have been possible without him. I would also like to thank Dr. Clare Leaver, my MPhil class teacher and internal examiner, for her selfless support along this academic journey. In addition, I appreciate critical contributions, at various stages of my thesis, from Professor Michael Waterson, Professor Marcel Fafchamps, Dr. John Thanassoulis, Dr. Takamisu Kurita, Dr. James Reade, Dr. Jennifer Castle, and Dr. Victoria Prowse.

A hearty thanks to my Oxford-based friends, who shared with me all the good and bad days, and who made time fly–Wasana Wongsurawat, Natalie Quinn, Raya Muttarak, Yaowaluck Roshorm, Narae Choi, Michael Cheng, Chinnawut Techanuvat, Songpol Chuenkhum, Pornthep Benyaapikul, Prach Panchakunakorn, Charles Lin, Cindy Ding, Ross Kang, Maki Aoyama, Wenjin Zeng, Marcus Eberhardt, Heiko Hesse, Bob Rijkers, Assad Bunto, Yugo Nagamura, Jan-Peter Siedlarek, Panos Nanos, Jean-Paul Carvalho, and Trang Le.

Being an average Thai citizen, my undergraduate degree from Duke University and my graduate degrees from the University of Oxford would not have been possible without financial aids. I am most grateful to the Thai government and all the tax payers in Thailand for the scholarship they provided throughout the past eleven years. I am also thankful to the Doctoral Studentship Scheme (GSS), department of economics, for providing me a financial aid as well as an opportunity to teach during my last year here.

Lastly, I am deeply and forever indebted to my family—papa, mommy, Khun Ta, Ice, and Poppy—for their unconditional love, sacrifice, and support. A special thanks goes to Thorn Pitidol for his selflessness, patient, and understanding.

Contents

Contents ..... **vi**

List of Tables ..... **ix**

List of Figures ..... **xi**

Introduction **1**

1 Investigating Market Power and Collusion under Demand Change: An Analysis of the UK Milk Market **6**

1.1 Introduction and Summary 7

1.2 Industry Background 12

1.3 Demand 15

1.3.1 Data for Demand Estimation 16

1.3.2 Estimation Issues 24

1.3.3 Empirical Demand Specification 34

1.3.4 Demand Estimates 35

1.4 Supply 44

1.4.1 The Model: Conduct Parameter Approach (θ) 45

1.4.2 Data on Marginal Costs 47

1.5 Results and Discussion 50

1.5.1 Conduct and Structure 50

1.5.2 Evolution of the Demand Curve (for Retail Milk) 52

1.6 Conclusion 58

1.A Calculation of the AIDS elasticities using the corrected Stone price index 61

2 Measuring the Effect of Collusion by UK Liquid Milk Suppliers 64

2.1 Introduction and Summary 65

2.2 Literature Review 69

2.3 Market for liquid milk in the UK 73

2.3.1 Supply Chain Structure 73

2.3.2 Demand and Market Equilibrium 75

2.4 Demand 78

2.4.1 The Data 78

2.4.2 Demand Curve Faced by the Retailers 81

2.4.3 Demand Curve Faced by the Manufacturers 87

2.5 Marginal Costs 92

2.5.1 Retailers' Marginal Costs (MC^r) 92

2.5.2 Manufacturers' Marginal Costs (MC^w) 98

2.6 Supply 99

2.6.1 Linear Pricing Vertical Contract 101

2.6.2 Non-Linear Pricing Vertical Contract 108

2.6.3 Perfect Vertical Collusion Model 109

2.7 Empirical Results 110

2.7.1 Menu of Outcomes 110

2.7.2 Encompassing Test 119

2.8 Conclusions 122

2.A Sequential two-step error correction 127

3 Estimating Agglomeration Incentives: The case of the Texas Lodging industry 131

3.1 Introduction and Summary 132

3.2 Literature Review 138

3.3 Data and Industry Background 147

3.3.1 The Texas Lodging Industry Overview 148

3.3.2 Hotel Location and Quality Choice 149

3.3.3 The Data 154

3.4 Test of Product Agglomeration and Dispersion (MTAD) 160

3.4.1 Application 163

3.4.2 Results 165

3.5 The Determinants of Quality Agglomeration 170

3.5.1 The Model 172

3.5.2 Problems with Empirical Estimation 176

3.5.3 Estimation Strategy 1: New Hotels Approach 181

3.5.4 Estimation Strategy 2: Static Bayesian Game Approach 186

3.5.5 Robustness Check 196

3.6 Conclusion 202

References 206

List of Tables

1.1 PCA of Prices of 56 Grocery Product Categories 31

1.2 AIDS Model Estimates 38

1.3 AIDS own-price, cross-price and expenditure elasticity 40

1.4 Robustness Check for Demand Estimation 43

1.5 Supermarket's Operational Costs by Year 50

1.6 Market Share of National-level Supermarkets 1994-2006 53

1.7 Value of the Conduct Parameter and Other Market Indicators by Year 54

2.8 Market Share of Supermarkets and of the Milkmen 75

2.9 Market Share by Types of Milk in 2005 78

2.10 Correlation Matrix of Private-Label Milk Price 2002-2005 81

2.11 Regression Results for the Linear Demand Estimation 88

2.12 Regression Results for the Log-linear Demand Estimation 89

2.13 Summary Statistics of Information for Marginal Costs Estimation 97

2.14 Supermarket's Operational Costs by Year 98

2.15 Asymmetric Capacity Constraints and Marginal Costs 107

2.16 Menu of Outcomes (Pre Mergers) 114

2.17 Menu of Outcomes (Post Mergers) 115

2.18 Menu of Outcomes (Before the First Farmers' Direct Action) 116

2.19 Menu of Outcomes (Between the First and the Second Farmers' Direct Action) 117

2.20 Menu of Outcomes (After the Second Farmers' Direct Action) 118

2.21 P-values of the pairwise encompassing test 123

3.22 Origins of Travelers, 2003 (predicted) 149

3.23 Hotel quality breakdown by year 150

3.24 Major hotel chains and number of affiliated hotels in Texas, 2004 151

3.25 Hotel rating comparison 152

3.26 Number of hotels with and without 9-digit zip code, 2004 160

3.27 Characteristics of all hotels in Texas, 2002 164

3.28 MTAD result for quality and size agglomerations 165

3.29 Variable Definitions and Sample Characteristics 187

3.30 Variable Definitions and Sample Characteristics (continued) 188

3.31 Probit Estimation Result: New Hotel Approach 189

3.32 Variable Definitions and Sample Characteristics 197

3.33 Variable Definitions and Sample Characteristics (continued) 198

3.34 Probit Estimation Result: Static Baysian Game Approach 199

3.35 Robustness Check 203

List of Figures

1.1 Frequency of Milk Price, May 1994- May 2000 20

1.2 Frequency Histograms of Retail vs. Doorstep Milk Price, May 2000 21

1.3 Price and Quantity of Retail and Doorstep Milk (1994-2006). 22

1.4 Budget Share of Milk and Total Grocery Spending (1994-2006). 23

1.5 Price Index of Different Grocery Product Categories 32

1.6 Plots of the First Four Principal Components 33

1.7 The Conduct Parameter and Other Structural Indicators (1994-2006) 54

1.8 Consumers' Demand Curve for Retail Milk by Year 57

2.9 Structure of UK Milk Supply Chain 74

2.10 Milk Quantity Supply by Firm 76

2.11 Price and Household Consumption 77

2.12 p^w- The Actual Wholesale Price (proxied by Wiseman's price). 94

2.13 Manufacturers' Average Marginal Cost. 100

2.14 Linear Pricing Illustration 102

3.15 Quality configurations in 5-digit zip code market, the entire Texas state, 2004. 153

3.16 Geographic location of all hotels in Texas, 2004. The flags represent low-quality hotels. The hollow circles represent high-quality hotels. 155

3.17 Percentage of low-quality hotels in 5-digit zip markets. 167

3.18 Percentage of low-quality hotels in county markets. 167

3.19 Percentage of small hotels in 5-digit zip markets. 168

3.20 Percentage of small hotels in county markets. 168

3.21 Illustration of distance bands in a sample market – Burnet county. The Flags represent low-quality hotels, hollow circles present high-quality hotels. 177

Introduction

The continuously increase in data availability, coupled with recent advances in estimation techniques and micro-economic theory, have made the analysis of many applied industrial organization issues become plausible. This thesis serves as an example of this claim.

The first chapter evaluates the horizontal competition among the major supermarkets during the period 1994-2006. In doing so, one needs to understand how the demand for retail milk and doorstep milk had changed during the same period. This is because, the change in demand elasticity contributes greatly to the change in firms' market power. The second chapter examines vertical competition among the UK milk processors and supermarkets from 2002-2005. It assesses whether there has been any evidence of collusion among these business entities. Finally, the last chapter studies quality choices made by hotels in Texas. This chapter investigates whether hotels of the same quality type enjoy positive spillovers from one another. It also allows firms to interact strategically using the static-Bayesian game setting (See, for example, Seim(2006), Orhun(2006), and Ellickson and Misra (2008)). The rest of the introduction discusses each of the chapters in more detail.

In the rst chapter, I investigate the long-term evolution of demand for fresh liquid milk in the UK from 1994-2006. Then, I assess the retailers' competition taking into account the demand change. During the sample period, UK households increasingly shifted away from doorstep milk towards retail milk. At the same time, the retailers' profits rapidly grew.

In 1994, the average profit-margin of the five largest supermarkets[1] was almost negative. By 2006, their profit-margin increased to about 26-27 percent on average[2]. This chapter, therefore, attempts to find the cause of such remarkable profit growth.

To do so, I adopt the conduct parameter approach, which suggests that a firm's ability to gain a positive profit-margin comes from three sources, i.e. collusion, market concentration, and low price elasticity of demand. I was able to obtain firms' profit-margin and market concentration[3] from reliable sources. On the other hand, demand elasticity is not observed and has to be estimated. After knowing the price-cost margins, market concentration, and demand elasticity, the level of collusiveness can then be assessed.

The almost ideal demand system (AIDS) is used to estimate the demand elasticity. My empirical specification allows the demand curve to shift and/or rotate through time. I obtained consumers' demand and price data from the national food survey (NFS) and the expenditure food survey (EFS). From assessing the change in the demand curve year-by-year, I found that consumers have become less sensitive to changes in the price of retail milk.

Results from the conduct parameter analysis suggest that the increase in the retailers' profit-margin was not due an improvement of collusion. But rather, was explained mainly by the significant decline in the price elasticity of demand for retail milk and, to a lesser extent, the increase in the market concentration.

---

1 Then, Tesco, Asda, Sainsbury, Morrisons and Safeway.

2 Profit margin = $\frac{\text{Price - Marginal Cost}}{\text{Price}}$. The price information is obtained from the National Food Survey (NFS) and the Expenditure Food Survey (EFS). The marginal cost information is obtained from the UK Milk Development Council (MDC).

3 The market share information was obtained from the Institute of Grocery Distribution's (IGD) report.

The second chapter further investigates collusion in the milk industry at both the vertical and horizontal levels. In 2007, major dairy manufacturers and supermarkets admitted to the UK competition commission that they were involved in some certain anti-competitive practices; particularly, exchanging commercially-sensitive information. This study, therefore, measures the supermarkets' gain from such practice. To do so, I generate a menu of prices and profits under different vertical and horizontal schemes. Then, I employ an encompassing test to find the scenario(s) most consistent with the actual market outcomes.

Despite the admitting of anti-competitive practices, I find that neither the manufacturers nor the retailers succeeded to earn higher than the as-if static noncooperative Cournot profit. Results from the encompassing tests suggest two most consistent competition scenarios; 1) the perfect Bertrand manufacturers and noncooperative Cournot retailers under a linear vertical contract, and 2) the noncooperative Cournot retailers under a non-linear contract with the manufacturers. To investigate why firms did not achieve more than non-collusive profits is beyond the scope of this paper. Some of the many possible explanations are, for example, failure to collude, threat of entry, and regulatory constraints.

This chapter also estimates the minimum welfare loss caused by the supermarkets. The results suggest that the average profits of the alleged dairy processors and supermarkets were 55.76 and 271.20 million pounds per year (or about 0.92 and 6.15 pounds per capita[4]) respectively. Compared with the most welfare-maximized scheme, where suppliers receive no profit (price = marginal cost) and quantity supplied reaches the highest possible level, the current market outcome could result in at least 27.04 million pounds per year (0.46 pounds

[4] Based on the Office for National Statistics' (ONS) population estimate, the population of the Great Britain (we exclude Northern Ireland in this research) was 58,136,000 in 2004.

per capita) of deadweight loss. The loss could have reached at least 772.17 to 814.25 million pounds per year (13.28 to 14.00 pounds per capita), had the manufacturers and retailers been perfectly collusive. Due to the data limitation, it is not possible to estimate the actual and the maximum deadweight loss without making a strong assumption about the shape of the demand curve.

The last chapter studies quality choice made by hotels in Texas during the period 1998-2004. This study is motivated by the agglomeration literature, which suggests that firms in some certain industries enjoy positive spillovers from co-locating. For example, antique stores, cheap food vendors, luxury fashion stores, etc. Positive spillovers can be created differently depending on the nature of each industry. Some explanations include reduction in consumers' search cost (Fischer and Harrington, 1996), firms' learning about demand (Toivanen and Waterson, 2005), and favourable investment on public goods (Enz et al., 2005).

I examine whether there exist such positive spillovers from co-locating of hotels of the same quality-type in Texas. In the first part of my analysis, I use the non-parametric Multinomial Test for Agglomeration and Dispersion (MTAD) to confirm that hotels in Texas were actually agglomerated in quality. In the second part, I employ econometric tools to investigate whether the competitors' quality choice can influence a hotel's own profit.

Such analysis involves difficulties in disentangling firms' strategic interactions, i.e. simultaneity of choices. Two different estimation approaches were adopted to solve the problem. In this study, the first approach is called the "new hotels" approach, while the

second one is called the "static-Bayesian game" approach. The new hotel approach assumes that firms are short sighted–i.e. only the decisions already made by the incumbents matter to them. This new hotel approach allows for a less complicated estimation process, but is only feasible only when the timing of decisions is available (Sorensen, 2006).

To relax the short-sighted assumption, the static-Bayesian game approach assumes that firms are capable of predicting their competitors' actions in equilibrium. Therefore, upon making their own decisions, firms take into account both the incumbents' and the future entrants' actions. Given this interlinked decision pattern, the estimation of the static-Bayesian game is cumbersome. However, due to the recent developments of computational technologies, this framework has recently been utilized more extensively. (See, for example, Seim(2006), Orhun(2006), and Ellickson and Misra (2008))

The results from the new hotel approach and the static-Bayesian game approach are consistent. After controlling for demand effects–through area-specific variables and area-specific unobserved heterogeneity–the incentive to agglomerate in quality choice still exists. In particular, the expected profit of a low-quality (high-quality) hotel increases with the number of low-quality (high-quality) competitors nearby. I find that high-quality hotels generate more positive spillovers than low-quality hotels counterpart. In both cases, the spillovers fade away with distance to competitors.

Chapter 1
Investigating Market Power and Collusion under Demand Change: An Analysis of the UK Milk Market

Abstract

During the period 1994-2006 UK households rapidly shifted their purchase of fresh liquid milk from the milkmen to the supermarkets. At the same time, the supermarkets' margin grew from almost negative to over 25 percent. Where did such market power come from? This chapter uses the conduct parameter approach to evaluate the sources of market power enjoyed by UK supermarkets from 1994-2006. The analysis accounts for the change in consumers' demand pattern, and uses direct measures of the marginal costs. My results suggest that the increase in the supermarkets' margin was explained mainly by a significant decline in the price elasticity of demand for supermarket milk and, to a lesser extent, an increase in the market concentration. I do not find sufficient evident to support a significant improvement of collusion.

1.1 Introduction and Summary

Before 1990 most UK households had their milk delivered at home by the milkman. Milk came in glass bottles and lasted for only a few days. After finishing their milk, households had to return the empty bottles to the milkman for re-use. Nowadays, this practice has become history for most British children. With the longer shelf-life of milk, better refrigeration, and more women working outside the home, etc., supermarket milk became more suitable to the modern life-style. By 2006, only 10 percent of households remained loyal to their milkman. The rest had already shifted towards retail milk[5]. Consumers now choose to purchase their milk from a variety of retail outlets. Marking up the price might seem difficult in this competitive, homogenous goods context. Yet what is observed are higher prices and increasing retailers' profit. What made this possible? Where did the market power come from? This paper investigates the sources of market power and collusiveness among UK milk retailers for the period 1994-2006.

I analyse the situation using the conduct parameter approach (see for example, Bresnahan (1989), Genesove and Mullin (1998), Wolfram (1999), Rosen (2007)). This method evaluates the market power from its three sources–demand elasticity, market concentration and collusion. If market power is defined as the power to earn above-competitive profit, then the conduct parameter approach proposes:

$$L_t = \sum_i s_{it} \frac{P_t - MC_{it}}{P_t} = -\frac{H_t \theta_t}{\varepsilon_t} \tag{1.1}$$

[5] Source: the monthly Expenditure Food Survey 2001-2006. Hereby, the term "retail milk" will be used to denote milk from supermarkets and other retail outlets.

where subscript i denotes firm and t denotes time, L_t is the industry's elasticity-adjusted Lerner's index, s_{it} is market share, P_t is price, MC_{it} is marginal cost, H_t is the Herfindahl index of concentration, ε_t is the price elasticity of demand *for retail milk* and finally, θ_t is the conduct parameter.

From expression (1.1), firms' ability to earn a higher margin comes from 1) low price elasticity of demand (ε_t) , 2) high market concentration (H_t) and, 3) high level of collusiveness (θ_t). To evaluate collusiveness, I obtain accurate values of price elasticity, market concentration and the retailers' margin. I find that the increase in the retailers' price-cost margin was mainly due to lower elasticity of demand and higher market concentration rather than higher collusiveness.

The demand elasticity is estimated using the almost ideal demand system (AIDS) framework (Deaton and Muellbauer, 1980b). I extend the original AIDS function in two ways. First, since my data set covers a long period of time, I allow the demand curve to shift and for the consumers' willingness to pay to change. Second, I use the principal component analysis (PCA) to compress the dimensionality of numerous candidates for control variables[6] (Hoderlein and Lewbel, 2007). I find that, over time, the equilibrium price elasticity of retail milk was falling. In addition, the aggregate demand curve for retail milk was shifting outward and becoming steeper.

In contrast to price elasticity, the evaluation of market concentration is straightforward. I calculate the Herfindahl index using the market share information obtained from

[6] Although the PCA has been proved useful given the increasing availability of data, it has not been used as much in the demand literature. The only, but most comprehensive, literature on the use of PCA in demand estimation is Hoderlein and Lewbel (2007).

the Institute of Grocery Distribution's (IGD) report. This report was constructed by using the Taylor Nelson Sofres' (TNS) homescan data base. Over time, major retailers gained more market share–resulting in a higher concentration.

Having obtained the demand elasticity of retail milk and the Herfindahl index, eq. (1.1) is eventually employed to find the conduct parameter. This can be done in two ways. One can either *estimate* the marginal cost (MC) simultaneously with the conduct parameter (θ); or find a direct measure of the marginal cost then *calculate* θ. The first method is known as the New Empirical Industrial Organization (NEIO), a technique comprehensively discussed in Bresnahan (1989). This method has become very useful and has made possible the evaluation of collusion in many cases, especially those with limited marginal cost data. Previous literatures adopting or assessing the NEIO approach are, for example, Genesove and Mullin (1998), Wolfram (1999), Kim and Knittel (2006) and Rosen (2007). Despite its utility, the NEIO technique has been heavily criticized. Corts (1999) shows that the method would give bias results if firms do not actually play a static conjectural variations game. To avoid this problem, I directly measure the retailers' marginal cost using highly reliable information. This method remedies the NEIO's potential bias from estimating MC and θ simultaneously.

Turning to an assessment of the conduct parameter, the values of θ obtained suggest no significant improvement of collusion. In fact, it is very close to the static Bertrand level from 1994-1996. Then, rises beyond the static noncooperative Cournot level between 1997-2000. Finally, it drops back and remains around the static noncooperative Cournot level from year 2001. Many explanations beyond the focus of this paper can be employed

to explain this result. For example, failure to collude, threat of entry and regulatory constraints, to name but a few.

Some caveats should be noted on the model choice. The conduct parameter approach is built upon a static conjectural variations framework, e.g. Bertrand, Cournot and Monopoly. Thus, if firms do not play those games in reality, the model is obviously not ideal for the analysis. In any case, this study does not intend to predict the actual game played by firms. Rather, it uses the well-studied conjectural variations game as the benchmark. Some attempts to relax the static conjectural variations assumption are such as those of Puller (2006) and Klemperer and Meyer (1989). Puller (2006) relaxes the static assumption of the conduct parameter model by allowing firms to maximize the sum of several periods' profit subject to a cheating constaint. However, the modelling of the cheating constaint would still be subjective. Klemperer and Meyer (1989), on the other hand, suggest alternative supply rules to the conjectural variations game[7]. However, without knowing exactly what game is played, I opt for the a more conventional conjectural variations game.

Another point worth making is that, in reality, supermarkets compete in a multi-product market. Ignoring this crucial characteristic of competition could result in an incomplete understanding of the situation. Nevo (2001) discusses that multi-product firms could gain market power from the portfolio effect, or the basket effect in the supermarket context. That is, if two products are imperfect substitutes, a firm selling both could charge more than two firms selling each. Bliss (1988) discusses that multi-product shop-

[7] Unlike the Bertrand competition which uses price as the strategic variable, or the Cournot competition which uses quantity as the strategic variable, their model allows the strategic variables to be both. In any case, the noncooperative outcomes from Klemperer and Meyer(1989)'s supply function are bounded by the Bertrand and Cournot outcomes.

ping outlets can save customers' shopping costs. In addition, Harrington (1987) proposed a theoretical model of collusion among multi-product firms. His model suggests that it is easier for multi-product firms to collude. This paper abstracts from the multi-product nature of the problem. It is important that readers keep in mind this crucial point.

This paper is guided by several variations of the conduct parameter literature. Genesove and Mullin (1998) investigate collusion in the US sugar industry. The paper compares results from the NEIO and the direct measure approach. They find that, although the NEIO's markups fall outside the 95-percent interval, the results do not give a significantly different economic implication. Wolfram (1999) and Clay and Trosken (2003) investigate the UK electricity spot market and the US whiskey industry respectively. They arrive at similar conclusions. In contrast, Kim and Knittel (2006) study the California electricity market and find that the NEIO poorly predicts the conduct parameter. Puller (2006) introduces a framework for the conduct parameter under repeated games, while Puller (2007) develops a conduct parameter model which takes into account the capacity constraint. Rosen (2007) suggests an alternative to the NEIO approach. To avoid bias caused by estimating cost and conduct simultaneously, he suggests fixing some firms' decision rules. Then estimates the range of marginal cost which is consistent with the pre-specified rules. This study adds to the conduct parameter literature by allowing the conduct parameter and price elasticity to change in response to the change in demand over time.

The analysis proceeds as follows. The next section describes the industry background. Section 2 discusses demand estimation, while section 3 discusses supply and conduct estimation. Section 4 reports and interprets the results. Section 5 concludes.

1.2 Industry Background

From raw milk to processed liquid milk: Over the past few decades, the UK dairy supply chain has experienced tremendous structural change. Prior to 1994, dairy production and marketing was regulated by one government enterprise called the "Milk Marketing Board" (MMB). This enterprise was established after the second world war to rescue the collapse of the UK milk market and to coordinate trading among farmers, milk processors and retailers. It was responsible for setting retail and wholesale milk prices and contracted between farmers and (mostly local) dairy processors. The MMB itself also owned a large-scale dairy processor, called Dairy Crest.

In the early 1990's, the board struggled with technological change and experienced high inefficiency (Empson, 1998) (Franks, 2001). In 1994, the government ended the statuory monopoly role of the MMB. A new entity called the Milk Marque was established instead as a voluntary farmer cooperative. This occured alongside with the privatization of Dairy Crest. Large dairy processors started to contract directly with farmers. Many other farmer cooperatives emerged. Although the transformation of the MMB into the Milk Marque started a new era of a competitive market for raw milk, the Milk Marque still managed a large volume of raw milk and had signifcant market power. In 1999, the competition commission (CC) ordered the Milk Marque, which then held about 49.6 percent of market share, to split into three entities, namely, Zenith Milk Ltd., Axis Milk Ltd. and Milk Link Ltd. (CC, 2000). Since then, the farmers have lost significant price-setting power.

In contrast to the end of the MMB and the splitting-up of the Milk Marque, dairy processors have become more consolidated. Over the past twenty years, better technology

has increased the economy of scale and enabled milk plants to operate on a large scale. In addition, better refrigeration and transportation eliminates the necessity for plants to locate near retail outlets. In 2003, about 97 percent of the UK's fresh liquid milk was processed by only four dairy processors–Arla, Dairy Crest, Express and Wiseman (CC, 2003b, p.93). Market concentration increased even further when Arla successfully acquired Express in 2004.

The industry was hit by the Foot and Mouth Disease (FMD) in 2001. This disease is highly contagious among cloven-hoof livestock, such as sheep, pigs, and cows. Transmission to human is possible, though has been rare. Thus, the impact of the FMD was mainly on agriculture rather than on human health. A large number of sheep and cattle were killed in attempt to stop the widespread of the disease. However, according to DairyCo, the total milk stock was decreased by only 0.22 percent (31 million liter) in 2001 (DairyCo, 2009). Also, about half of the total milk stock was usually processed into fresh liquid milk. Therefore, it was possible to process a larger proportion of raw milk into fresh liquid milk to stabilize the supply. The top panel of figure (1.3) shows that household consumption of fresh liquid milk did not drop dramatically during the FMD epidemic. Thus, milk consumption was not greatly affected by the outbreak of the disease.

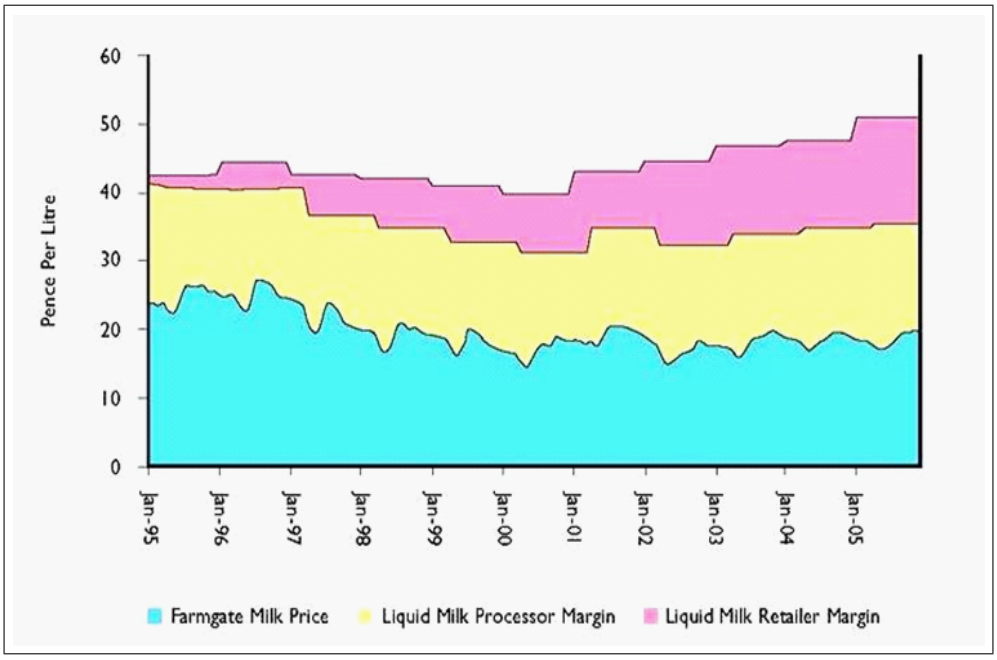
From processed liquid milk to consumers: Apart from in the raw milk market, there has also been significant changes in the retail structure and consumption pattern. In 1980, doorstep milk accounted for almost 90 percent of the UK households' liquid milk consumption (CC, 2003a). However, with better refrigeration and the longer shelf-life of milk, there has been a rapid shift away from doorstep milk towards retail milk. According

to the MDC, the UK's total doorstep consumption was 2,565 million litres per annum (45.5 litres per capita) in 1995 and dropped to only 467 million litres per annum (7.9 litres per capita) in 2006. Retail milk consumption, in contrast, has been growing steadily from 3,076 million litres per annum (54.6 litres per capita) in 1995 to 4,293 million litres per annum (73.0 litres per capita) in 2006.

Many factors can be claimed to contribute to this shift. For example, (i) when household size gets smaller, a household may need less regular supply of milk; (ii) availability of milk in retail outlets reduces the importance of the milkman; (iii) more women working away from home means there is no one to wait for milk delivery; (iv) new households are likely to be unfamiliar with the doorstep delivery practice. Given such changes in the households' consumption pattern, it is doubtful that the demand curve for retail milk has remained static. Sections 1.3 and 1.5.2 of this thesis investigate the evolution of the demand curve for retail milk over time.

Along with the change in consumers' consumption pattern, there has also been a change in the suppliers' relative profit. The figure below shows prices and margins for liquid milk from January 1995 to January 2006. The plot suggests that farmgate milk price fluctuates seasonally. However, the average farmgate price was falling from 1995-2000. It then, increased at a very slow rate from 2001 onward. The processor's price was smoother than the farmgate price, but seemed to reflect the adjustment of the farmgate price on average. The processor margin, therefore, remains rather constant over time. The retailer margin, on the other hand, seems to be growing over time. In 1995, the retailers hardly made any profit from selling milk. However, by the end of 2005, their profit margin

had grown to about 27.4 percent[8]. The main focus of this paper is to explain the retailers' growing profit margin during the period 1994-2006. The next section investigates one of the possible explanations, the change in demand.



Prices and Margins for Liquid Milk. Sources: TNS, MDC, Defra.

1.3 Demand

This study uses the Almost Ideal Demand System (AIDS) framework to estimate demand for retail milk. The AIDS is expressed in terms of budget share. It was developed from the double logarithmic demand model where the log-linear demand function was modified in order to satisfy the adding-up condition[9]. The AIDS also allows for the homogeneity and

[8] According to the TNS and MDC data reported in table(1.7), the average retail price was 51.60 pence per litre(ppl), while the average marginal cost was 37.44 ppl in 2005. Therefore, the average price-cost margin $(\frac{P-MC}{P})$ was 27.4 percent.

[9] Four properties which ensure that the demand function is derived from consumers' utility maximisation

Slutsky symmetry conditions to be tested or imposed (Deaton and Muellbauer, 1980a). I use the same notation as in Deaton and Muellbauer (1980b) to express the AID function as follows:

$$w_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log(x/P) \tag{1.2}$$

where w_i is the consumers' budget share of good i, p_j is the price of good j, x is the total expenditure, α_i, γ_{ij}, β_i are parameters, and P is the translog AID price index defined as

$$\log P = \alpha_0 + \sum_j \alpha_j \log p_j + \frac{1}{2} \sum_i \sum_j \gamma_{ij} \log p_i \log p_j \tag{1.3}$$

where α_0, α_j and γ_{ij} are parameters. Using parameter estimates from the AIDS budget share equation (1.2), the price elasticities of demand and expenditure can be calculated as follows:

AIDS Price elasticities:

uncompenstated (ε_{ij})

$\varepsilon_{ij} = \gamma_{ij}/w_i - \beta_i\alpha_j/w_i - (\beta_i \Sigma_j \gamma_{ij} \ln p_j)/w_i - \delta_{ij}$

compensated (ε^*_{ij})

$\varepsilon^*_{ii} = \eta_{ij} + w_j(1 + \beta_i/w_i)$

where δ_{ij} = 1 for $i = j$ and δ_{ij} = 0 for $i \neq j$

AIDS Expenditure elasticity:

$(e_i) = \beta_i/w_i + 1$

1.3.1 Data for Demand Estimation

The data used in demand estimation can be regarded as falling into two groups. The first group contains those specified in the AIDS model, e.g. milk price, quantity, total expenditure and prices of substitutes and complementary goods. The second group contains the control variables and instrumental variables required to remedy potential endogeneity of price and total expenditure. I first discuss the sources of data generally required by the

---

are adding up, homogeneity, Slutsky symmetry, and negative semi-definiteness of the Slutsky matrix(Mas-Colell et al., 1995).

AIDS model. Data sources of control and instrumental variables will be discussed when I analyse those issues in the next section.

Retail and doorstep milk price and quantity are obtained from the National Food Survey (NFS) for the years 1994-2000; and from the Taylor Nelson Sofres (TNS) for the years 2001-2006. Since the data were obtained from two different sources, I will have to include a time dummy = 1 for TNS observations. The TNS data set was downloadable from the MDC's website. It contains four-weekly price, expenditure and quantity of retail and doorstep milk from 2001 to 2008. On the other hand, the NFS data set was available on the UK data archive's website. This survey collected data from UK households from 1961 to 2000. However, it did not cover households in Northern Ireland until 1975. The NFS originally aimed to "monitor the adequacy of the diet of urban working class households in wartime" (UKDA). Questions were occasionally modified to serve more up-to-date research issues. About 8,000 households took part in the NFS each year. They were first interviewed by an NFS staff for general household characteristics. Then, the households were asked to keep a diary, recording their food consumption information, for seven consecutive days. The information includes each household's expenditure on every food category, quantity and price. Since households were participating at different times of the year, it is possible to construct data at different frequency levels, i.e. weekly, monthly, quarterly and annually. This paper analyses data on a monthly basis.

Unfortunately, the information of whether households' milk consumption was from doorstep delivery or retail store was not available in the NFS data until 1999. Therefore, I use my own identification method to separate doorstep and retail purchases. Since the

doorstep milk price was always higher than the retail milk price, I cross-check this fact by generating histograms of price frequency. Figure (1.1) shows frequency histograms of milk price from May 1994- May 2000. The histograms have twin peaks in all panels. Figure (1.2) shows the actual frequencies of retail and doorstep milk price in May 2000 (when the retail/doorstep information is available). This figure confirms that the lower-value peak is the mode of the retail price, whereas the higher-value peak is the mode of the doorstep price. Through this frequency plot experiment, I obtain the mode price of retail and doorstep milk. To separate retail from doorstep purchase, I arbitraily assume a cut-off point = mean price after dropping all the outliers with value above 100 pence per litre. Then, assume that all purchases with price below mean belong to retail milk, while the rest belongs to doorstep milk. By cross-checking this with year 2000 data–when the actual doorstep identification is available–my identification method is 83.7 percent accurate. Other reasonable methods were also examined. For example, using the mid-point between the two modes (retail price and doorstep price) yields 82.3 percent accuracy, while using the local antimode between the two modes yields 66.5 percent accuracy.

Monthly per-capita expenditure on grocery was obtained from the NFS for the years 1994-2000 and from the Expenditure Food Survey (EFS) for the years 2001-2006. The EFS was designed to replace the old NFS and Family Expenditure Survey (FES)–both of which were ended in 2001. Therefore, my expenditure data obtained from the NFS and the EFS should still be consistent.

Prices of substitute and complementary products are from the Retail Prices Index (RPI) data set available on the Office for National Statistics (ONS) website. This data

set reports the current price (year 1987=100) of different retail product categories such as bread, cereals, cheese, eggs, biscuits, beef, poultry, etc. All together, prices of 30 different food and non-alcoholic beverage categories are available. The fact that only current price– not price level–is available does not affect my demand analysis. This is because only the variation in price is needed for identification. In my estimation, I change the base year of all price variables to 1994. This also applies to the retail milk price, the doorstep milk price and the total grocery expenditure. Thus, the price of all products would be normalized 1 in 1994.

The top panel of figure (1.3) plots prices of retail and doorstep milk over the period 1994-2006. Since the NFS data for the years 1994-2000 is available at per-item level, I am able to find mode price in addition to mean price. The bottom panel of figure (1.3) plots the average consumption of retail and doorstep milk (in litre/person/day) over the sample period. Retail consumption was rising, while doorstep comsumption was falling through time. On average, retail milk consumption was 0.10 litres/person/day in 1994 and increased to 0.21 litres/person/day in 2001. Doorstep milk consumption was 0.12 litres/person/day in 1994 and dropped to 0.02 litres/person/day in 2006.

The budget shares of retail and doorstep milk are shown in the top panel of figure (1.4); total grocery spending is shown in the bottom panel. In this context, total grocery spending accounts for spending only on food and non-alcoholic beverages. This spending, reported in nominal terms, was increasing steadily. On the other hand, spending on milk has been falling with an increasing proportion contributed to retail milk and a decreasing proportion contributed to doorstep milk.

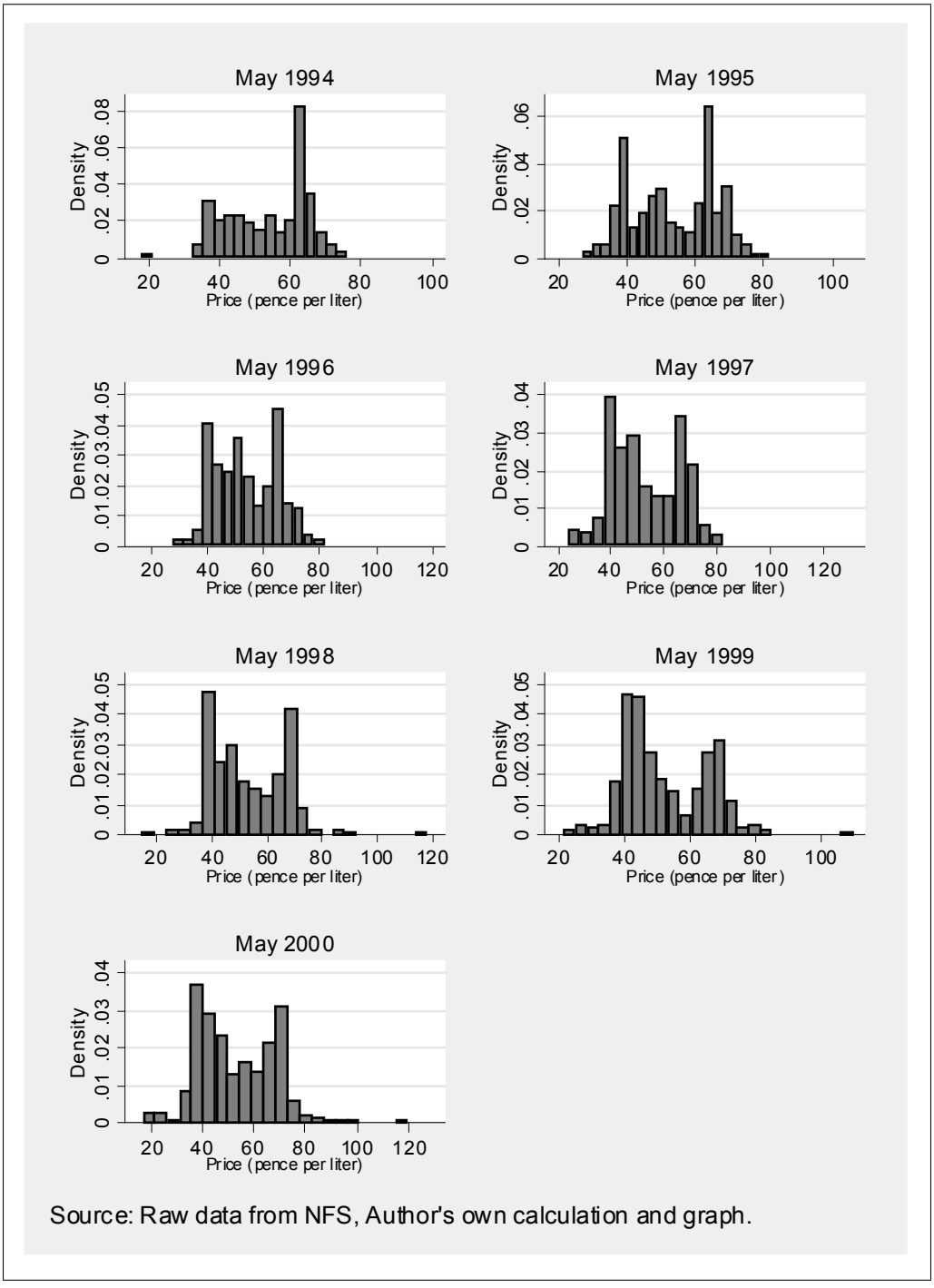


Fig. 1.1. Frequency of Milk Price, May 1994- May 2000

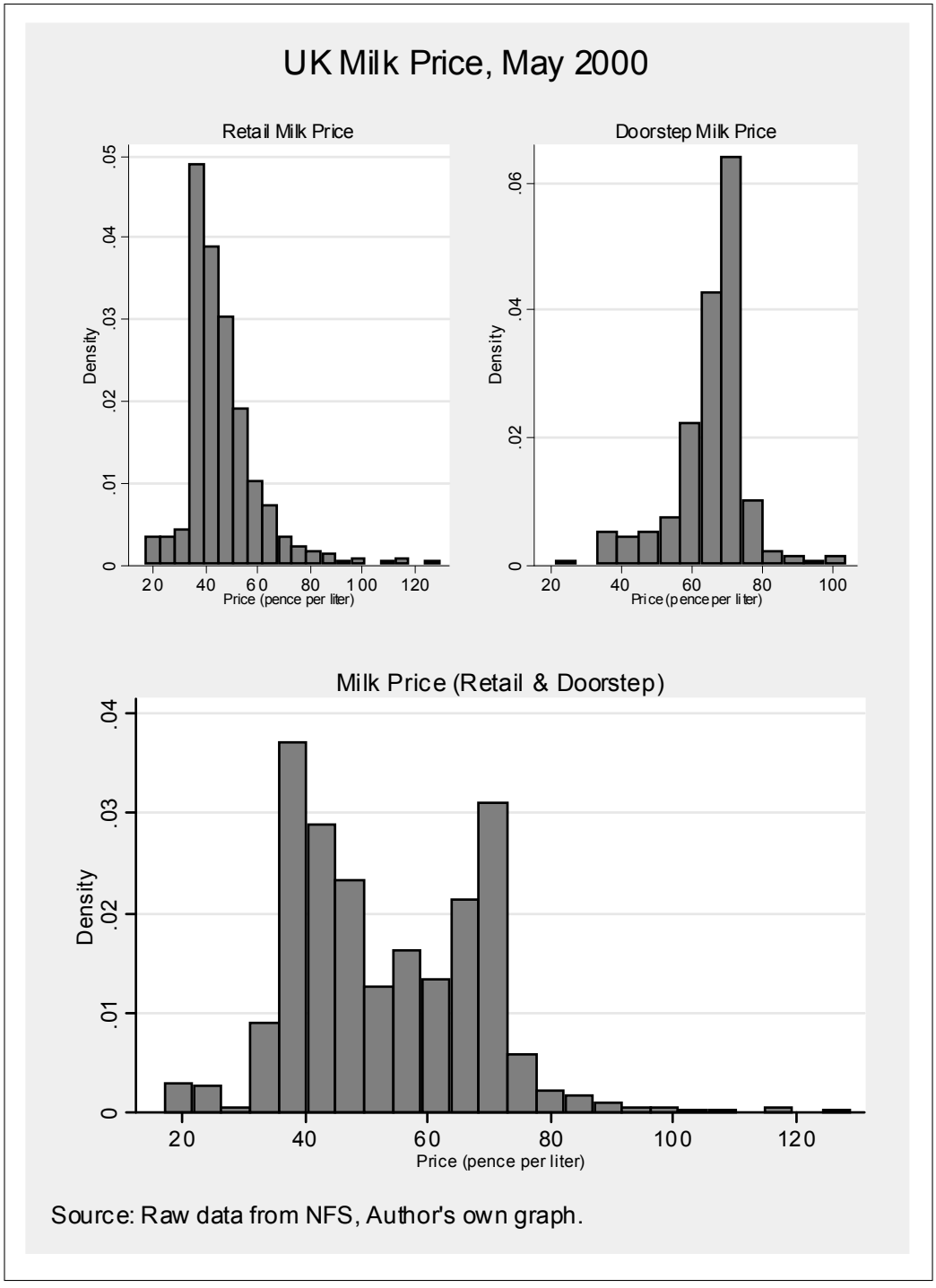


Fig. 1.2. Frequency Histograms of Retail vs. Doorstep Milk Price, May 2000

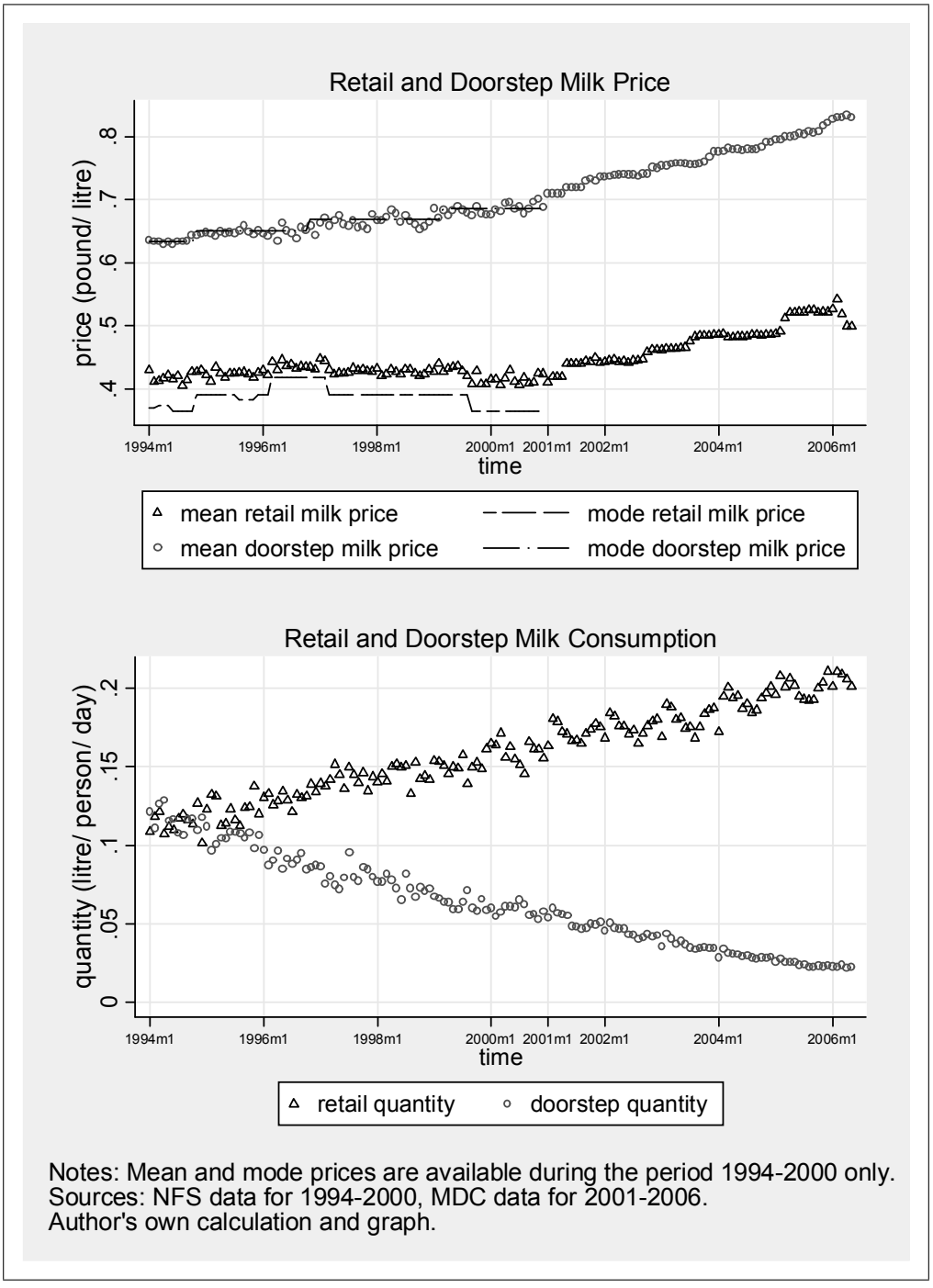


Fig. 1.3. Price and Quantity of Retail and Doorstep Milk (1994-2006).

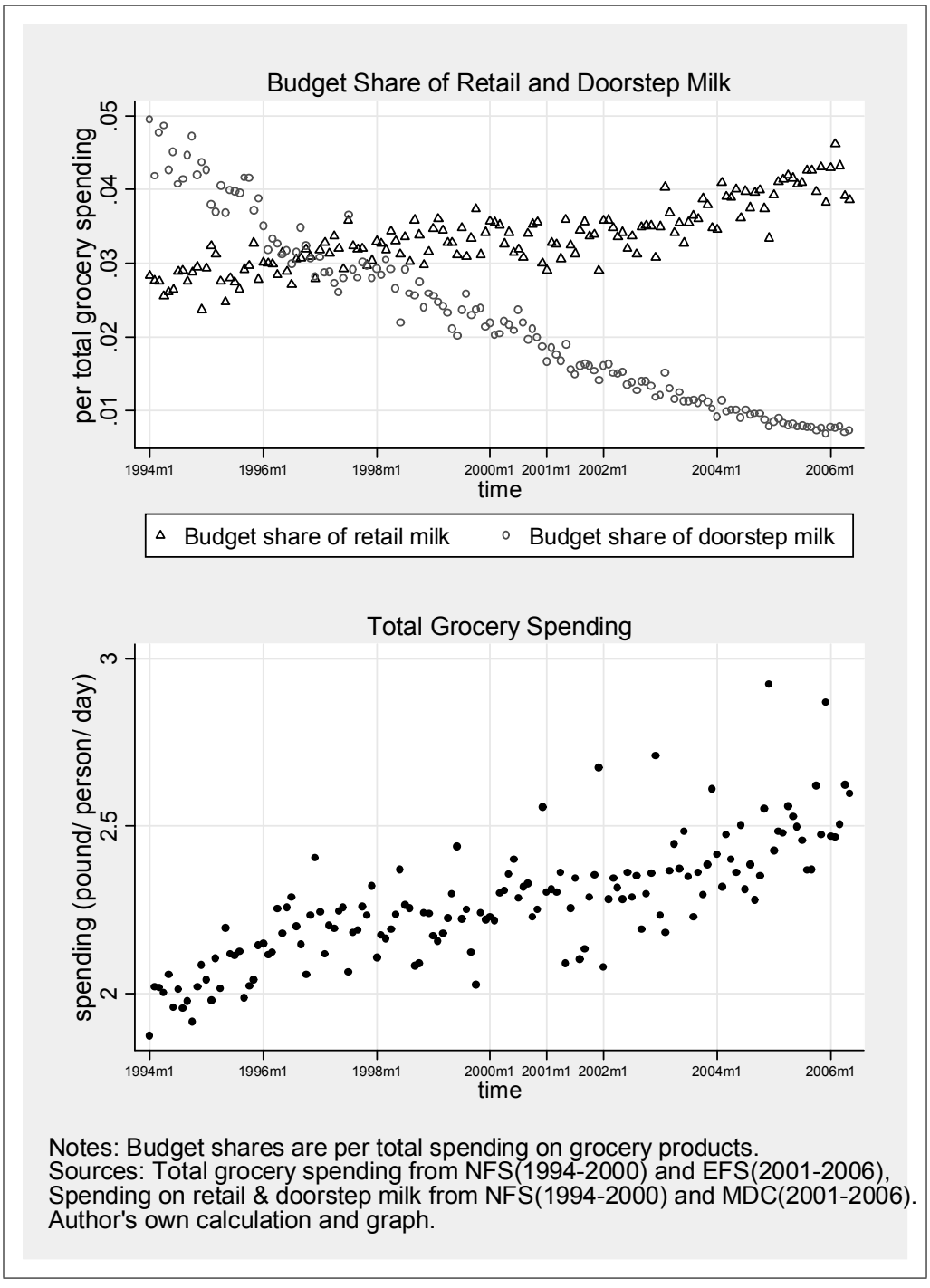


Fig. 1.4. Budget Share of Milk and Total Grocery Spending (1994-2006).

1.3.2 Estimation Issues

This section discusses estimation problems encountered when estimating the AID function (1.2). First, I discuss the classic problem of demand estimation–the endogeneity of price and expenditure. Then, I discuss the dimensionality problem, e.g. too many explanatory variables, in which I use the principal component analysis (PCA) to resolve it.

Endogeneity and Identification of Price and Expenditure

Endogeneity of price and expenditure is a classic problem often encountered in demand estimation. The endogeneity of price usually comes from two sources– 1) simultaneity of price and quantity and, 2) lack of control variables. The simultaneity problem arises when the two-way relationship between price and quantity is not accounted for. Thus, an econometric specification which writes quantity as a function of price $q_i(p_i)$ would be incorrect. In fact, price is also a function of quantity, $q_i(p_i(q_i))$. So, price is explained by quantity(or demand) shocks and would be correlated with the error term in the AID function. This fact violates the OLS assumption and would make OLS parameters inconsistent.

Another type of endogeneity arises from the lack of control variables, usually price of substitute and complementary products. If the omitted variables are correlated with price (p_i), p_i would be correlated with the error term. Hence again, the OLS assumption would be violated and the estimation would be inconsistent.

In this paper, I use the instrumental variable (IV) to remedy the potential endogeneity of price. Valid instrumental variables are those strongly correlated with milk price; while orthogonal to demand (or budget share in the case of AIDS) shocks. I obtained cost vari-

ables to serve this purpose. This is because they are likely to be correlated with price, but unlikely to be correlated with demand. From the MDC website, I obtained the monthly Intervention Milk Price Equivalent (IMPE) and, the monthly actual milk price equivalent (AMPE). I also obtained, directly from the MDC, the wholesale price of liquid milk sold by Wiseman plc. IMPE is constructed using data from several raw milk buyers. It has been used as the benchmark for raw milk price. AMPE measures the market price of butter and skim milk powder. This is, then, the opportunity cost of selling milk for liquid milk production. The Wiseman price is the average price Wiseman charged supermarkets for their bottled liquid milk. This price includes delivery to supermarket outlets. Thus, Wiseman plc's price should also be strongly correlated with the retail price set by supermarkets.

Having discussed the potential endogeneity of price, we now turn to the potential endogeneity of expenditure. Similar to the case of price, expenditure could suffer from the simultaneity problem. At the household (or individual) level, this problem seems more serious because a household's expenditure is very likely to correlate with its demand shocks. For example, when a family invites guests or organizes a big party, their demand for grocery increases. This positive demand shock obviously affects their expenditure. At the aggregate level, however, expenditure is less likely to be affected by any particular household's demand shocks. Aggregate expenditure is more likely to be affected by systematic demand shocks such as seasons, holidays, economic booms, recessions, etc. Therefore, including relevant seasonal and time dummies would suffice to remedy endogeneity of expenditure in an aggregated data[10]. To this effect, my empirical AID model includes a new year dummy

[10] Blundell et al.(1993) also found that their expenditure variable is only endogeneous at the household level. The variable becomes exogenous at the aggregate level.

for December-January, a year dummy for each year in the sample, and a quarterly dummy for each quarter of the year.

Dimensionality Problem

The dimensionality problem usually refers to two issues; 1) too many explanatory variables and, 2) multi-colinearity among explanatory variables. When either one or both of the problems occur, demand estimation would not be identified. To tackle the two problems, past research suggested many solutions. Popular remedies include the multi-stage budgeting approach (Gorman, 1959), the characteristic approach (Pinkse and Slade, 2004), the generalized composite commodity approach (GCCT) (Lewbel, 1996) and the principal component analysis (PCA) (Hoderlein and Lewbel, 2007).

The multi-stage budgeting approach suggests that consumers' decision can be depicted as a decision tree where budget allocation on big categories of consumption is decided first, i.e. food, transportation, recreation, etc. Then, conditional on the allocated budget shares, consumers decide on how much of each individual product to consume. Thus, rather than having to estimate a large demand system, the researcher only have to focus on products under the same branch of the decision tree (or the same category). This implies that products under different branches are weakly-separable. The multi-stage budgeting approach could be convenient and efficient if the consumption pattern assumed is correct. However, the more complicated the consumption pattern, the higher the chance of specifying it incorrectly.

The characteristic approach is appropriate when information on product characteristics is available. Rather than estimating all the cross-price elasticity parameters in the demand function, the parameters are separately determined as functions of product characteristics. Products with more similar characteristics would have higher cross-price elasticity and vice-versa.

The GCCT approach aggregates products together and constructs a price index of each product group. This way, the researcher could reduce the number of price regressors from the total number of individual products to the total number of product groups. One drawback of this method is that the aggregation has to satisfy certain restrictions in order for the demand system to satisfy the adding-up, homogeneity, Slutsky symmetry and negative semi-definiteness (Lewbel, 1996). One restriction is the independence between each relative price and all the price indexes, which is very difficult (if not impossible) to prove. See more in Lewbel (1996).

The PCA is similar to the GCCT in the sense that we replace individual prices with aggregated-product price indices. However, by construction of the principal component, the aggregation automatically satisfies the four properties of a rational demand system. Hoderlein and Lewbel (2007) provide a comprehensive proof of properties in the case of the almost ideal demand system. They also estimate demand for gasoline in the US as an example. This paper follows Hoderlein and Lewbel (2007) in using the PCA to remedy the dimensionality problem.

Principal Component Analysis: A Remedy for the Dimensionality Problem

The PCA approach is new compared to all other remedies for the dimensionality problem. To the best of my knowledge, the first formal work using it in demand estimation is Hoderlein and Lewbel (2007). They proved that the PCA-modified AIDS still satisfies the four slutsky properties of a rational demand system. They also showed how to estimate own- and cross- price elasticities between US gasoline and various goods of which the price had been reduced by the PCA.

PCA is a statistics method that is popular for compressing multi-dimensional data into lower dimensions. Prior to its introduction to empirical economics, PCA was applied to a wide range of applications from the computer network traffic system to face recognition. The main idea of the PCA is that it finds common characteristics of the multi-dimensional data set and reconstructs each characteristic into a new principal component. Therefore, if variables in the orginal data set are highly correlated, we would be able to use small numbers of components to account for much of the variation.

Using a more technical and formal explanation, PCA provides orthogonal linear transformations of the orginal data into a new coordinate system. Each coordinate is constructed from common variations in the original data. Thus, the most prominent coordinate accounts for the greatest variation in the original data. In the PCA, each coordinate is called a principal component. The most prominent coordinate is called the first principal component, etc. In this study, the principal components were constructed according to the following steps:

1. **Set-up**: suppose my original data set contains K variables of N observations and I would like to reduce the dimension to M variables. Thus, $M < K$.

   As mentioned in the data section, I obtained price indices of 30 $(K = 30)$ grocery product categories from the ONS website. Since my sample size is small $(N = 149$ monthly observations from January 1994 - May 2006$)$, I would like to use as few control variables as possible to account for as much variation as I can. The PCA is used here to reduce the number of dimensions from $(K = 30)$ to M (where $K > M$).

2. **Standardization of the original data**: since the PCA's scaling depends on the unit of the original data, I standardize each of the price variables into zero mean and unit standard deviation. This standardization simplifies the calculation without affecting elasticity estimates because only the change in price–not the absolute value of price–matters in the demand estimation.

3. **Covariance matrix:** calculate a covariance matrix of the standardized data. The covariance matrix is of $K \times K$ dimension.

4. **Eigenvector**: find eigenvectors and associated eigenvalues of the covariance matrix. Each eigenvector is of dimension $K \times 1$. When the covariance matrix has full rank, I obtain the maximum number of eigenvectors which is K. In a case where the covariance matrix does not have full rank, I would have less than K eigenvectors.

5. **Rank the eigenvectors**: rank the eigenvectors according to the magnitude of their eigenvalues. The first eigenvector accounts for the most variation in the data set. The second eigenvector accounts for the second most variation of the data set, etc.

In the worst case ,where variables are completely independent of each other, each eigenvector would account for $1/K$ proportion of the variation (each accounts for equal amounts of different aspects of variations). In more usual cases where variables are correlated, the first component should account for more than $1/K$ proportion of the variation. Thus, the higher the correlation among the original data, the fewer components are needed to account for a given level of variation.

6. **Choose number the of components:** pick the first M eigenvectors to reduce the data set to M dimensions (M principal components). The more eigenvectors chosen, the more variation in the original data set is accounted for.

7. **Construct principal components:** transform the original data set (K variables, N observations) to the PCA data (M principal components, N observations).

$$PCA \text{ data}_{N\times M} = \text{original data set}_{N\times K} \times \text{eigenvector}'_{K\times M}$$

Thus, the PCA transforms the original data set from K variables to M variables, which accounts for $x \in [100/K, 100]$ percent of the variation.

This study uses the PCA to transform a data set of 30 price indices (30 different grocery categories excluding liquid milk) into lower dimensions of principal components. Each of the 30 grocery category prices are calculated as follows:

$$p_{kt} = \log \frac{\rho_{kt}}{\rho_{k1}} \tag{1.4}$$

where p_{kt} is the price index of product category k at time t (deflated by its own price at time 1), $\rho_{k_j t}$ is the price index of individual product j in category k at time t.

Table (1.1) shows the proportion of variation accounted for by each of the first M components. Since it is solely up to the researcher's discretion how many components should be included, it is best to test whether altering numbers of components is robust to the results. In the result section, I will show that after a certain number of components are included, adding more components would not be robust to the result.

To show how different components represent different aspects of variations in the data, figure (1.5) plots prices of eggs, cereals, potatoes and soft drinks against time; while figure (1.6) plots the first four principal components against time.

Table 1.1. PCA of Prices of 56 Grocery Product Categories

M^{th} component	Eigen-value	% of variation	Cumulative	Mean	Std. Dev.	Min.	Max.
1	21.23	64.32	64.32	0.00	7.96	-14.24	19.50
2	3.63	10.99	75.31	0.00	2.64	-5.21	3.74
3	2.10	6.37	81.68	0.00	2.28	-4.26	7.05
4	2.00	6.07	87.75	0.00	1.48	-3.94	2.87
5	0.92	2.80	90.55	0.00	1.24	-2.94	2.93
6	0.73	2.20	92.75	0.00	1.03	-3.85	1.86
7	0.61	1.86	94.61	0.00	0.89	-2.07	2.05
8	0.35	1.05	95.66	0.00	0.80	-2.60	2.04
9	0.32	0.96	96.62	0.00	0.69	-1.80	2.30
10	0.27	0.81	97.43	0.00	0.66	-1.48	1.73
11	0.17	0.51	97.94	0.00	0.58	-1.68	1.33
12	0.15	0.46	98.40	0.00	0.52	-1.47	1.17
13	0.11	0.33	98.73	0.00	0.33	-0.94	1.00
14	0.10	0.31	99.04	0.00	0.32	-0.79	0.81
...	...	...	...	...	...	...	...
30	0.00	0.01	100.00	0.00	0.01	-0.03	0.03

Notes: the above principal components used price indices specified in equation (1.4) as inputs.
Source: raw data from ONS, Author's own calculation and table.

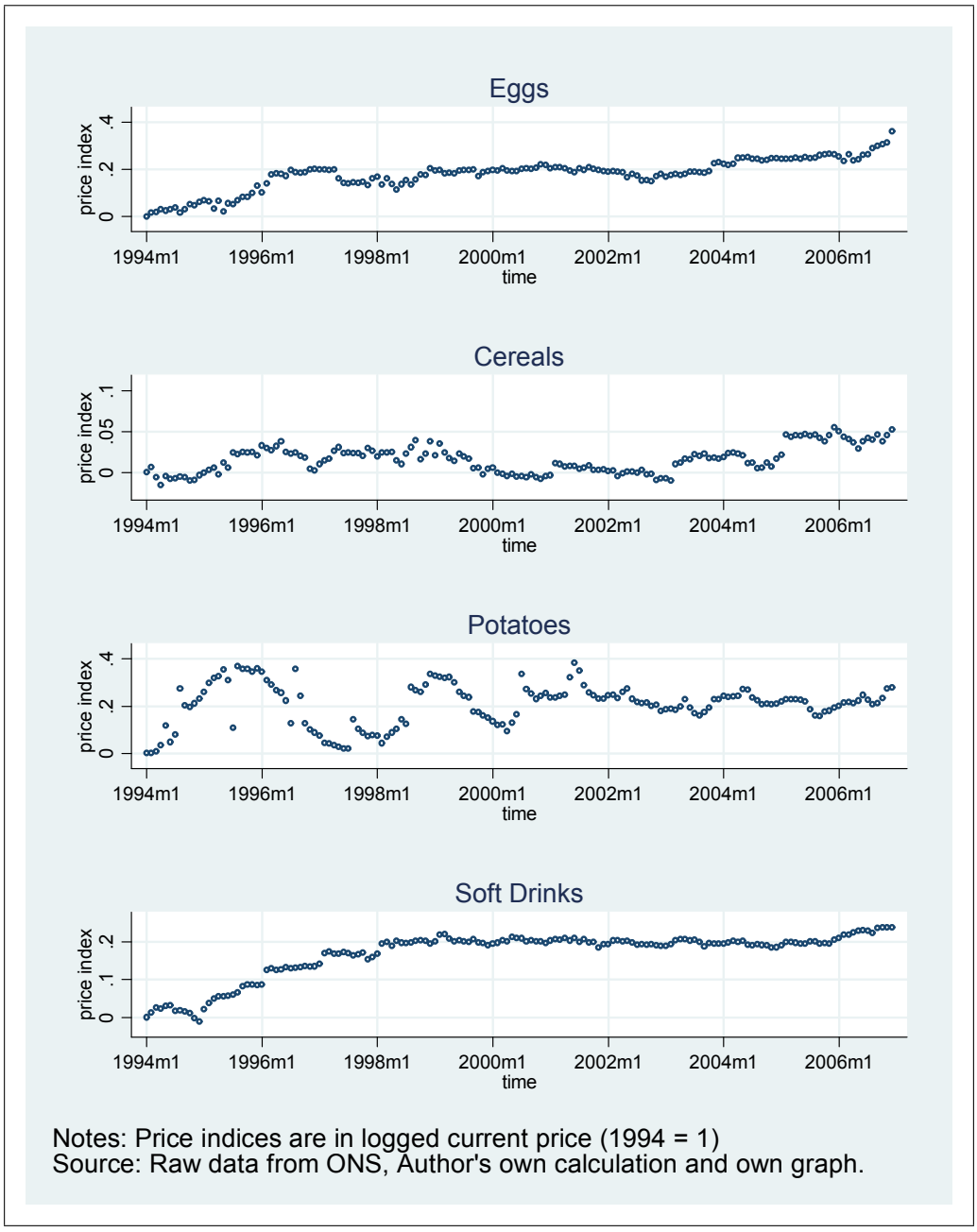


Fig. 1.5. Price Index of Different Grocery Product Categories

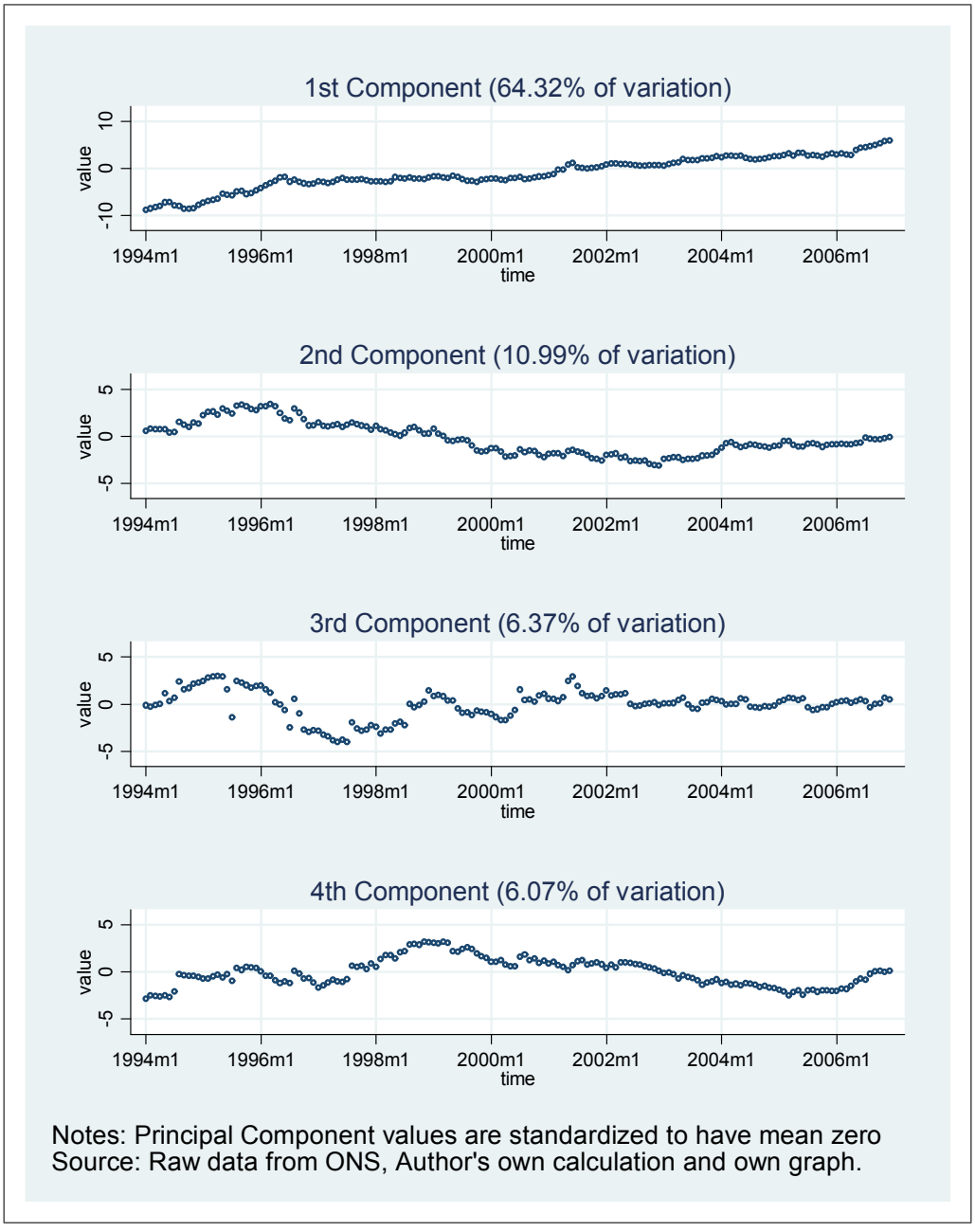


Fig. 1.6. Plots of the First Four Principal Components

1.3.3 Empirical Demand Specification

Having discussed the solutions to endogeneity and dimensionality problems, we now turn to the actual estimated demand specifications. The original AIDS equation is modified in two ways. First, I add some control variables which remedy the potential endogeneity of price and expenditure. Second, the AIDS price index P (defined in eq.1.3) is replaced by the monthly corrected Stone's price index P^{cs}. The corrected Stone's price index is one of the various linear approximations to the AIDS translog price index. This is because the translog price index imposes estimation difficulties to the demand system. The use of a linear approximation or pre-specified price indices is common in practice. In particular, Asche and Wessells (1997) show that the corrected Stone price index[11] can be used to derive consistent results in the AIDS model. The corrected stone's price index is used in Blundell et al. (1993) and Cho (2007); while the publicly available retail price index (RPI) is used in Lechene (2000).

In my case, the estimated AIDS model can be expressed as follows:

$$\begin{aligned} w_{rm,t} &= \alpha_0 + \beta_0 \log \frac{x_t}{P_t^{cs}} + \gamma_1 \log p_{rm,t} + \gamma_2 \log p_{dm,t} + \sum_{m=1}^{M} \varphi_m \widetilde{pc}_{mt} + \alpha_1 \Upsilon_t \\ &\quad + \gamma_3 \Upsilon_t * \log p_{rm,t} + \sum_n \delta_n T_n + u_t \end{aligned} \tag{1.5}$$

$$\log P_t^{cs} = \sum_k w_{kt} \log(p_{kt}) \tag{1.6}$$

where subscript t represents each monthly period, $w_{rm,t}$ is consumers' budget share of retail milk, p_t is the price of retail milk at time t normalized by p_1, x_t is total expenditure (on food and non-alcoholic drinks) at time t normalized by x_1, $\widetilde{pc}_{mt}$ is the m^{th} principal com-

[11] $\log P_t^{CS} = \sum_{k=1}^{K} w_{kt} \log \frac{p_{kt}}{p_{k1}}$,where w_{kt} is the weight given to product category $k \in (1, ..., K)$, p_{kt} is price of product category k at time t and p_{k1} is price of product category k at time 1.

ponent (which used price indices defined in (1.4) as inputs), M denotes the total number of principal components included in the specification, Υ denotes exogenous variables that may shift or rotate demand at time t, T_n denotes any time dummy for time n (e.g. year n or quarter n), and $u_{it} \sim N(0, \sigma^2)$ is the error component. Parameters to be estimated are the $\alpha, \beta, \gamma, \varphi$ and δ.

The corrected Stone price index P_t^{cs} is defined by eq.(1.6), where p_{kt} is the price of product category k normalized by its price in the first period (see eq.1.4). w_{kt} is the budget share of product category k in period t.

Since the corrected Stone price index P_t^{cs} is used instead of the original AIDS translog price index, the elasticities need to be recalculated. Appendix (1.A) shows my calculation which yields the following results:

$$\begin{aligned} \varepsilon_{rm,t} &= \frac{\partial q_{rm}}{\partial p_{rm}} \frac{p_{rm}}{q_{rm}} = \gamma_1 / w_{rm,t} + \Upsilon_t * \gamma_3 / w_{rm,t} - \beta_0 - 1 \\ \varepsilon_{rm,dm,t} &= \frac{\partial q_{rm}}{\partial p_{dm}} \frac{p_{dm}}{q_{rm}} = \gamma_2 / w_{rm,t} - \beta_0 w_{dm,t} / w_{rm,t} \\ e_{rm,t} &= \frac{\partial q_{rm}}{\partial x} \frac{x}{q_{rm}} = \beta_0 / w_{rm} + 1, \end{aligned} \tag{1.7}$$

where parameters $\gamma_1, \gamma_2, \gamma_3, \beta_0$ are as shown in equation (1.5). Their estimated values are reported in table (1.2, specification 3).

1.3.4 Demand Estimates

This section reports and discusses the estimation result of the AIDS model (eq.1.5). In addition to the AIDS' model variables, I include exogenous variables which may shift or rotate the demand curve over time. Such exogenous variables include female participation in the labour force, a new-year dummy, one-year dummies and quarterly dummies.

Thus, variable Υ_t in eq. (1.5) represents the quarterly rate of female participation in the labour force. It is also possible to expect other exogenous factors to affect consumers' preference for retail milk. I have considered household size[12], refrigerator ownership and car ownership. However, those variables varied very little during the period 1994-2006. For example, the average household size only decreased from 2.44 persons in 1994 to 2.37 persons in 2006[13], refrigerator ownership was already 98 per cent of households in 1994[14], car ownership increased from 69 percent of households in 1994 to 79 percent of households in 2006[15], but this information is available only on a yearly basis. Including yearly statistics of car ownership would be redundant to the year dummies.

Table (1.2) reports estimated coefficients from the almost ideal demand function for retail milk. The null hypotheses are rejected at the 5 percent level of significance, unless stated otherwise. All specifications include 149 observations and the adjusted-R^2 are above 0.87. Specification 1 first employs the OLS technique. This OLS specification passes the Breusch-Pagan test for heteroskedasticity (Breusch and Pagan, 1979) and the Ramsey's RESET test for omission of non-linear regressors (Ramsey, 1969). However, it does not pass the Breusch-Godfrey test (Godfrey, 1978) for serial correlation at lag = 2 and 3. The final specification, therefore, will need to account for the serial correlation in the error terms.

---

[12] I expect household size to be negatively correlated with demand for retail milk. This is because, for bigger households, there is a higher chance that at least one person would be home to bring the milk in. In addition, bigger households may demand more milk. Since milk is heavy to carry, they may prefer their milk to be delivered.

[13] According to the UK's General Household Survey 1993-1994 to 2005-2006.

[14] Data from the National Food Survey 1994 to 2006.

[15] According to the UK's General Household Survey 1993-1994 to 2005-2006.

Specification 2 remedies the potential endogeneity of $\log p_{rm}$, $\log p_{dm}$ and $\Upsilon * \log p_{rm}$ using $\log(IMPE)$, $\log(AMPE)$, $\log(Wiseman\ price)$ and $\Upsilon * \log(IMPE)$ as instrumental variables (IV). I test for the validity of instrumental variables using the Sargan test of overidentifying restriction (Sargan, 1958). This test exploits the model's excess instruments to test the null hypothesis of no correlation between the set of instruments and the residuals, i.e. the instrumental variables are valid. Since specification 2 passes the Sargan test, I cannot reject the null hypothesis that the instrumental variables are valid. In any case, the Wu-Hausman F-test of exogeneity is also performed to test the null hypothesis that $\log p_{rm}$, $\log p_{dm}$ and $\Upsilon * \log p_{rm}$ are exogenous. Since the p-value is greater than 0.05, I cannot reject the null hypothesis. Therefore, the IV estimation is not required in this case because price variables are not endogenous. In this case, the OLS coefficients would be consistent and more efficient.

Finally, specification 3 reports the OLS estimation with Newey-West robust standard errors (Newey and West, 1987) to account for the serial correlation. Since the data is on a monthly basis, I allow for serial correlation up to 12 lags. The coefficients associate with $\log p_{rm}$, $\log(x/P^r)$, $\Upsilon * \log p_{rm}$ and new-year are significant at 5 percent level. This suggests that other than the price of retail milk, total expenditure, female participation in the labour force and the new-year dummy also affect the budget share of retail milk (w_{rm}). Principal components are not statistically significant. However, if the insignificance is caused by my small sample size, excluding them could cause omitted variable bias. I decide to trade off some inefficiency by keeping the principal components.

Table 1.2. AIDS Model Estimates

Dependent Variable = w_{rm}			
Explanatory Variable	(1-OLS)	(2-IV)	(3-OLS,Newey)
constant	0.12* (0.07)	0.16 (0.16)	0.12* (0.06)
$\log p_{rm}$	-0.53**(0.26)	-0.74 (0.71)	-0.53**(0.25)
$\log p_{dm}$	0.03 (0.02)	0.10* (0.06)	0.03 (0.02)
$\log (x/P^r)$	-0.03**(0.00)	-0.03**(0.00)	-0.03**(0.00)
$\Upsilon * \log p_{rm}$	1.07**(0.50)	1.47 (1.38)	1.07**(0.48)
Υ	-0.18 (0.13)	-0.25 (0.31)	-0.18 (0.12)
principal components			
$\widetilde{pc}_1$	0.03 (0.51)	-0.23 (0.97)	0.03 (0.54)
$\widetilde{pc}_2$	0.48 (0.35)	0.67 (0.42)	0.48 (0.27)
$\widetilde{pc}_3$	-0.06 (0.18)	-0.05 (0.23)	-0.06 (0.17)
$\widetilde{pc}_4$	-0.10 (0.33)	0.16 (0.39)	-0.10 (0.19)
$\widetilde{pc}_5$	-0.06 (0.29)	-0.15 (0.38)	-0.06 (0.24)
$\widetilde{pc}_6$	0.13 (0.33)	0.27 (0.45)	0.13 (0.37)
$\widetilde{pc}_7$	0.06 (0.39)	0.45 (0.68)	0.06 (0.39)
$\widetilde{pc}_8$	-0.14 (0.41)	0.15 (0.48)	-0.14 (0.58)
time dummies $(\times 10^3)$			
new year	-1.52**(0.50)	-1.21**(0.55)	-1.52**(0..65)
year dummy	yes	yes	yes
quarter dummy	yes	yes	yes
observations	149	149	149
adjusted-R^2	0.90	0.88	0.90
Specification tests (p-values)			
Breusch-Pagan	0.15	–	–
Ramsey (RESET)	0.54	–	–
Breusch-Godfrey, lag1	0.71	–	–
Breusch-Godfrey, lag2	0.09	–	–
Breusch-Godfrey, lag3	0.01	–	–
Sargan	–	0.51	–
Wu-Hausman	–	0.25	–

Notes: ** significant at 5%, * significant at 10%.
The Breusch-Pagan test tests the null hypothesis of no heteroskedasticity, i.e. the error terms are normally and independently distributed with a constant variance. The Ramsey (RESET) test tests the null hypothesis of no omitted variables, specifically, higher powers of the regressors. The Breusch-Godfrey test tests the null hypothesis of no higher-order serial correlation in the error terms. In my case, I examined up to lag = 3. The Sagan test tests the null hypothesis of valid instruments. The Wu-Hausman test tests the null hypothesis of consistent OLS estimates counterpart.

Having discussed the AIDS model estimation results, we now turn to the corresponding elasticities. Table (1.3) reports the AIDS own-price elasticity of demand for retail milk (ε_{rm}), cross-price elasticity with demand for doorstep milk ($\varepsilon_{dm,rm}$) and income elasticity of demand for retail milk (e_{rm}). Over the period 1994-2006, the budget share of retail milk (w_{rm}), and female participation in the labour force (Υ) gradually increased. This coincided with the decrease in consumers' price elasticity of demand for retail milk. Demand for retail milk was the most elastic in 1994 where $\varepsilon_{rm} = -1.288$; and was becoming less and less elastic to $\varepsilon_{rm} = -0.335$ in 2006. Each elasticity has a standard error because it was calculated using *estimated* parameters from the AID function (1.5).

In contrast to retail milk, doorstep milk was becoming more elastic over time. My AIDS estimates suggest that the own-price elasticity of demand for doorstep milk (ε_{dm}) was $-1.792, -2.487, -4.576$ in 1994, 2000, and 2006 respectively. In addition, the expenditure elasticity of doorstep milk (e_{dm}) was decreasing from 0.689 in 1994 to 0.340 and -0.847 in 2000 and 2006 respectively. Full estimation results are available upon request.

The directions of change in the price elasticity and in the expenditure elasticity of doorstep milk were the opposite to those of retail milk. Demand for retail milk was becoming less sensitive to price, while demand for doorstep milk was becoming more sensitive to price. This could be one factor supporting the growth of demand for retail milk, and the increase of supermarkets' market power, over time.

Table 1.3. AIDS own-price, cross-price and expenditure elasticity

Year	Variables			Elasticities (std. err.)		
	w_{rm} (retail share)	w_{dm} (doorstep share)	Υ (female work)	ε_{rm} (own-price)	$\varepsilon_{dm,rm}$ (cross-price)	e_{rm} (expenditure)
1994	0.027	0.045	0.500	-1.03**(0.32)	1.11(0.86)	-0.15(0.13)
1995	0.028	0.039	0.505	-0.85**(0.23)	1.05(0.82)	-0.09(0.12)
1996	0.030	0.032	0.502	-0.96**(0.27)	1.01(0.79)	-0.06(0.12)
1997	0.032	0.029	0.503	-0.91**(0.23)	0.94(0.74)	0.01(0.11)
1998	0.033	0.027	0.508	-0.76**(0.18)	0.92(0.73)	0.03(0.11)
1999	0.034	0.023	0.513	-0.63**(0.15)	0.89(0.70)	0.06(0.10)
2000	0.034	0.021	0.513	-0.62**(0.15)	0.89(0.70)	0.06(0.10)
2001	0.030	0.015	0.516	-0.49**(0.16)	0.98(0.78)	-0.05(0.11)
2002	0.031	0.013	0.521	-0.35* (0.19)	0.94(0.75)	-0.00(0.11)
2003	0.034	0.011	0.520	-0.42**(0.17)	0.88(0.70)	0.06(0.10)
2004	0.035	0.009	0.520	-0.45**(0.16)	0.83(0.67)	0.11(0.10)
2005	0.038	0.007	0.525	-0.35* (0.19)	0.77(0.62)	0.18*(0.09)
2006	0.039	0.007	0.525	-0.36* (0.19)	0.75(0.60)	0.19**(0.09)

Notes: ** significant at 5%, * significant at 10%.
Elasticity calculations use parameter estimates from table(1.2) specification 3.

Robustness Check

Obtaining accurate estimates of the demand elasticity (ε_{rm}) is important to the analysis of market power and firms' collusion. As will be discussed further in section (1.4.1: The Conduct Parameter Approach), inelastic demand, high market concentration and collusion are the three sources of market power–the power to earn a positive price-cost margin. If the demand elasticity is inaccurately estimated, the analysis of market power and collusion would also be biased.

The chosen specification (specification 3-OLS with Newey standard errors) suggests that demand elasticity was falling from -1.03 in 1994 to -0.36 in 2006. I employ some robustness tests to investigate whether changing the specification could alter the magnitude and direction of the demand elasticity adjustment. The first test concerns the use of $\Upsilon * \log p_{rm}$−female participation in the labour force interacts with price. This variable is the

only interaction term I include in the chosen specification. Therefore, other than the change in the budget share of retail milk (w_{rm}), $\Upsilon * \log p_{rm}$ provides the only source of elasticity adjustment through time. To see this more clearly, consider the AIDS elasticity eq.1.7:

$$\varepsilon_{rm,t} = \gamma_1 / w_{rm,t} - \Upsilon_t * \gamma_3 / w_{rm,t} - \beta_0 - 1.$$

Here, only the budget share of retail milk $(w_{rm,t})$, and the time-varying exogenous variable Υ_t determine how the elasticity adjusts through time. Although it is possible that Υ_t contains many exogenous variables, my chosen specification uses only female participation in the labour force. I have considered other exogenous variables such as household size, refrigerator ownership and car ownership. However, as mentioned in section 3.5, those variables vary very little and are only observed on a yearly basis. Their effects, thus, are controlled by using yearly dummies.

To this effect, the first robustness test examines the use of the time trend as an alternative to female participation in the labour force. I add to the chosen specification the time trend (T) and time trend interaction with price $(T * \log p_{rm})$. This addition makes the coefficient associates with $\Upsilon * \log p_{rm}$ become statistically insignificant. However, this is not so surprising given that female participating in the labour force (Υ) varies every quarter, while the time trend (T) varies every period (every month) so it can explain more variation. This robustness test does not change the direction of the elasticity adjustment. Demand elasticity still becomes more inelastic over time. However, the magnitude of the elasticity changes. Compared with the chosen specification, demand is more elastic in 1994 and becomes more inelastic at a quicker rate. By 2006, this specification predicts demand to be perfectly inelastic. That is $\varepsilon_{rm} \approx 0$. Although the time trend is a possible candidate

for the exogenous demand shifter and rotator (Υ), one problem in using it is that the variable increases linearly every period. Thus, by including the time trend in the specification, one assumes that price elasticity changes linearly with time. Since this is unlikely to be the case, I decide not to include the time trend and time trend interactions in my specifications.

The next robustness test examines the role of the principal components. I first remove all the principal components, meaning that prices of other grocery products are not controlled for. This gives lower estimates of elasticity (in the absolute term) than the chosen specification. Such bias is expected, given that most of the other grocery products are substitutes for retail liquid milk.

The last robustness test increases the number of principal components from eight to ten. This addition does not affect the elasticity estimates from 1994-2000. However, it gives slightly higher elasticity estimates from 2001-2006.

In sum, the robustness tests show that the direction of elasticity adjustment does not change but the magnitude does. Therefore, the qualitative finding which suggests that the retailers were gaining higher market power from declining demand elasticity still holds. On the other hand, the magnitude of demand elasticity changes significantly when the time trend is used instead of female participation in the labour force. However, since the time trend increases linearly over time, I do not believe it is a more reasonable alternative. I also find that the magnitude of demand elasticity changes slightly when more principal components are added. This change becomes very small when the first eight components have already been included in the specification. I include eight components in my chosen

specification and expect the chosen specification to reasonably approximate to the true magnitude of the demand elasticity.

Table 1.4. Robustness Check for Demand Estimation

Variable	Chosen Specification	Add Time Trend	Only Time Trend	No PC^{a}s	More PCs
constant	0.12* (0.06)	-0.03 (0.08)	0.02**(0.01)	0.12**(0.03)	0.10* (0.06)
$\log p_{rm}$	-0.53**(0.25)	0.23 (0.34)	-0.01**(0.01)	-0.53**(0.19)	-0.44* (0.27)
$\log p_{dm}$	0.03 (0.02)	0.03 (0.02)	0.03 (0.02)	0.02 (0.02)	0.03 (0.02)
$\log (x/P^r)$	-0.03**(0.00)	-0.03**(0.00)	-0.03**(0.00)	-0.03**(0.00)	-0.03**(0.00)
$\Upsilon * \log p_{rm}$	1.07**(0.48)	-0.48 (0.68)	-	1.07**(0.37)	0.88* (0.53)
Υ	-0.18 (0.12)	0.10 (0.16)	-	-0.18**(0.07)	-0.14 (0.12)
$T * \log p_{rm}$	-	0.46**(0.13)	0.39**(0.07)	-	-
T	-	-0.02 (0.08)	0.00 (0.08)	-	-
Other control variables					
PC$_s$(number)	yes(8)	yes(8)	yes(8)	no	yes(10)
new year dummy	yes	yes	yes	yes	yes
quarter dummy	yes	yes	yes	yes	yes
year dummy	yes	yes	yes	yes	yes
Corresponding annual-average own-price elasticity (ε_{rm})					
1994	-1.03**(0.32)	-1.34**(0.23)	-1.41**(0.21)	-0.87**(0.27)	-1.02**(0.36)
1995	-0.85**(0.23)	-1.21**(0.18)	-1.22**(0.18)	-0.69**(0.21)	-0.86**(0.26)
1996	-0.96**(0.27)	-0.97**(0.20)	-1.06**(0.15)	-0.81**(0.23)	-0.96**(0.30)
1997	-0.91**(0.23)	-0.82**(0.19)	-0.91**(0.13)	-0.77**(0.20)	-0.92**(0.26)
1998	-0.76**(0.18)	-0.72**(0.14)	-0.77**(0.12)	-0.61**(0.16)	-0.78**(0.20)
1999	-0.63**(0.15)	-0.63**(0.11)	-0.64**(0.11)	-0.48**(0.13)	-0.67**(0.16)
2000	-0.62**(0.15)	-0.46**(0.12)	-0.50**(0.12)	-0.47**(0.13)	-0.66**(0.15)
2001	-0.49**(0.16)	-0.32**(0.13)	-0.34**(0.13)	-0.37**(0.13)	-0.58**(0.15)
2002	-0.35* (0.19)	-0.25* (0.14)	-0.23* (0.13)	-0.24* (0.14)	-0.47**(0.16)
2003	-0.42**(0.17)	-0.13 (0.14)	-0.15 (0.14)	-0.31**(0.12)	-0.52**(0.15)
2004	-0.45**(0.16)	-0.02 (0.16)	-0.07 (0.15)	-0.35**(0.12)	-0.55**(0.14)
2005	-0.35* (0.19)	-0.02 (0.15)	-0.02 (0.15)	-0.25* (0.13)	-0.46**(0.17)
2006	-0.36* (0.19)	0.09 (0.17)	0.07 (0.16)	-0.26**(0.13)	-0.47**(0.17)
observations	149	149	149	149	149
adjusted-R^2	0.90	0.92	0.92	0.91	0.91

Notes: all specifications are estimated by OLS using Newey-West standard errors to allow for serial correlation up to 12 lags. The "Chosen Specification" is specification(3-OLS, Newey) in table(1.2).
^{a} : PCs stands for principal components, ** significant at 5%, * significant at 10%.

1.4 Supply

Having estimated the demand for each year, we now turn to analyse supply. The conduct parameter approach will be used to study changes on the supply side, e.g. the retailers' competitiveness. Brenahan (1989) comprehensively discussed the conduct parameter approach as a means to estimate firms' marginal cost and conduct simultaneously. The technique is also known as the New Empirical Industrial Organization (NEIO). In 1999, Corts (1999) provided proof that the NEIO technique could give inconsistent marginal cost and conduct estimates if firms do not play a conjectural variations game. In the case where marginal cost is observed, however, conduct estimates would still be reliable. Another critique of the conduct parameter is that by Salvo (2004). Using the case of the Brazilian Cement Industry, the paper suggests that the conduct parameter understates collusion if the threat of entry by foreign suppliers is not accounted for.

This paper overcomes the Corts Critique by obtaining the direct approximation of retailers' marginal cost. As for the Salvo Critique on the threat of entry by foreign competitors, I believe that this case does not apply at the retailer level[16]. This is because fresh liquid milk would be very costly, if not impossible, to be imported by the households themselves.

[16] Salvo Critique is more relevant at the manufacturer level. Given that the retailers are profit-maximizing, they would import fresh processed milk from abroad whenever it is cheaper to do so. Here, if the price of imported milk is higher than the monopoly price of manufacturers, the threat of import has to be accounted for.

1.4.1 The Model: Conduct Parameter Approach (θ)

The conduct parameter approach was built upon the static conjectural variations game. In such a game, players form their beliefs of other players' actions. Then, given these beliefs, they choose the best response. If the researcher observes the market outcome–particularly the equilibrium price and quantity–they can derive the corresponding conduct parameter θ. Different values of the conduct parameter would suggest different levels of collusiveness.

Although based on the same idea, the value of θ can be defined slightly differently (see for example, Genesove and Mullin (1998), Villas-Boas and Hellerstein (2006), and Puller (2007)). In what follows, I show my derivation of the conduct parameter and discuss its definition for use in this paper. First of all, I assume that firms, the supermarkets in this context, are profit-maximizing and use quantity as the strategic variable. This results in the following optimality condition:

$$\arg\max_{q_{it}} \Pi_{it} = (P_t(Q_t) - MC_{it})q_{it}$$

where subscript i denotes supermarket and t denotes time, Π_{it} denotes supermarket i's profit at time t, P_t denotes market price of the product (which is a function of total market quantity Q_t), MC_{it} denotes marginal cost of supermarket i at time t and q_{it} denotes quantity supplied by supermarket i at time t. Solving the first order condition gives:

$$\begin{aligned}
0 &= P_t + (q_{it}\frac{\partial P_t}{\partial Q_t}\frac{t\partial Q_t}{\partial q_{it}}) - MC_{it} \\
\frac{P_t - MC_{it}}{P_t} &= -q_{it} \times \frac{\partial P_t}{\partial Q_t} \times \frac{\partial Q_t}{\partial q_{it}} \times \frac{1}{P_t} \times \frac{Q_t}{Q_t} \\
\frac{P_t - MC_{it}}{P_t} &= \underbrace{-\frac{\partial P_t}{\partial Q_t} \times \frac{Q_t}{P_t}}_{-1/\varepsilon_t} \times \frac{\partial Q_t}{\partial q_{it}} \times \underbrace{\frac{q_{it}}{Q_t}}_{s_{it}}
\end{aligned}$$

where ε_t is demand elasticity at time t, s_{it} is supermarket i's market share at time t. The term on the left-hand side is supermarket i's Lerner's index. Here, the weighted-average Lerner's index of the industry can be calculated as follows:

$$\begin{aligned} \sum_i s_{it} \frac{P_t - MC_{it}}{P_t} &= -\sum_i s_{it} \times \frac{1}{\varepsilon_t} \times \frac{\partial Q_t}{\partial q_{it}} \times s_{it} \\ L_t &= -\underbrace{\sum_i s_{it}^2}_{H_t} \times \frac{1}{\varepsilon_t} \times \underbrace{\frac{\partial Q_t}{\partial q_{it}}}_{\theta_t} \\ L_t &= -\frac{H_t \theta_t}{\varepsilon_t} \end{aligned} \tag{1.8}$$

where H_t is the Herfindahl index of concentration at time t, L_t is the industry's Lerner's index at time t, and θ_t is the conduct parameter[17] at time t. The conduct parameter $\theta_t = \partial Q_t / \partial q_{it}$ measures how much output of one supermarket matters to the total output of the industry. When the output of one supermarket does not affect the total output of the industry, θ gets close to zero. Thus, the conduct implies that supermarkets are close to perfect competition. When the output of one supermarket increases the total output of the industry by the same unit, θ equals 1. The same would result when supermarkets play a noncollusive Cournot game. Finally, when supermarkets are perfectly collusive, they would promptly increase or decrease their quantity according to their market share. An x percent increase in q_{it} would result in an x percent increase in the industry's quantity. Thus,

$$1 = \frac{\partial Q_t}{\partial q_{it}} \times \frac{q_{it}}{Q_t} \tag{1.9}$$

[17] Note here that the conduct parameter θ in this paper is similar but not identical to the one in Genesove and Mullin(1998) and Cort(1999). Although using the same logic to derive, our θ ranges from 0 to $1/\Sigma_i s_i^2 = 1/H$ whereas θ in the above two papers ranges from 0 to 1. Our version of θ allows the elasticity-adjusted Lerner's index (L) to be expressed in terms of demand elasticity ε , the Herfindahl index H and the conduct parameter θ. See equation(1.8).

for perfectly collusive supermarkets. I can use the above condition to find the upper bound of conduct value–the one from perfectly collusive supermarkets θ^{mono}. The Lerner's index in a perfectly collusive industry would be:

$$\begin{aligned} L_t &= -\sum_i s_{it}^2 \times \frac{1}{\varepsilon_t} \times \frac{\partial Q_t}{\partial q_{it}} \\ &= -\underbrace{\sum_i s_{it}}_{1} \times \frac{1}{\varepsilon_t} \times \underbrace{\frac{\partial Q_t}{\partial q_{it}} \times \frac{q_{it}}{Q_t}}_{1}. \\ &= -\frac{1}{\varepsilon_t} \end{aligned} \tag{1.10}$$

Since $L_t = -\frac{H\theta}{\varepsilon_t}$, we have $\theta^{mono} = 1/H$ when supermarkets are perfectly collusive. Therefore, the values of θ range from 0 (perfectly competitive) to $1/H$ (perfectly collusive).

Having defined our conduct parameter, the following sections discuss the data sources and estimation results.

1.4.2 Data on Marginal Costs

Given the price elasticity, the conduct parameter can be obtained through two means. First, if the marginal cost information is not available, one can follow the NEIO approach to simultaneously estimate marginal cost and conduct. Second, if the marginal cost information is available, one can use it to directly calculate the conduct parameter. Previous literature, however, has extensively discussed the potential bias from estimating marginal cost and conduct parameter simultaneously. In particular, Corts (1999) shows that, if firms do not adopt a conjectural variations strategy, the marginal cost would be inconsistently esti-

mated. As a result, the corresponding conduct would be incorrect. To avoid this problem, the actual marginal cost or the direct approximation of marginal cost should be used.

I obtained information from various sources to directly approximate retailers' marginal cost. According to Smith and Thanassoulis (2008), supermarkets' marginal cost of liquid milk consists of two main components–1) cost of raw milk (p_i^w) and 2) supermarkets' operational cost (c_i), e.g:

$$MC_{it} = p_{it}^w + c_{it}.$$

where subscript i denotes supermarket, subscript t denotes time and MC_{it} denotes the marginal cost of supermarket i at time t.

The cost of raw milk (p_{it}^w) is proxied by the monthly-average price of processed liquid milk, inclusive of delivery, charged by Wiseman. This data is available on request to the MDC. Price charged by other companies are not available. However, I believe it is reasonable to use Wiseman's price as an industry average. This is because all major milk processors did not differ in their characteristics so much that it would seem reasonable they could not charge significantly different prices. In 1999, only four milk processors held as much as 79.2 percent market share–27 percent by Express, 25.1 percent by Dairy Crest, 13.8 percent by Wiseman, and 13.3 percent by Arla (CC, 2000).

Other than for raw milk, supermarkets have to pay the operational costs (c_{it}). This is largely the cost for shelving (labour) and storage (energy). I was able to obtain sufficient information about the four biggest firms in 2004 (Asda, Morrisons, Sainsbury, Tesco); and Safeway in 2003. The labour cost is calculated by first finding the total number of full-time-equivalent employees in each firm. Multiply this by the average earnings per hour in

the retail sector for 8 hours per day will give the total cost of labour for each firm. To find the contribution of this cost by each litre of liquid milk, I multiply the total labour cost with the proportion of milk items (or packs) to all grocery items in store. Then, I divide by the total volume (in litre) of milk sold.

The number of full-time-equivalent employees was obtained from each company's annual reports. The average earnings per hour in the retail sector was obtained from the ONS website. The proportion of milk items to all grocery items, and the total volume of liquid milk sold by each supermarket firm is obtained from the TNS homescan data set.

The calculation of energy cost is guided by the DEFRA report on the impact of UK liquid milk consumption (DEFRA, 2007). According to the report, an average supermarket uses 37 percent of their electricity for refrigeration; while liquid milk occupies about 5 percent of the refrigerated shelves. Thus, liquid milk accounts for about 1.85 percent of a store's electricity costs. I obtained the average per-outlet electricity cost from a survey conducted by the Green Liberal Democrats, an environmental campaign group of the Liberal Democrats party (UK) (Baker, 2004). Although the report disguises the identity of each supermarket, I was able to identify them through the annual turnovers, number of employees, number of stores and other information given in the report.

Since the approximated operational costs are for 2004 only, I inflate and deflate them as appropriate for all other years. I use the average earnings index to inflate and deflate the labour cost, and use the average of gas and electricity price index to inflate and deflate the energy cost. All these indices are available on the ONS website.

Table (1.5) reports the marginal cost components of fresh processed liquid milk. All figures are in pence per litre.

Table 1.5. Supermarket's Operational Costs by Year

Firm	Asda		Morrisons		Safeway		Sainsbury's		Tesco		Firms' avg
Year	Labour	Ener.	Labour	Ener.	Labour	Ener.	Labour	Ener.	Labour	Ener.	Wiseman price
1994	0.63	0.0003	0.63	0.0002	1.10	0.0005	1.11	0.0005	0.70	0.0003	41.43
1995	0.65	0.0003	0.66	0.0002	1.14	0.0005	1.15	0.0005	0.73	0.0003	42.11
1996	0.67	0.0002	0.67	0.0002	1.18	0.0004	1.18	0.0004	0.75	0.0002	41.56
1997	0.70	0.0002	0.70	0.0002	1.23	0.0004	1.23	0.0004	0.78	0.0002	38.79
1998	0.74	0.0002	0.74	0.0002	1.29	0.0004	1.29	0.0004	0.82	0.0002	36.02
1999	0.77	0.0002	0.77	0.0002	1.35	0.0004	1.35	0.0004	0.86	0.0002	34.24
2000	0.82	0.0002	0.82	0.0002	1.44	0.0004	1.44	0.0004	0.92	0.0002	32.46
2001	0.86	0.0003	0.86	0.0002	1.50	0.0005	1.50	0.0005	0.96	0.0003	35.21
2002	0.88	0.0003	0.88	0.0003	1.54	0.0005	1.55	0.0005	0.98	0.0003	33.79
2003	0.91	0.0003	0.91	0.0002	1.59	0.0005	1.60	0.0005	1.02	0.0003	34.65
2004	0.95	0.0003	0.95	0.0003	1.66	0.0006	1.67	0.0006	1.06	0.0003	35.62
2005	0.98	0.0003	0.98	0.0005	1.72	0.0010	1.72	0.0010	1.09	0.0006	36.26
2006	1.03	0.0005	1.03	0.0004	1.80	0.0009	1.80	0.0009	1.15	0.0005	37.00

Notes: all figures are in pence per litre. Labour and electricity costs are from the author's own approximation. Labour = cost of labour, Ener. = cost of electricity and gas, Wiseman Price = mean price of fresh processed milk inclusive of delivery (to store's doorstep) charged by Wiseman plc. I assume that this is the mean across all retailers.

1.5 Results and Discussion

1.5.1 Conduct and Structure

This section discusses the estimation results–that is, how the market concentration, demand elasticity and conduct parameter changed over time. To evaluate the market concentration, table (1.6) reports the market share of each national-level supermarket chain in the UK for the period 1994-2006. The last column of the table shows the Herfindahl index ($H = \Sigma_i s_i^2$, where s_i denotes the market share of chain i). I obtained this market share information

from an IGD report, which was constructed using the TNS homescan data base. Over time, the big firms gainned more market share and caused the Herfindahl index (H) to rise from 0.048 in 1994 to 0.113 in 2006. The values of H observed, however, were still significantly lower than the monopoly's H, which equals 1.

The adjustment of the Herfindahl index is then compared with the adjustments of other market indicators. Table (1.7) reports retail price, retailers' marginal costs, demand elasticity, price-cost margin, Herfindahl index, the conduct parameter under perfectly collusive retailers $(\theta^{mono} = 1/H)$, and the annual-average values of the conduct parameter (θ). Over time, both the average retail price and the marginal cost steadily increased. However, since the average retail price increased at a faster rate than the marginal costs, the retailers' price-cost margin increased. Here, if the market power is defined by firms' ability to earn positive profits, the retailers clearly enjoyed increasing market power during 1994-2006.

One contribution to the increase in market power is the aforementioned increase in market concentration. Another contribution is the decline in consumers' price sensitivity. The estimated own-price elasticity of demand for retail milk (ε_{rm}) became less sensitive, from ε_{rm} = -1.03 in 1994 to ε_{rm} = -0.36 in 2006. Unless firms are perfectly competitive, inelastic demand gives firms more power to markup their price. The next section will examine how this reflects into the movement of the demand curve.

In contrast to market concentration and demand elasticity, I find that the conduct parameter did not always change in favour of higher market power. The value of θ implied higher market power when it increased steadily from -0.184 in 1994 to 2.003 in 2001. After 2001, the value dropped and remained around 1, or even fell below 1. Hence, the increase

in market power can partly be explained by more collusive conduct from 1994-2001. After that, the increase in market power can only be explained by higher market concentration and declining price elasticity alone.

Using the as-if conjectural variations benchmarks, we can say that firms were achieving similar price-cost margins to those of the static perfect-Bertrand competition from 1994-1996. They earned higher than the noncooperative Cournot profits from 1997-2000. However, due to higher competition intensity, firms' profitability dropped to about the same level as that of the static noncooperative Cournot competition from 2001-2006.

There is some evidence of below-marginal-cost pricing from 1994-1995, where the average retail prices were lower than the marginal costs. This could have been caused by many practices. For example, the retailers' use of predatory pricing against the milkman, or an adoption of the lost-leader strategy which uses cheap liquid milk to attract consumers to retail stores. In any case, examining the motives of this below-marginal cost pricing is beyond the focus of this study.

1.5.2 Evolution of the Demand Curve (for Retail Milk)

I have discussed in the previous section that lower price elasticity of demand for retial milk explains why the retailers were able to gain higher margins under low values of conduct. In this section, I examine the movement of the demand curve that resulted in such change. Three things could explain a decrease in the price elasticity 1) an increase in the willingness to pay, 2) an increase demand, and 3) an increase in both the willingness to pay and demand. When demand increases, i.e. there are simply more consumers to sell the product for, the

Table 1.6. Market Share of National-level Supermarkets 1994-2006

Year	Tesco	Sains.	Asda	Safe.	Some.	Kwik.	M&S	Morr.	Waitrose	Iceland	Wm Low	Co-ops	Total	H
1994	0.11	0.12	0.07	0.08	0.04	0.04	0.03	0.02	0.02	0.02	0.01	0.07	0.56	0.048
1995	0.13	0.12	0.07	0.07	0.04	0.04	0.03	0.02	0.02	0.02	0.00	0.07	0.57	0.053
1996	0.14	0.12	0.08	0.08	0.04	0.04	0.03	0.03	0.02	0.02	0.00	0.06	0.59	0.056
1997	0.15	0.12	0.08	0.08	0.04	0.04	0.03	0.03	0.02	0.02	0.00	0.06	0.60	0.058
1998	0.15	0.12	0.09	0.08	0.07	0.00	0.03	0.03	0.02	0.02	0.00	0.06	0.60	0.062
1999	0.16	0.12	0.09	0.08	0.07	0.00	0.03	0.03	0.02	0.02	0.00	0.06	0.60	0.063
2000	0.16	0.12	0.10	0.08	0.01	0.00	0.03	0.03	0.02	0.02	0.00	0.06	0.61	0.060
2001	0.19	0.13	0.11	0.08	0.03	0.00	0.03	0.01	0.02	0.02	0.00	0.03	0.63	0.073
2002	0.20	0.12	0.12	0.08	0.03	0.02	0.03	0.01	0.02	0.02	0.00	0.03	0.66	0.078
2003	0.22	0.12	0.12	0.08	0.03	0.02	0.03	0.01	0.03	0.02	0.00	0.03	0.68	0.086
2004	0.23	0.12	0.13	–	0.03	0.02	0.03	0.10	0.03	0.01	0.00	0.03	0.73	0.098
2005	0.25	0.12	0.13	–	0.03	0.01	0.03	0.09	0.03	0.01	0.00	0.03	0.74	0.104
2006	0.26	0.13	0.13	–	0.04	0.01	0.04	0.09	0.03	0.01	0.00	0.03	0.76	0.113

Source: Institute of Grocery & Distribution (IGD).

Notes: Sains = Sainsbury's, Safe. = Safeway, Kwik. = KwikSave, Morr. = Morrisons. H is the Herfindahl index, H = $\Sigma_i s_i^2$ where s_i denotes market share of store i. Total = the sum of all national supermarkets' share. I assume that market share of all other stores are small, and thus, have little effect on the value of H. Safeway was acquired by Morrisons in 2004. So, its market share is not available.

Table 1.7. Value of the Conduct Parameter and Other Market Indicators by Year

Year	P	MC	ε_{rm}	$\frac{P-MC}{P}$	H	$\theta^{mono} = \frac{1}{H}$	$\theta = \frac{P-C}{P} \times \frac{-\varepsilon_{rm}}{H}$
1994	41.95	42.32	-1.03	-0.009	0.048	20.70	-0.184
1995	42.35	43.01	-0.85	-0.016	0.053	18.81	-0.250
1996	43.47	42.48	-0.96	0.023	0.056	17.82	0.386
1997	43.11	39.75	-0.91	0.078	0.058	17.25	1.223
1998	42.69	37.02	-0.76	0.133	0.062	16.21	1.636
1999	42.51	35.28	-0.63	0.179	0.063	15.95	1.707
2000	41.59	33.56	-0.62	0.200	0.060	16.73	2.003
2001	43.41	36.36	-0.49	0.162	0.073	13.71	1.090
2002	44.90	34.97	-0.35	0.221	0.078	12.80	0.991
2003	47.38	35.85	-0.42	0.243	0.086	11.59	1.184
2004	48.48	36.77	-0.45	0.242	0.098	10.19	1.109
2005	51.60	37.44	-0.35	0.274	0.104	9.59	0.921
2006	51.74	38.24	-0.36	0.261	0.113	8.83	0.829

Notes: P is the mean price in pence per litre of retail milk, MC is the retailers' marginal cost, ε_{rm} is the own-price elasticity estimate from table(1.3), $\frac{P-MC}{P}$ is the price-cost margin ,$H = \Sigma_i s_i^2$ is the Herfindahl index of market concentration,$\frac{1}{H}$ marks the value of conduct parameter (θ^{mono}) when firms are perfectly collusive, θ is the average value of conduct parameter in each calendar year.

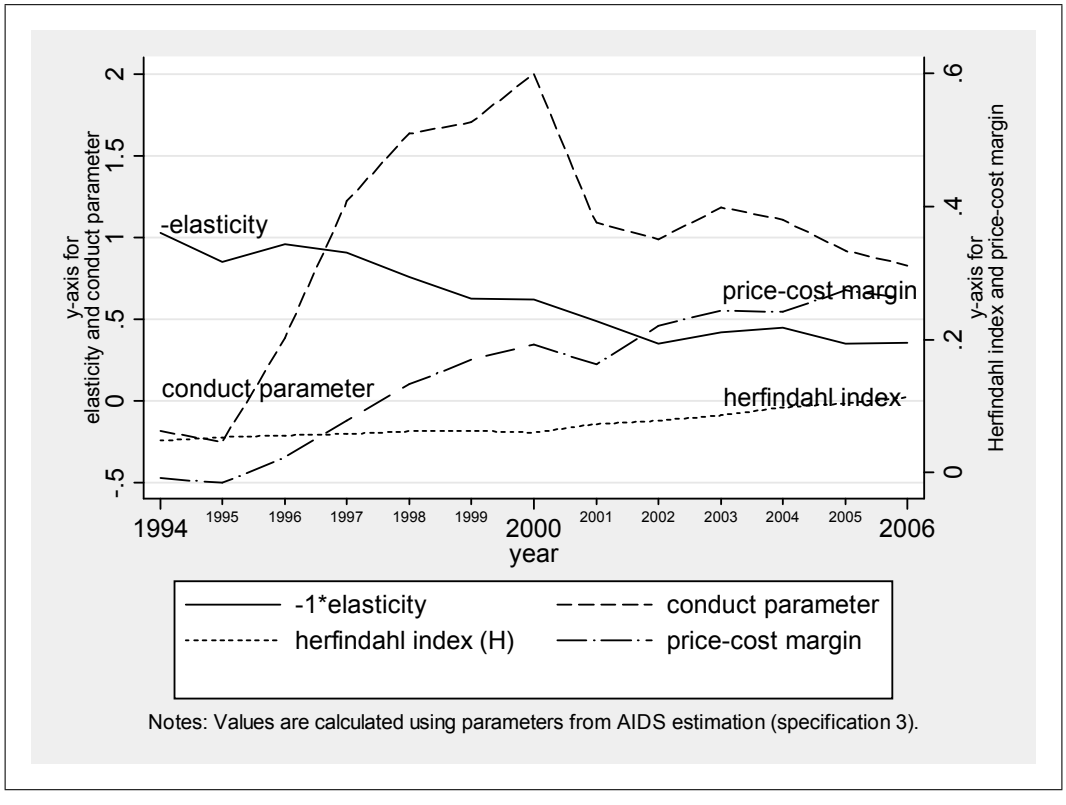


Fig. 1.7. The Conduct Parameter and Other Structural Indicators (1994-2006)

demand curve shifts out. If the willingness to pay by the new consumers is the same as that of the old consumers, the new demand curve would be parallel to the old one. However, if the willingness to pay changes, the demand curve rotates. To decompose the movement of the demand curve from the price elasticity, consider the following:

$$\varepsilon_{rm} = \frac{\partial q_{rm}}{\partial p_{rm}} \frac{p_{rm}}{q_{rm}},$$

where ε_{rm} is the own-price elasticity of demand for retail milk, q_{rm} is the equilibrium quantity, and p_{rm} is the equilibrium price. The formula suggests that demand elasticity can change due to 1) the slope of the demand curve $\frac{\partial q_{rm}}{\partial p_{rm}}$ (the willingness to pay), 2) the equilibrium quantity q_{rm} (demand), and 3) the equilibrium price p_{rm}. It is also possible that the change in price elasticity was affected by all three factors.

In the case of retail milk, the equilibrium quantity and equilibrium price increased steadily over time. The slope of the demand curve could also change if consumers' preference for retail milk had changed. But since the change of the slope cannot be observed immediately from the data or from the AIDS result, an analytical form of the slope has to be derived. Consider again the empirical AIDS function (eq. 1.5):

$$\begin{aligned} w_{rm,t} &= \alpha_0 + \beta_0 \log \frac{x_t}{P_t^r} + \gamma_1 \log p_{rm,t} + \gamma_2 \log p_{dm,t} + \sum_{m=1}^{M} \varphi_m \widetilde{pc}_{mt} + \alpha_1 \Upsilon_t \\ &\quad + \gamma_3 \Upsilon_t * \log p_{rm,t} + \sum_n \delta_n T_n + u_t \end{aligned}$$

where all variables are as defined in eq. 1.5. Substituting the expression for budget share $w = \frac{p_{rm}q_{rm}}{x}$ and rearrange, we get:

$$\begin{aligned} q_{rm,t} &= (\alpha_0 + \beta_0 \log \frac{x_t}{P_t^r} + \gamma_1 \log p_{rm,t} + \gamma_2 \log p_{dm,t} + \sum_{m=1}^{M} \varphi_m \widetilde{pc}_{mt} + \alpha_1 \Upsilon_t \\ &\quad + \gamma_3 \Upsilon_t * \log p_{rm,t} + \sum_n \delta_n T_n + u_t) \frac{x_t}{p_{rm,t}} \end{aligned} \tag{1.11}$$

now, assume that the total spending on grocery x_t and the retail price index P_t^r are invariant to price under a reasonably small price range. Therefore, taking derivative of (1.11) w.r.t. $p_{rm,t}$ and inverting the result would give the slope of the demand curve:

$$\begin{aligned} \frac{\partial p_{rm,t}}{\partial q_{rm,t}} &= \{(\alpha_0 + \beta_0 \log \frac{x_t}{P_t^r} + \gamma_1 \log p_{rm,t} + \gamma_2 \log p_{dm,t} + \sum_{m=1}^{M} \varphi_m \widetilde{pc}_{mt} + \alpha_1 \Upsilon_t \\ &\quad + \gamma_3 \Upsilon_t * \log p_{rm,t} + \sum_n \delta_n T_n + u_t) \frac{x_t}{p_{rm,t}^2} + \frac{x_t}{p_{rm,t}} (\frac{\gamma_1}{p_{rm,t}} + \frac{\gamma_3 \Upsilon_t}{p_{rm,t}}) \}^{-1} \\ &= \{w_t \frac{x_t}{p_{rm,t}^2} + \frac{x_t}{p_{rm,t}} (\frac{\gamma_1}{p_{rm,t}} + \frac{\gamma_3 \Upsilon_t}{p_{rm,t}}) \}^{-1} \\ &= \frac{p_{rm,t}^2}{x_t(w_t + \gamma_1 + \gamma_3 \Upsilon_t)}. \end{aligned}$$

which is increasing in p_{rm} and decreasing in $x_t, w_t, \gamma_1, \gamma_3$ and Υ_t. In other words, demand becomes more elastic at a higher price, and less elastic at a higher budget share and a higher level of female participation in the labour market. However, since p_{rm} x_t, w_t and Υ_t all increased over time, it is difficult to see how the slope had changed.

Figure 1.8, thus, uses parameter estimates from the AIDS estimation (table 1.2, specification 3) to illustrate the demand curve for the years 1994, 1997, 1999, 2004 and 2006. From this figure, we can see that the demand curve shifted outward and became steeper over time. The outward shift was due to an expansion of demand for retail milk. The steeper slope was due to a higher willingness to pay for retail milk.

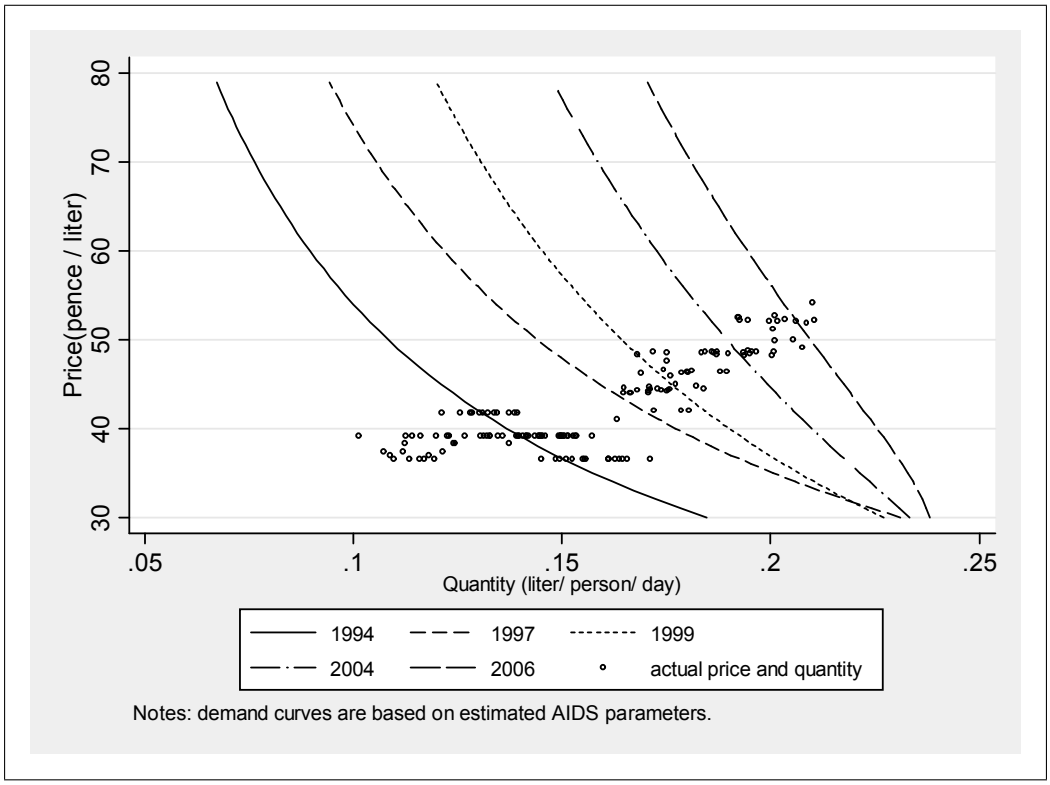


Fig. 1.8. Consumers' Demand Curve for Retail Milk by Year

Continuing the discussion in section 3, it was most likely that the demand for retail milk was growing because consumers were shifting away from doorstep milk. It was unlikely that such demand expansion was gainned from elsewhere because the overall demand for milk was declining during the sample period. As for the increase in the consumers' willingness to pay for retail milk, this could have been due to the change in the consumers' shopping habit. Households increasingly shifted away from traditional markets to supermarkets. This made it more convenient for households to combine their milk purchase with other grocery items. Bliss (1988) suggests that multi-product shopping outlets can save customers' shopping costs. This cost reduction could essentially increase the consumers' willingness to pay for retail milk.

1.6 Conclusion

This paper studies market power and investigates collusion in the UK retail milk industry for the period 1994-2006. The conduct parameter approach is used as the main framework. This framework analyses the market power from its three sources–inelastic demand, high market concentration and collusion. Using the direct measures of the marginal costs, I find that the retailers' price-cost margin grew from -0.9 percent in 1994 to 26.1 percent in 2006. Nevertheless, the estimated conducts suggest no improvement of collusion. I find that the retailers' growing profit margin was due to a decrease in the demand elasticity and a higher market concentration. Using the static conjectural variations game benchmark, firms did not achieve a higher profit-margin than the noncooperative Cournot level.

To arrive at the above conclusion, I estimate the demand curve for retail milk for the period 1994-2006. I find that the aggregate demand for retail milk expanded, while the consumer valuation of each litre of milk also increased. Price became less elastic from $\varepsilon_{em} =$ -1.03 in 1994 to $\varepsilon_{em} =$ -0.36 in 2006. In other words, when price doubles, demand for retail milk would be driven to zero in 1994; but would fall by only 36 percent in 2006. This change is significant. It is the main contribution to the impressive growth of the retailers' profit margin. Another factor that also changed in favour of market power is market concentration. Because large supermarket chains had gained more market share during the sample period, the Herfindahl index increased from 0.048 in 1994 to 0.113 in 2006.

The observed price, marginal cost, market concentration and the estimated demand elasticity are finally used in the conduct parameter analysis. Using the static conjectural variations games as benchmarks, I find very close to perfect Bertrand behaviour during the period 1994-1997. In fact, firms earned negative profits from 1994-1995. This could be explained by perfect competition among the retailers or even predatory pricing against the milkman and other competitors. In 1996, the retailers started to earn a positive profit. The level of conduct exceeded that of the noncooperative Cournot level during the period 1997-2000. However, it dropped back and remained around the noncoorporative Cournot level from 2001. In any case, the estimated price elasticity and market concentration suggest that firms could have earned a much higher profit if they had been perfectly collusive. Reasons beyond the focus of this paper may have prevented them from doing so. Some possibilities

are the breakdown of collusion attempts, threat of entry, and regulatory constraints, to name but a few.

Last but not least, readers should be reminded that my results are by no means indicative of the actual game(s) played. Although the static conjectural variations benchmarks (e.g. Bertrand and Cournot) are used extensively for discussion, they are mainly points of reference. In reality, it is possible for firms to adopt other decision rules. See for example Klemperer and Meyer (1989). Another caveat one should take is that this paper assumes a linear vertical pricing contract between milk processors and retailers. Although this assumption would not change the relative level of the estimated conduct, it would affect the absolute level. For example, if the milk processors and retailers were using an efficient non-linear contract, the retailers' marginal costs would be lower. As a result, the conduct would suggest a higher level of collusiveness. I carry out the analysis on vertical contract in the second chapter of this thesis.

1.A Calculation of the AIDS elasticities using the corrected Stone price index

The general definition of price elasticity of demand (ε_{ij}) and expenditure elasticity of demand (e_i) can be expressed as:

$$\begin{aligned} \varepsilon_{ij} &= \frac{\partial q_i}{\partial p_j}\frac{p_j}{q_i} = \frac{\partial \log q_i}{\partial \log p_j}, \\ e_i &= \frac{\partial q_i}{\partial x}\frac{x}{q_i} = \frac{\partial \log q_i}{\partial \log x} \end{aligned}$$

where ε_{ij} can be regarded as the ratio of percentage changes between quantity demanded of a product i and changes in the price of product j. On the other hand, e_i represents the ratio of percentage changes between quantity demanded of product i and changes in total expenditure.

To find the expression of ε_{ij} and e_i in my case, let's first consider my AID specification defined in eq.1.5:

$$\begin{aligned} w_{rm,t} &= \alpha_0 + \beta_0 \log \frac{x_t}{P_t^{cs}} + \gamma_1 \log p_{rm,t} + \gamma_2 \log p_{dm,t} + \sum_{m=1}^{M} \varphi_m \widetilde{pc}_{mt} + \alpha_1 \Upsilon_t \\ &\quad + \gamma_3 \Upsilon_t * \log p_{rm,t} + \sum_n \delta_n T_n + u_t \\ \log P_t^{cs} &= \sum_k w_{kt} \log(p_{kt}) \end{aligned}$$

where all variables take the same definition as in eq.1.5 and eq.1.6. The budget share of retail milk w_t can alternatively be written as:

$$w_{rm,t} = \frac{p_{rm,t} q_{rm,t}}{x_t} \tag{1.12}$$

where $p_{rm,t}$ is the price of retail milk, $q_{rm,t}$ is demand for retail milkst, x_t is total expenditure on grocery. Take log on both sides, we can express eq.1.12 as:

$$\begin{aligned} \log w_{rm,t} &= \log p_{rm,t} + \log q_{rm,t} - \log x_t \\ \log q_{rm,t} &= \log w_{rm,t} - \log p_{rm,t} + \log x_t \end{aligned} \tag{1.13}$$

Take the derivative of $\log q_{rm,t}$ with respect to $\log p_{rm,t}$, and assume that total expenditure is constant, we get the own-price elasticity of demand for retail milk:

$$\begin{aligned} \varepsilon_{rm,t} &= \frac{\partial \log q_{rm,t}}{\partial \log p_{rm,t}} = \frac{\partial \log w_{rm,t}}{\partial \log p_{rm,t}} - 1 \\ &= \frac{1}{w_{rm,t}} \frac{\partial w_{rm,t}}{\partial \log p_{rm,t}} - 1 \\ &= \frac{1}{w_{rm,t}} (-\beta_0 \frac{\partial P_t^{cs}}{\partial \log p_{rm,t}} + \gamma_1 + \gamma_3 \Upsilon_t) - 1 \\ &= \frac{1}{w_{rm,t}} (-\beta_0 w_{rm,t} + \gamma_1 + \gamma_3 \Upsilon_t) - 1 \\ &= \frac{\gamma_1}{w_{rm,t}} + \frac{\Upsilon_t \gamma_3}{w_{rm,t}} - \beta_0 - 1 \end{aligned}$$

Similarly, the cross-price elasticity of demand for retail milk is:

$$\varepsilon_{rm,j,t} = \frac{\log q_{rm,t}}{\log p_{j,t}} = \frac{\log w_{rm,t}}{\log p_{j,t}} - \frac{\log p_{rm,t}}{\log p_{j,t}}$$

where $p_{j,t}$ is the price of product j at time t. If we assume that a change in $p_{j,t}$ does not have any effect on $p_{rm,t}$, at least in the short-run, we can express $\varepsilon_{rm,j,t}$ as:

$$\begin{aligned} \varepsilon_{rm,j,t} &= \frac{1}{w_{rm,t}} (-\beta_0 w_{j,t} + \gamma_1 + \gamma_3 \Upsilon_t) \\ &= \frac{\gamma_1}{w_{rm,t}} + \frac{\Upsilon_t \gamma_3}{w_{rm,t}} - \beta_0 \frac{w_{j,t}}{w_{rm,t}} \end{aligned}$$

Finally, the expenditure elasticity of demand for milk can be derived by taking the derivative of $\log q_{rm,t}$ in eq.1.13 with respect to $\log x_t$:

$$\begin{aligned} e_i &= \frac{\partial \log q_i}{\partial \log x_t} \\ &= \frac{\partial \log w_{rm,t}}{\partial \log x_t} - \frac{\partial \log p_{rm,t}}{\partial \log x_t} + \frac{\partial \log x_t}{\partial \log x_t} \\ &= \frac{\beta_0}{w_{rm}} + 1. \end{aligned}$$

.

Chapter 2
Measuring the Effect of Collusion by UK Liquid Milk Suppliers

Abstract

In 2007, the major dairy processors and retailers admitted to the UK competition commission that they were involved in anti-competitive practices, such as exchanging commercially-sensitive information. This chapter attempts to assess the outcome of such anti-competitive behaviours. It generates a menu of outcomes under different vertical and horizontal competition scenarios. Then, it uses an encompassing test to find the scenario most consistent with the reality. The results suggest that the observed market prices are most consistent with two scenarios; 1) the perfectly competitive (Bertrand) manufacturers and noncooperative Cournot retailers under a linear vertical arrangement, and 2) the noncooperative Cournot retailers under a non-linear vertical contract with the manufacturers. Hence, this paper does not obtain sufficient evidence to support that the supermarkets achieved benefits beyond the static noncoorporative level.

2.1 Introduction and Summary

In September 2007, the Office of Fair Trading (OFT) accused the major UK supermarkets (Asda, Morrisons, Safeway, Sainsbury's and Tesco) and major dairy processors (Arla, Dairy Crest, Wiseman, Lactalis McLelland and The Cheese Company) for fixing the price of liquid milk. The alleged parties were subject to a combined penality of over 116 million pounds. In response to the case, the majority admitted to having been involved in anti-competitive practices–mostly exchange, or facilitating the exchange, of commercially sensitive information in respect of retail price. The OFT, in turn, had given a significant reduction in the financial penality, under condition that the parties continued to provide full coorporation (OFT, 2007a).

To what extent did the dairy processors and supermarkets actually collude? How much does the "exchange of commercially-sensitive information" facilitate firms to achieve a higher profit than at the competitive level? In this chapter, I attempt to find evidence of successful collusion by UK supermarkets and dairy processors during the period 2002-2005. I will also assess the magnitude of firms' profits and deadweight loss associated with such practice. In doing this, I adopt the menu approach suggested in Nevo (1998), Villas-Boas (2007), Puller (2007) and Bushnell et al. (2008). For each of the different vertical and horizontal competition scenarios, the menu reports corresponding price, quantity and the lower bound of the society's deadweight loss[18]. To formally justify the scenario that best

[18] As will be discussed in section 2.3.2, the lower bound of the deadweight loss can be obtained by using the linear demand curve. As there is not enough information to estimate how curvy the demand curve would be, it is not possible to obtain the actual, and the upper bound of the deadweight loss.

matches the actual data, an encompassing test is used (Davidson and MacKinnon, 1981, 1993).

Three conjectural variation schemes are used as benchmarks to analyse the horizontal relation among the supermarkets and among the dairy processors. These schemes are perfect Bertrand, noncooperative Cournot and joint-monopoly. The perfect Bertrand and noncooperative Cournot models, respectively, predict the lower and upper bound for prices in a static noncoorporative competition. The joint-monopoly model predicts the upper bound for prices in a static collusive equilibrium. Despite being used as benchmarks, there is no reason one should expect firms to strictly behave according to the models. In practice, firms could alternate between different competition strategies; e.g. Rotemberg and Saloner (1986) suggest that firms successfully collude when demand is low and play perfect Bertrand when demand is high. Green and Porter (1984) suggest otherwise. Alternatively, firms could adopt a different decision rule than Bertrand, Cournot and monopoly. For example, Klemperer and Meyer (1989) propose the supply function equilibrium (SFE) concept, where firms optimize according to a supply function which depends on both quantity and price. The SFE's outcomes are bounded by the perfect Bertrand and Cournot results (Klemperer and Meyer, 1989).

In addition to examining different horizontal competition schemes, I investigate the vertical relation between the supermarkets and the processors. This is done using two benchmarks, namely the linear and efficient non-linear contracts. It is only recently that empirical literature on vertical arrangement between upstream and downstream supply-chain entities has been widely examined. This is mainly because of the data availability.

Villas-Boas (2007) and Bushnell et al. (2008) highlighted that without accounting for the correct non-linear vertical arrangement, analysis on horizontal market structure would not be accurate. Villas-Boas (2007) finds evidence supporting non-linear vertical arrangement between the manufacturers and retailers in the American yogurt industry, while Bushnell et al. (2008) examine the case of the US electricity industry.

Having generated a menu of different horizontal and vertical competition scenarios, I then estimate the outcomes associated with them. Accurate estimations of outcomes under different scenarios require information from both the consumer side and the supplier side. In particular, the knowledge of consumers' demand elasticity and firms' marginal costs are needed. I obtain consumer-level data to estimate the consumers' demand curve and demand elasticity. In addition, I obtain reasonable approximates of dairy processors' and supermarkets' marginal costs. Along with equilibrium price and quantity, I will estimate the deadweight loss associated with each competition scenario. Such welfare analysis puts on view the magnitude of losses resulting from different degrees of market power exploitation.

The results suggest that the actual weekly price of liquid milk falls between that predicted by 1) perfect Bertrand dairy processors and perfect Bertrand supermarkets under a linear vertical arrangement and 2) noncooperative Cournot supermarkets under an efficient non-linear vertical arrangement. The formal encompassing tests further suggest that the first and second most consistent scenarios are; 1) perfect Bertrand dairy processors and noncooperative Cournot supermarkets under a linear vertical arrangement, and 2) the noncooperative Cournot supermarkets under a non-linear vertical arrangement. Considering that perfect Bertrand and noncooperative Cournot are both static noncooperative models,

I do not have sufficient evidence to suggest that firms achieved profits beyond the static noncooperative level.

Turning to the welfare consequence, my analysis predicts the actual profits of the alledged dairy processors and supermarkets to be 54.25 and 361.85 million pounds per year in 2004 (or 0.92 and 6.15 pounds per capita[19]) respectively. Compared with the most welfare-maximized scheme where suppliers receive no profit (price = marginal cost) and quantity supply reaches the highest possible level, the current scenario suggests the minimum deadweight loss of 27.14 million pounds per year (or 0.46 pounds per capita). Due to the data limitation, it is not possible to estimate the actual and the maximum deadweight loss without imposing a strong assumption on the shape of the demand curve.

This chapter proceeds as follows. Section 1 reviews the relevant literature. Section 2 gives a general overview of the UK liquid milk market. Section 3 conducts the demand estimation. Section 4 discusses the marginal costs. Section 5 uses the estimates of demand and marginal costs to generate a menu of outcomes. Section 6 discusses empirical results; while section 7 concludes.

I hereafter use the term "manufacturers" interchangeably with dairy processors, and "retailers" interchangeably with supermarkets.

[19] Based on the Office for National Statistics' (ONS) population estimate, the population of the Great Britain (we exclude Northern Ireland in this research) was 58,136,000 in 2004.

2.2 Literature Review

This chapter was inspired by a variation of both theoretical and empirical research on collusion. My work is the closet to Villas-Boas (2007), Bushnell et al. (2008) and Smith and Thanassoulis (2008). From Villas-Boas (2007), I adopt the consideration on vertical integration and the use of the menu of price approach. The paper examines different possible horizontal and vertical competition schemes among the retailers and the manufacturers in the US yogurt industry. After obtaining prices predicted by different possible schemes, the author selects the one that is most consistent with the actual data. The model developed in Villas-Boas (2007) allows firms to sell multi products, i.e. different brands of yogurt. Therefore, it has a large potential for the future use when more data become available. To incorporate the multi-product feature of the problem, this model requires the knowledge of price and cost of all products. Since I do not have such information (for all the products in all supermarkets), it is not possible to utilize the multi-product feature in my case.

Bushnell et al. (2008) adopts a menu of price approach similar to that of Villas-Boas (2007). Using the data from the US electricity industry, their model allows electricity retailers and wholesalers to engage in vertical contracts. After predicting prices resulted from different competition schemes, they use an encompassing test to find the model that is most consistent with the actual price. This encompassing test is similar to that developed by Davidson and MacKinnon (1993).

In both Villas-Boas (2007) and Bushnell et al. (2008), the dynamic feature of competition was not considered. Puller (2006) and Harrington (2005) proposed a model in which the dynamic nature of the competition can be incorporated. However, with limited data, the

estimation of a dynamic model would not significantly differ from the estimation of a static model. For example, according to Puller (2006, 2007), this would involve adding firm dummies to a static model.

In terms of the context, my paper studies the same industry as Smith and Thanassoulis (2008). The authors investigate collusion in the UK liquid milk supply chain using a bargaining model. This model allows the farmers, milk processors, and supermarkets to bargain for their share of profits. Thus, introduces an alternative means to account for vertical relations. I acquire much for the understanding on the UK milk industry from this paper.

In addition to the abovementioned papers, Harrington (2005) has provided a thorough discussion on empirical methods that can be used to detect cartels. These methods are categorized into four main groups. 1) the ones that test whether the behavior of suspected firms differ from that of competitive firms, 2) the ones that examine whether there is a structural break in firms behavior, 3) the ones that first identify competitive firms. Then, test whether the suspected firms behave the same way as the competitive firms, finally 4) the ones that predict outcomes of collusive and competitive conducts. Then, test whether the actual data is consistent with the collusive predictions. While the first two methods can only be used to reject competitive conducts, the last two methods can be used to identify collusion. The menu of price used in this paper can be counted as the last method. Harrington (2005) further discusses each of the methods as follows:

To implement the first method, the researcher finds behaviors that are only consistent with competitive conducts. Then, test whether the behaviors of suspected firms differ from

the competitive behaviors. Such competitive behaviors are, for example, positive correlations between price and cost, strong correlations between price and demand shocks, and independence of each firm's price/bidding values. Bajari and Ye (2003) uses this method to test whether firms' bids for seal coating[20] projects in the US are independent. This type of method was also used to detect bid-riggings in Porter and Zona (1993, 1999) and Pesendorfer (2000).

The second method, on the other hand, examines whether there exist any structural break in firms behaviors. In particular, the method is suitable for industries which were forming a cartel, breaking a cartel, or industries which experienced collusive and competitive episodes. This method requires the information prior to the suspected collusive period. In addition, to prevent misspecification, other factors that could cause structural breaks should be accounted for. Block et al. (1981) studies the US bread market. They find that a government's suspicion of price-fixing in one city could reduce the price in neighboring cities.

Some pitfalls of the first and second methods are that they do not give an indicative sign of collusion. Rather, they only suggest whether firms' behaviors are consistent with the competitive practices specified by the researcher. Therefore, the researcher needs to be aware that if the competitive practices, or the characteristics of the structural break, are incorrect, the results could also be incorrect.

The third method first identifies competitive firms. Then, test whether the behaviors of the suspected firms differ from those of the competitive ones. Funderburk (1974) uses

[20] Seal coating is a highway maintenance process, which is used to extend the life of the road. It involves coating the surface of the road with oil and other materials to reduce cracking.

this method to detect collusion in the biddings for asphalt contracts in Oklahoma. He finds that the same firms were able to offer much lower prices in other states, which were assumed to be more competitive. Other examples of this method can be seen in Porter and Zona (1993, 1999). The pitfall of this method is the potential endogeneity of the competitive firms, which are used as benchmarks. If some firms in the same industry are competitive, while others aren't, they could be competitive for a reason. To obtain consistent results, the researcher needs to control for factors (if any) which make them competitive.

The last method predicts outcomes of collusive (and competitive) conducts. Then, test whether the actual data is consistent with the collusive predictions. The menu of price approach can be counted under this method. Some examples can be seen in Villas-Boas (2007), Bushnell et al.(2008), Baldwin et al. (1997), Bajari and Ye (2003), and Banerji and Meenakshi (2004). When using this method, the researcher is required to construct and predict outcomes from different firms' conducts. Thus, its pitfall lies in the risk of misspecification of the collusion and competition models.

In terms of reflecting reality, empirical work on detecting collusion still lacks behind the theoretical work counterpart. While many important features of competition have been modeled theoretically, only a handful has been implemented with the data. This is mainly due to the lack of data. Most models require the knowledge of all the prices and costs of all products and firms in the industry over a sufficiently long period of time. Such a complete dataset is rare, if any does exist. Some important considerations that are crucial for the future development of empirical work on collusion are such as 1) the multi product

feature (Harrington, 1987), 2) the dynamic nature of competition (Puller, 2006), (Rotemberg and Saloner, 1986), (Green and Porter, 1984), (Haltiwanger and Harrington, 1991) and (Harrington, 2005, 2006), and 3) competition in the presence of an antitrust authority (Harrington, 2004, 2006).

2.3 Market for liquid milk in the UK

2.3.1 Supply Chain Structure

The UK liquid milk supply chain constitutes four main parties–farmers, cooperatives, dairy processors and supermarkets. Farmers run farms and milk their cow to get raw milk. Each farmer sells their raw milk to either a dairy cooperative or a milk processor. Dairy cooperatives can sell their milk to processors of different types of dairy products, i.e. liquid milk, butter, cheese, powder milk, etc. Milk processors buy milk either directly from farmers or from cooperatives. They process milk into different fat contents, and purify and bottle these for sale in the retail market. Consumers either buy milk directly from a store or have it delivered to their doorsteps by a milkman. Most milkmen are managed by a dairy processor. Supermarkets make direct contracts with dairy processors on price, volume and transportation arrangement of bottled fresh liquid milk. Figure (2.9) depicts the structure of the liquid milk supply chain.

In 2006, there were about 19,011 dairy farms throughout the UK. This number had fallen from 34,570 in 1996, but the average farm-size had grown (MDC, 2008a). The total number of liquid milk processors in the UK was 523 in 2006. However, only a few of them

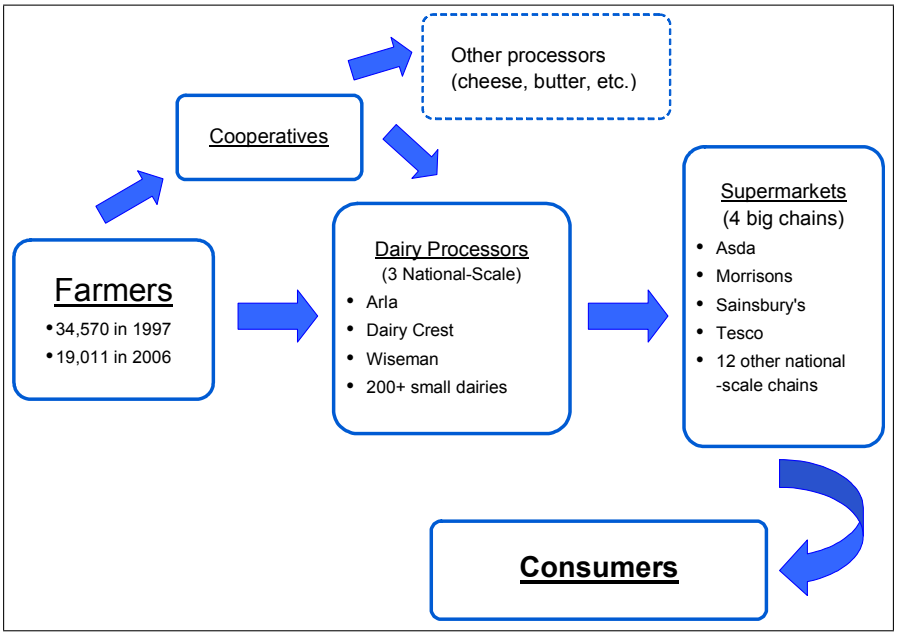


Fig. 2.9. Structure of UK Milk Supply Chain

were operating at the national level. In 2003, only Arla, Dairy Crest, Express and Wiseman were accounted for more than 90 percent of the milk supplied to the Big4 supermarkets. The number of major processors fell to only 3 after Arla acquired Express in 2004 (CC, 2007).

Turning to the supermarkets. In 2007, there were 17 supermarket chains operating nationwide. After Morrisons acquired Safeway in 2004, only the four largest firms–Asda, Morrisons, Sainsbury's and Tesco–have been holding the majority of the market share. These four supermarket firms are known as "the Big4". The market share of the Big4 has been growing throughout the past few years. In 2005, they accounted for approximately 70 percent of liquid milk sold to households[21].

[21] According to the TNS homescan data which will be discussed in the data section of this chapter.

Table 2.8. Market Share of Supermarkets and of the Milkmen

Firm	Market Share (%) by Year			
	2002	2003	2004	2005
Aldi	1.46	1.54	1.40	1.25
Asda	14.83	15.84	16.24	15.84
Budgen	0.34	0.37	0.34	0.31
Co-op	4.46	4.08	3.91	3.51
Iceland	1.06	1.15	1.30	1.35
Kiwk Save/Somerfield	6.34	6.22	5.75	5.72
Lidl	1.37	1.29	1.28	1.33
Marks & Spencer	0.63	0.59	0.51	0.59
Morrisons	5.91	6.26	12.95	12.26
Netto	1.09	1.18	1.33	1.40
Safeway[22]	7.41	6.66	–	–
Sainsbury's	11.70	11.58	11.39	11.78
Somerfield	2.82	2.87	3.01	3.48
Tesco	23.17	24.59	26.83	28.98
Waitrose	1.30	1.36	1.40	1.41
Other firms	3.28	3.40	4.05	4.21
Milkman	15.64	13.89	11.34	10.04
Herfindahl Index $(H) = \Sigma_{i=1}^{N} s_i^2$				
All firms; $\forall i$	0.13	0.13	0.15	0.16
Big4 Only; $i \in Big4$	0.25	0.25	0.28	0.29
The Big4's market share	63.02	64.94	67.40	68.86

In contrast to the increasing sales by the Big4, doorstep milk consumption has been falling. Table (2.8) shows the market share of all national supermarket chains and of the milkmen during the period 2002-2005.

2.3.2 Demand and Market Equilibrium

Despite seasonal fluctuations of milk demand and supply, liquid milk price has been relatively stable. During the sample period (September 2002 – August 2005), the price of private-label[23] liquid milk sold in the Big4 supermarkets adjusted twice. Both adjustments, first in the 30^{th} week of 2003 and second in the 13^{th} week of 2005, were due to dairy farm-

[23] Store's own brand, e.g. Tesco's milk, Sainsbury's milk, etc.

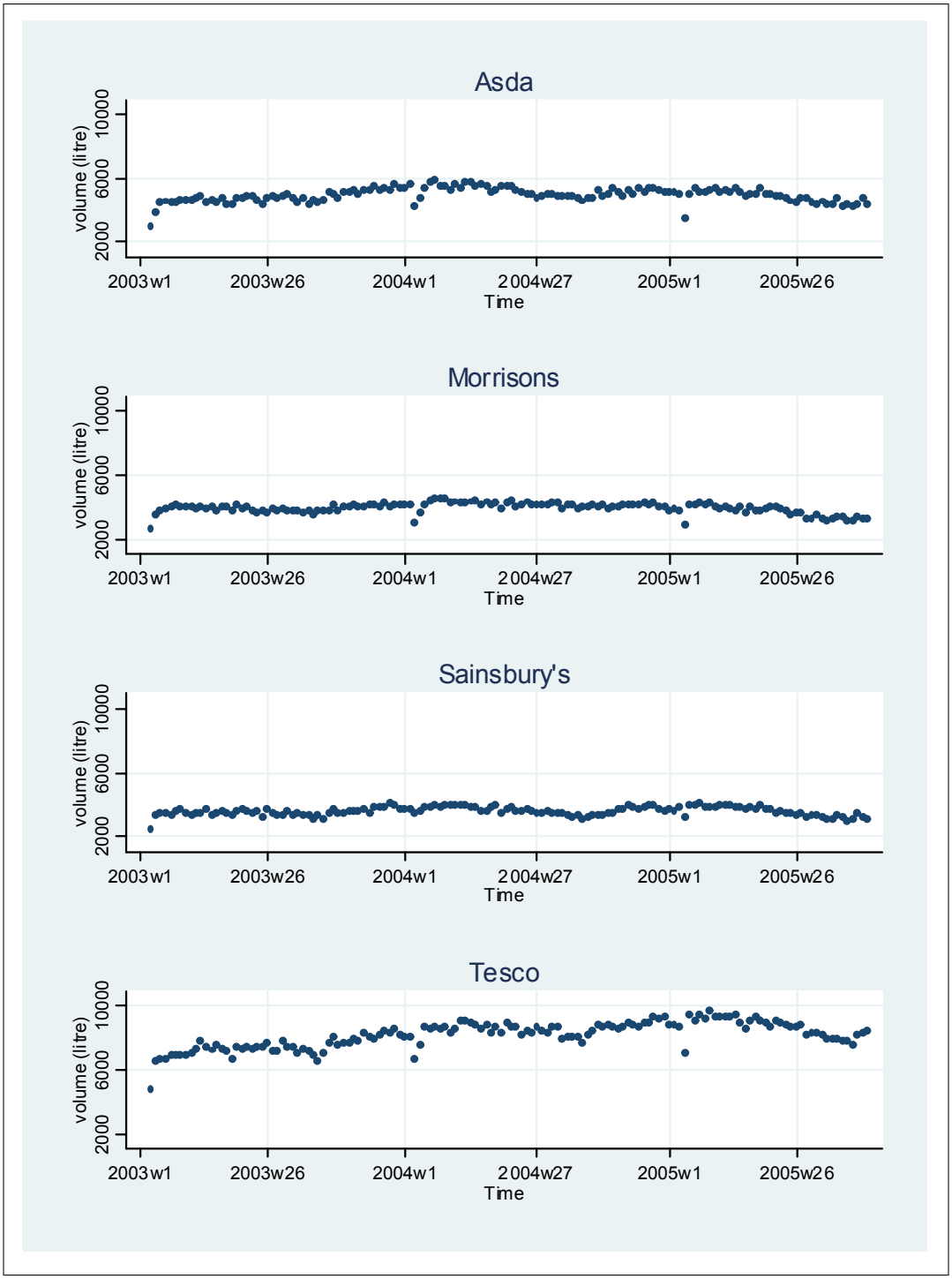


Fig. 2.10. Milk Quantity Supply by Firm

ers' direct actions where farmers demonstrated to put pressure on the supermarkets to raise the retail milk price. Figure (2.11) plots price and consumption of liquid milk in the UK.

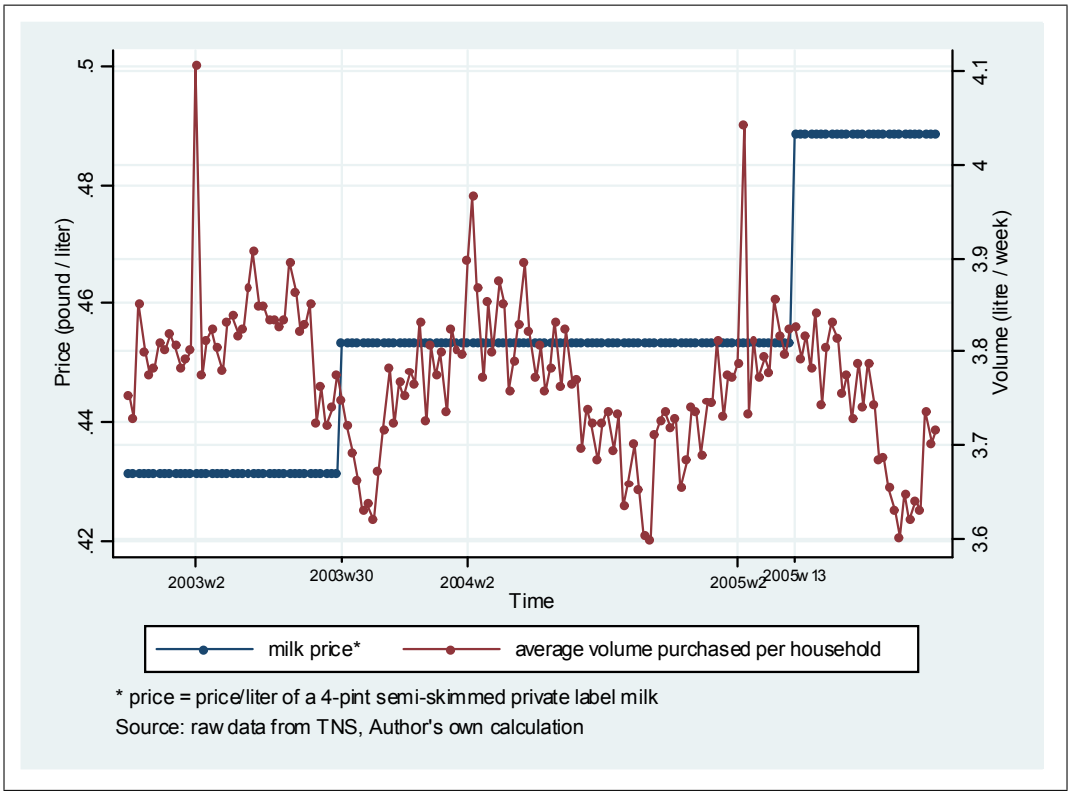


Fig. 2.11. Price and Household Consumption

Turning to household consumption, liquid milk constitutes about 3.7 percent of a household's total spending on grocery. This makes milk the third largest category in terms of spending, after vegetables (10 percent) and fruits (5.5 percent). Each day, over 11 million litres of liquid milk is consumed by UK households (MDC, 2008b). Liquid milk is available in different varieties according to fat content, purifying technology, brand, flavour, etc. Table (2.9) summarizes the proportion of different milk varieties sold in UK supermarkets

Table 2.9. Market Share by Types of Milk in 2005

Classification	Type (market share, %)*
Fat Content	Semi-Skimmed(58.1), Skimmed(15.6), Whole(26.3)
Purifying Technology	Pasteurized(86.3), U.H.T(9.1), Filtered(3.4), Sterilized(0.7)
Farm Production Process	Non-Organic(97.3), Organic(2.7)
Brand	Private Label (81.7), National Brand(18.3)

* Based on sales in calendar year 2005.
Source: raw data from TNS, Author's tabulation.

in 2005. Private-label semi-skimmed milk accounted for more than half of household milk consumption in the UK.

2.4 Demand

This study aims to investigate different market outcomes generated by firms under different competition scenarios. By outcomes, I mean different price and quantity equilibria as well as the corresponding consumer surplus, producer surplus and deadweight loss. Knowledge of demand is crucial to the analysis because it gives the level of market demand at any given price. If firms' marginal costs are also known, and a competition scenario is assumed, it would be possible to estimate the market outcomes. Before reaching that point, this section discusses the estimations of demand curves–one faced by the retailers and another faced by the manufacturers. Since the data dictates the setup of the estimations, I discuss the data first.

2.4.1 The Data

Data on consumers' characteristics and their grocery purchases are obtained from Taylor Nelson Sofres plc (TNS). The company used a randomization sampling method to recruit

households and ensure that their demographic variations represented that of the UK population. On agreeing to join the programme, each participating household was given a homescan equipment. They were asked to scan every item bought from retail grocery outlets and pass the information to the TNS. For each grocery item purchased, I observed price, quantity, product category and the store identification code. As for milk items in particular, I was able to acquire additional information on brand, fat content, pack size, purifying technology and whether or not it was labelled organic.

My data set covers a three-year period from September 2002 to August 2005. This includes shopping records of 26,133 different households across the UK. Since households were not obliged to participate for the entire period, their participation durations vary.

I adopt a household selection method similar to that of Hausman and Leibtag (2005) in order to avoid information distortion from uncommitted households. The term uncommitted households refers to those who participated for a very short duration of time or who were not serious about recording their purchases. I dropped records from any given household if they participated for less than 12 months. In addition, I dropped those who appeared to be active for less than $\frac{10}{12}$ of their participating duration. For example, if household A participated in the homescan programme for 12 months, they passed the first criterior. However, if their purchase records appeared for less than 10 months in the 12-month participating period, I dropped them because the participation rate would be less than $\frac{10}{12}$ of the time. A total of 14,480 households passed my sample selection criteria. From now on, all the information will be referred to these 14,480 households.

In this paper, the measurement of market power and the degree of collusion among liquid milk suppliers is done at the aggregate level. I evaluate the market power of the Big4 as one aggregate entity. Then, I assess whether there is any evidence of collusion among them. To do this, the elasticity of the residual demand for milk faced by the Big4 needs to be estimated. The demand estimation is discussed in the next section.

Since there are different numbers of participating households in each period, demand (or consumption) is measured in per-household unit. This is to make demand comparable across periods. The price of milk in the Big4 supermarkets is the weighted average price defined as follows:

$$p_{m,t} = \sum_{i=1}^{n} \left(\rho_{m_i,t} \times \frac{volume_{m_i,t}}{\Sigma_{i,t} volume_{m_i,t}} \right); m_i \in m$$

where subscript t denotes the weekly period, $p_{m,t}$ denotes the price index of the Big4's liquid milk in week t, $\rho_{m_i,t}$ denotes the price of liquid milk item i ($i = 1$ to n), $volume_{m_i,t}$ denotes the total volume of milk item i in week t, $\Sigma_{i,t} volume_{m_i,t}$ denotes the total volume of liquid milk sold by all the Big4 stores in week t.

Similar to the price of milk, the prices of other product categories are updated on a weekly basis. I calculate the weighted-average (by spending) price index of each product category as follows:

$$p_{k,t} = \sum_{i=1}^{n} \left(\rho_{k_i,t} \times \frac{spending_{k_i,t}}{\Sigma_{i,t} spending_{k_i,t}} \right); k_i \in k \tag{2.14}$$

where subscript t denotes the weekly period, $p_{k,t}$ denotes the price index of product category k in week t, $\rho_{k_i,t}$ denotes the price of item i ($i = 1$ to n) in category k, $spending_{k_i,t}$ denotes the total spending on item i in week t, $\Sigma_{i,t} spending_{k_i,t}$ denotes the total spending

on product category k in week t. I use spending as the weight for all other product categories because the units of measurement of other product categories are not unified in my data set. For example, for the fruit category, some types of fruits are measured in unit (e.g. 1 orange), some are in kilogram, etc.

Table 2.10. Correlation Matrix of Private-Label Milk Price 2002-2005

product		1 pint bottle			2 pint bottle			4 pint bottle		
type*		wh.	sm.	sk.	wh.	sm.	sk.	wh.	sm.	sk.
1	wh.	1.000								
pt.	sm.	0.985	1.000							
	sk.	1.000	0.985	1.000						
2	wh.	1.000	0.985	1.000	1.000					
pt.	sm.	1.000	0.985	1.000	1.000	1.000				
	sk.	1.000	0.985	1.000	1.000	1.000	1.000			
4	wh.	0.997	0.983	0.997	0.997	0.997	0.997	1.000		
pt.	sm.	0.997	0.983	0.997	0.997	0.997	0.997	1.000	1.000	
	sk.	0.983	0.997	0.983	0.983	0.983	0.983	0.987	0.987	1.000

Source: Raw data from TNS, Author's calculation and tabulation.
* Price of each type of private-label milk is identical across all Big4 stores.

2.4.2 Demand Curve Faced by the Retailers

Demand Function

The demand curve and demand elasticity of liquid milk faced by the Big4 supermarkets are estimated using a linear demand function. Although the linear assumption of demand may be unrealistic, it comes with two merits. First, it makes the calculation more simplistic; and second, it provides the lower bound for the price of any demand curve that is non-convex to the origin (Hausman, 1997). To test the robustness of the estimation, the

results will be compared with those obtained from a log-linear demand specification. The log-linear demand function is widely used in studies where only the point-estimate of demand elasticity is of interest (see for example Graddy (1995) Puller (2007)). When the entire demand curve is of interest, however, this demand function imposes an unrealistic assumption, e.g. it assumes a constant demand elasticity for the entire curve.

Demand estimation is usually subject to the simultaneity bias problem. This is because price and quantity are simultaneously determined in the demand vs. supply equilibrium. Therefore, the price variable in the demand function is not exogenous and would result in inconsistent OLS coefficient estimates. I remedy the simultaneity problem using the instrumental variable (IV) approach. Cost components of the retailers should serve as valid instruments because, while they are strongly correlated with price, they should not be correlated with demand.

The retailers' cost components are acquired from various sources. I obtained the monthly Intervention Milk Price Equivalent (IMPE), the monthly Actual Milk Price Equivalent (AMPE) and the Milk for Cheese Value Equivalent (MCVE) from the MDC website. I also obtained the wholesale price of liquid milk sold by Wiseman plc directly from the MDC research unit. The IMPE is constructed using data from several raw milk buyers. It has been used as the benchmark for raw milk price. AMPE measures market price of butter and skim milk powder. This is, then, the opportunity cost of selling milk for liquid milk production. Similarly, MCVE is the MDC's estimate of the value of milk sold to butter and cheese processors. This is practically the farmers' opportunity cost of raw milk. Finally, the "wiseman price" is the average price that Wiseman charged supermarkets for their bot-

tled liquid milk. This price includes delivery to supermarkets' outlets. Thus, wiseman price should also be strongly correlated with retail price set by supermarkets. Apart from the cost of milk, I obtained the average earnings index (AEI), which can proxy for the labour cost from the Office for National Statistics (ONS) website.

In addition to the simultaneity bias, I remedy the omitted variable bias by including a number of control variables. Firstly, a new year dummy and a spring season dummy are used to account for seasonal variations of milk consumption. The new year dummy takes the value of one for the first and last week of every year and zero otherwise. This is when milk consumption abruptly increases (see figure 2.11). The spring dummy takes the value of one for observations in March - June and zero otherwise. During these months of the year, milk consumption is higher than average. Secondly, prices of other grocery products are used to control for the opportunity cost (or relative price) of retail milk. For example, if product A is a substitute for milk, *ceterus paribus*, a reduction in its price should decrease the demand for retail milk. Since I observe prices of over 30 product categories, which could reasonably be substitutes for or complements of retail milk, I use the principal component analysis (PCA) to compress the number of regressors. Further detail on the implementation of the PCA can be found in chapter 1 section(1.3.2).

Finally, the empirical specification of my linear demand function can be expressed as follows:

$$q_{m,t} = \alpha + \theta_0 x_t + \theta_m p_{m,t} + \sum_{j}^{J} \theta_{mj} pc_{j,t} + \delta_1 * new_year + \delta_2 * spring + \varepsilon_t \quad (2.15)$$

where subscript t denotes different weekly periods, $q_{m,t}$ is the average consumption of milk per household per week, $p_{m,t}$ is the weighted-average (by volume) price of milk sold

in Big4 supermarkets, x_t is household's total spending on grocery, pc_j is the j^{th} principal component, J denotes the total number of principal components included in the specification, new_year denotes the new year dummy, $spring$ denotes the spring dummy, ε_t is the error component, $\alpha, \theta_0, \theta_m, \theta_{mj}, \delta_1$ and δ_2 are parameters to be estimated. The resulting point-elasticity of price (η_m) and point-elasticity of expenditure (e_m) can be calculated by:

$$\begin{aligned} \eta_m &= \frac{\partial q_{m,t}}{\partial p_{m,t}} \times \frac{p_{m,t}}{q_{m,t}} = \theta_m \frac{p_{m,t}}{q_{m,t}} \\ e_m &= \frac{\partial q_{m,t}}{\partial x_t} \times \frac{x_t}{q_{m,t}} = \theta_0 \frac{x_t}{q_{m,t}}. \end{aligned}$$

Thus, using the average values $\bar{p}_{m,t}, \bar{q}_{m,t}$ and $\bar{x}_t$ would yield the average price and income elasticities over the observed price range.

Adopting the same notation as the linear demand case, the log-linear demand function can be expressed as follows:

$$\log q_{m,t} = \varpi + \sigma_0 \log x_t + \sigma_m \log p_{m,t} + \sum_j^J \sigma_{mj} \widetilde{pc}_{j,t} + \phi_1 * new_year + \phi_2 * spring + \epsilon_t \tag{2.16}$$

where $\log(\bullet)$ denotes the log operation, $\widetilde{pc}_{j,t}$ denotes the j^{th} principal component of logged prices of different product categories, and $\varpi, \sigma_0, \sigma_m, \sigma_{mj}, \phi_1$ and ϕ_2 are parameters to be estimated. This specification readily gives the estimate of price elasticity $(\eta_m = \sigma_m)$ and expenditure elasticity $(e_m = \sigma_0)$.

Demand Estimation Results

Table 2.11 shows estimation results from linear demand specifications. Only specification 1 uses the OLS method which does not account for the endogeneity of price (p_{mt}). Specifications 2, 3 and 4 use the IV approach to account for the endogeneity of p_{mt}. The

error terms in all specifications are calculated using the Huber-White sandwich estimator to account for the heterogeneity in the error terms (Huber, 1967) (White, 1980). For all the IV specifications, I report the p-value of the Hansen J-test–which tests the null-hypothesis of p_{mt} being exogenous. The null hypotheses are rejected at the 5 percent level of significance, unless stated otherwise. According to the final specification (specification 4), I cannot accept the null hypothesis that p_{mt} is exogenous (p-value = 0.204). Finally, the last two rows report the point-estimates of mean price elasticity η_m and mean expenditure elasticity e_m.

Specification (1)OLS includes all the control variables but leaves p_{mt} subject to the endogeneity bias. The specification gives a point-estimate of price elasticity of η_m = -0.265, which is less elastic than those suggested by the IV specifications. This direction of bias implies that price and quantity are positively correlated in the supply relation, which is intuitive. Without accounting for the endogeneity of price in the demand function, price would be correlated with quantity shocks in the error term. The sign of the correlation is the same as that of the correlation between price and quantity in the supply relation. If price and quantity are positively correlated in the supply relation, η_m would be biased upward (more positive) and vice-versa.

Specifications 2,3 and 4 remedy the endogeneity of price using the IV approach. Cost components namely AMPE, IMPE, MCVE, wiseman price and AEI are used as instruments. We can see that, after the endogeneity is accounted for, coefficients on price become more negative. Specifications 2 and 3 explore the role of different sets of control variables. Specification 2 includes only the seasonal dummies and one retail price index

(RPI), while specification 3 includes only the principal components $(pc_{j,t})$.Since controlling for both sets of variables are robust to the results, I include all of them in specification 4.

Overall, the values of coefficients are intuitive. The significant and negative coefficient on price (p_{mt}) suggests that demand decreases when price increases. The significant and positive coefficient on expenditure (x_t) implies that retail liquid milk is a normal good. The different values of coefficients on the principal components, both positive and negative, suggest that they represent prices of both substitutes and complements of milk. The significant and positive coefficient on RPI suggests that most retail products are substitutes for retail milk.

As for the effects of seasonal dummies, the coefficient on $spring$ is positive and significant. This is expected because it is the period when milk consumption is higher than on average. On the other hand, the coefficient on new_year is insignificant. This could be because the consumption shocks during the new year period has already been accounted for by expenditure on grocery (x_t), which also peaks during the new year period.

The robustness to different numbers of principal components has already been explored in table (1.4) of chapter 1. I found that the results become unrobust to additional principal components once a certain number of principal components have been included– in this case, after the first 12 components.

I also test the robustness of the linear demand function (eq.2.15) by replicating demand estimation using the log-linear function (eq.2.16). Table 2.12 reports the results from the log-linear demand estimation. Overall, the two functional forms do not give signifi-

cantly different estimates of elasticities. In the final specification (4)IV, the linear demand function gives point-estimates of mean $\eta_m = -0.403$ and mean $e_m = 0.246$, while the log-linear demand function gives estimates of the constant $\eta_m = -0.472$ and constant $e_m = 0.266$[24].

2.4.3 Demand Curve Faced by the Manufacturers

Apart from demand elasticities faced by the supermarkets, knowledge of demand elasticities faced by the manufacturers is also crucial. Such information allows us to predict wholesale prices set by the manufacturers under different competition schemes. Hence, it enables us to assess market power among the manufacturers.

Estimation of demand elasticities faced by the manufacturers can be done in two ways. First, one can *directly estimate* it in the same way I did in the retailers case. Second, one can *infer* it from the retailers' consumption behaviour. This is because the retailers are the only demanders for liquid milk from the manufacturers.

In this study, I do not have sufficient information to directly estimate elasticities faced by the manufacturers. Therefore, the manufacturers' elasticities are obtained through the second means. The demand curve faced by the manufacturers will be inferred from three different retailers' competition behaviours–perfect Bertrand, joint-monopoly and Cournot competition. The same consumers' linear demand function is assumed in all cases, which can be written as:

$$P_t = \alpha + \frac{\partial P}{\partial Q} Q_t \tag{2.17}$$

[24] By functional form, the log-linear demand function assumes constant elasticities for the entire demand curve.

Table 2.11. Regression Results for the Linear Demand Estimation

Dependent variable = q_{mt}				
Variable	(1) OLS	(2) IV	(3) IV	(4) IV
constant	-0.108 (0.431)	-0.183 (0.265)	2.467 (0.332)	0.043 (0.434)
p_{mt}	-0.522 (0.195)	-1.500 (0.254)	-0.851 (0.314)	-0.795 (0.296)
x_t	0.482 (0.077)	0.268 (0.096)	0.501 (0.079)	0.482 (0.079)
$pc_{1,t}$	-6.862 (1.612)	–	-4.479 (1.730)	-6.776 (1.629)
$pc_{2,t}$	-3.971 (3.011)	–	12.356 (3.821)	-1.373 (3.506)
$pc_{3,t}$	-11.621 (1.483)	–	-11.927 (1.416)	-11.004 (1.490)
$pc_{4,t}$	-11.523 (3.058)	–	-27.163 (2.631)	-12.609 (3.066)
$pc_{5,t}$	1.121 (2.379)	–	4.984 (2.219)	2.585 (2.546)
$pc_{6,t}$	11.117 (3.666)	–	8.570 (2.809)	11.611 (3.701)
$pc_{7,t}$	-2.033 (2.629)	–	-1.346 (2.579)	-2.209 (2.632)
$pc_{8,t}$	-5.618 (2.418)	–	-8.141 (2.968)	-7.248 (2.759)
$pc_{9,t}$	4.432 (2.865)	–	8.498 (2.997)	4.341 (2.783)
$pc_{10,t}$	-3.437 (2.567)	–	-1.828 (3.096)	-4.018 (2.642)
$pc_{11,t}$	5.397 (2.897)	–	10.561 (3.248)	5.705 (2.913)
$pc_{12,t}$	-0.424 (2.946)	–	-0.731 (3.298)	-0.160 (2.971)
RPI	0.022 (0.004)	0.035 (0.005)	–	0.023 (0.004)
$spring$	0.029 (0.009)	0.048 (0.008)	–	0.024 (0.010)
new_year	-0.009 (0.022)	-0.024 (0.018)	–	-0.005 (0.022)
no. of obs	156	156	156	156
Adjusted-R^2	0.779	0.414	0.706	0.776
Hansen-J (p-value)	–	0.001	0.001	0.101
Own-price (η_m)	-0.265 (0.099)	-0.761 (0.129)	-0.432 (0.159)	-0.403 (0.150)
Income (e_m)	0.246 (0.039)	0.137 (0.049)	0.255 (0.040)	0.246 (0.040)

Notes: All error terms are calculated by the Huber-While estimator to account for heteroskedasticity in the error terms. * Hansen-J tests the null hypothesis of p_{mt} being exogenous. ** η_m is evaluated at mean price $p_{m,t}$ = 0.464 ppl, and mean quantity/household/week $q_{m,t}$ = 2.097 litre.

Table 2.12. Regression Results for the Log-linear Demand Estimation

Dependent variable = $\log q_{mt}$				
Variable	(1) OLS	(2) IV	(3) IV	(4) IV
constant	-0.380 (0.201)	-0.922 (0.210)	0.752 (0.011)	-4.924 (1.025)
$\log p_{mt}$	-0.276 (0.101)	-0.776 (0.126)	-0.481 (0.168)	-0.472 (0.156)
$\log x_t$	0.265 (0.045)	0.153 (0.054)	0.271 (0.047)	0.266 (0.046)
$pc_{1,t}$	-3.586 (0.910)	–	-0.520 (1.059)	-3.156 (0.947)
$pc_{2,t}$	1.485 (1.396)	–	-6.437 (1.780)	-0.152 (1.625)
$pc_{3,t}$	-5.695 (0.701)	–	-5.953 (0.676)	-5.311 (0.696)
$pc_{4,t}$	-4.813 (1.470)	–	-12.731 (1.269)	-5.482 (1.472)
$pc_{5,t}$	1.473 (1.127)	–	3.178 (1.002)	2.466 (1.193)
$pc_{6,t}$	5.253 (1.766)	–	3.912 (1.460)	5.354 (1.800)
$pc_{7,t}$	-1.028 (1.248)	–	-0.825 (1.228)	-1.221 (1.249)
$pc_{8,t}$	2.849 (1.158)	–	4.389 (1.483)	4.006 (1.359)
$pc_{9,t}$	1.688 (1.322)	–	3.546 (1.409)	1.522 (1.275)
$pc_{10,t}$	-2.020 (1.224)	–	-1.515 (1.510)	-2.544 (1.273)
$pc_{11,t}$	2.157 (1.342)	–	4.494 (1.525)	2.293 (1.343)
$pc_{12,t}$	0.312 (1.488)	–	-0.463 (1.634)	0.547 (1.525)
RPI	0.011 (0.002)	0.017 (0.002)	–	1.226 (0.222)
$spring$	0.014 (0.004)	0.023 (0.004)	–	0.011 (0.005)
new_year	-0.006 (0.010)	-0.012 (0.009)	–	-0.004 (0.010)
no. of obs	156	156	156	156
Adjusted-R^2	0.776	0.414	0.696	0.770
Hansen-J* (p-value)	–	0.001	0.001	0.204
Own-price** (η_m)	-0.276 (0.101)	-0.776 (0.126)	-0.481 (0.168)	-0.472 (0.156)
Income (e_m)	0.265 (0.045)	0.153 (0.054)	0.271 (0.047)	0.266 (0.046)

Notes: All error terms are calculated by the Huber-While estimator to account for heteroskedasticity in the error terms. * Hansen-J tests the null hypothesis of p_{mt} being exogenous.

where subscript t denotes time, P_t is the retail price paid by the consumers, Q_t is the total consumer demand, α is a constant intercept and $\frac{\partial P}{\partial Q}$ is the slope of the demand curve obtained from the previous consumer's demand estimation. I also assume the following marginal cost structure for the retailers:

$$MC_t^r = p_t^w + c_t^r$$

where p^w is the wholesale price per unit charged by the manufacturers and c^r is the retailers' additional marginal cost. The optimal p^w will be estimated given the manufacturers' and retailers' competition scheme. The component c^r will be directly approximated using the market information. Finally, the demand function faced by the manufacturers $(p^w = f(Q_t))$ can be estimated using the functions below:

- Given perfectly competitive retailers: retailers set their price equal to marginal cost.

$$\begin{aligned} MC_t^r &= P_t = \alpha + \frac{\partial P}{\partial Q} Q_t \\ p_t^w &= \alpha - c_t^r + \frac{\partial P}{\partial Q} Q_t \end{aligned} \tag{2.18}$$

The demand curve faced by the manufacturers is as steep as the one faced by the retailers.

- Given joint-monopoly retailers: quantity demanded by the retailers would be the quantity that maximizes their joint-profit.

$$\begin{aligned}
\arg\max_{Q_t} \Pi_t &= (P_t - MC_t^r)Q_t \\
0 &= (\alpha + \frac{\partial P}{\partial Q}Q_t - MC_t^r) + \frac{\partial P}{\partial Q}Q_t \\
MC_t^r &= \alpha + 2\frac{\partial P}{\partial Q}Q_t \\
p_t^w &= \alpha - c_t^r + 2\frac{\partial P}{\partial Q}Q_t
\end{aligned} \tag{2.19}$$

  Therefore, the demand curve faced by the manufacturers is twice as steep as the one faced by the retailers.

- Given Cournot retailers: each retailer unitarily maximizes their own profit given other retailers' best responses.

$$\arg\max_{q_{it}} \Pi_{it} = (P_t - MC_{it}^r)q_{it}$$

  Solving the first order condition, we get:

$$\begin{aligned}
0 &= (\alpha + \frac{\partial P}{\partial Q}q_{it} + \sum_{j\neq i}\frac{\partial P}{\partial Q}q_{jt} - MC_{it}^r) + \frac{\partial P}{\partial Q}q_{it} \\
MC_{it}^r &= \alpha + 2\frac{\partial P}{\partial Q}q_{it} + \sum_{j\neq i}\frac{\partial P}{\partial Q}q_{jt} \\
\frac{MC_{it}^r}{Q_t} &= (\alpha + 2\frac{\partial P}{\partial Q}q_{it} + \sum_{j\neq i}\frac{\partial P}{\partial Q}q_{jt})/Q_t \\
\frac{MC_{it}^r}{Q_t} &= \frac{\alpha}{Q_t} + 2\frac{\partial P}{\partial Q}s_{it} + \sum_{j\neq i}\frac{\partial P}{\partial Q}s_{jt} \\
MC_{it}^r &= \alpha + (s_{it}+1)\frac{\partial P}{\partial Q}Q_t \\
p_t^w &= \alpha - c_t^r + (s_{it}+1)\frac{\partial P}{\partial Q}Q_t \\
\bar{p}_t^w &= \alpha - \bar{c}_t^r + (\overline{s_{it}+1})\frac{\partial P}{\partial Q}Q_t
\end{aligned} \tag{2.20}$$

where s_{it} is the market share of firm i, $\bar{p}_t^w$ is the industry-average manufacturers' wholesale price, $\bar{c}_t^r$ is the industry-average retailers' marginal costs and $(\overline{s_{it}+1})$ is the industry-average of $s_{it}+1$. The demand curve faced by the manufacturers is steeper than the one faced by the retailers. The slope of the manufacturers' demand curve increases with market concentration at the retailers level.

2.5 Marginal Costs

In contrast to the demand curve, marginal costs of the retailers and the manufacturers are approximated (rather than estimated) using information from reliable sources. I first discuss the retailers' marginal costs, and then, the manufacturers' marginal costs.

2.5.1 Retailers' Marginal Costs (MC^r)

The retailers marginal costs (MC^r) consist of two parts. First, the marginal cost of raw milk paid to the manufacturers (MC^w) and, second, the supermarket's own operational marginal cost (c^r). As discussed earlier, the marginal costs of raw milk under linear and non-linear contracts are different.

For linear contracts, manufacturers set a flat wholesale price per unit (p^w). The retailers make their payment according to this wholesale price only. Therefore, the retailers'

marginal cost is p^w plus c^r.

$$\begin{aligned} \Phi_i^r &= p^w q_i \\ p^w &= \frac{\Phi_i^r}{q_i} \\ &\Rightarrow \quad MC^r = p^w + c^r \end{aligned} \tag{2.21}$$

where Φ_i^r is the total payment by retailer i, and q_i is the total quantity bought by retailer i.

As for non-linear contracts, I assume an efficient two-part tariff. The manufacturers charge a flat price per unit equal to their marginal cost (MC^w). They then, extract a lump sum profit (F) from the retailers. Therefore, the retailers' marginal cost is MC^w plus c^r.

$$\begin{aligned} \Phi_i^r &= MC^w q_i + F_i \\ p^w &= \frac{\Phi_i^r}{q_i} = \frac{MC^w q_i + F_i}{q_i} \\ &\Rightarrow \quad MC^r = MC^w + c^r \end{aligned} \tag{2.22}$$

The wholesale price (p^w) and the supermarket's operational marginal cost (c^r) are needed for the calculation of MC^r. In this study, they are obtained as follows:

The Wholesale Price (p^w)

I obtained the direct approximation of p^w from the MDC. This p^w is the annual-average price per litre of milk, inclusive of transportation to supermarket outlets, charged by Wiseman plc. The MDC estimated this price by dividing the company's turnover by the volume of liquid milk sold per year. Since Wiseman's business is mainly devoted to liquid milk, where 70% is supplied to the Big4 supermarkets (Smith and Thanassoulis, 2008), I believe that Wiseman's price is a reasonable approximation of the average wholesale price

per unit. This direct approximation of the actual p^w will be used for comparison with the ones predicted by different vertical/horizontal competition scenarios. Figure(2.12) plots Wiseman plc's price during the sample period.

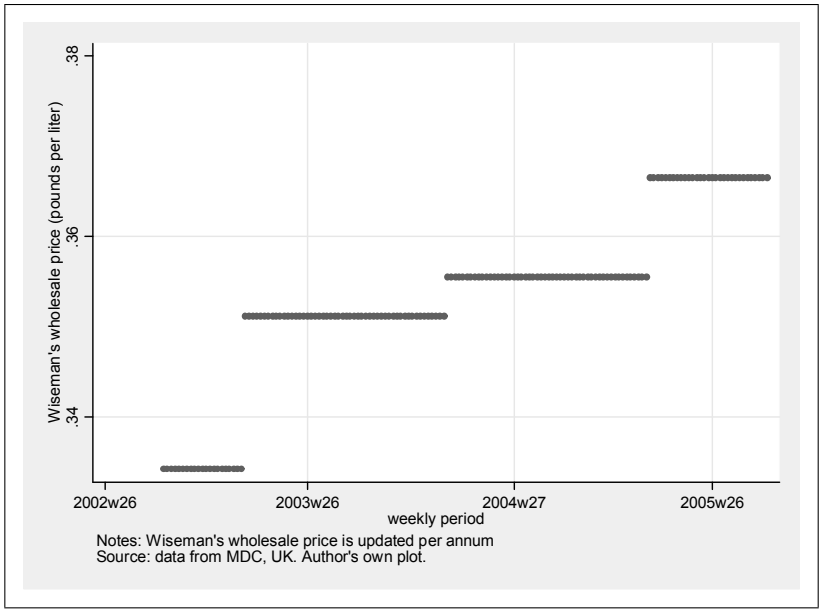


Fig. 2.12. p^w- The Actual Wholesale Price (proxied by Wiseman's price).

Supermarket's operational marginal costs (c_i^r)

On top of the wholesale milk, supermarkets have to pay for the operational costs (c_{it}). This is largely the cost for shelving (labour) and storage (energy). I was able to obtain sufficient information of the four biggest firms in 2004 (Asda, Morrisons, Sainsbury's, Tesco); and Safeway in 2003. The labour cost is calculated by first, finding the total number of fulltime-equivalent employees in each firm. Then, multiplying the number of employees by the average earnings in the retail sector for 8 hours per day. This gives an estimate of the labour cost per day. To find the contribution of this cost by each litre of liquid milk, I

multiply the total labour cost with the proportion of milk items (packs) to all grocery items (packs) in store. I then, divide the amount by the total volume (in litre) of milk sold per day. The numbers of fulltime-equivalent employees are obtained from each company's annual reports. The average earnings of labour in the retail sector is obtained from the Office of National Statistics (ONS) website. The proportion of milk items to all grocery items, and the total volume of liquid milk sold by each supermarket chain are obtained from the TNS homescan data set.

My calculation of the energy cost is guided by the DEFRA report on the impact of UK liquid milk consumption (DEFRA, 2007). According to the report, an average supermarket uses 37 percent of its electricity for refrigeration, while liquid milk occupies about 5 percent of the refrigerated shelves. Thus, liquid milk accounts for about 1.85 percent of stores' electricity cost. The total eletricity cost is calculated as follows:

$$\text{Firm } i's \text{ total electricity cost} = A_i \times B_i \times C_i$$

where A_i is firm $i's$ average electricity usage per outlet, B_i is the total number of outlets owned by firm i and C_i is the average electricity price per unit. The average electricity usage per outlet is obtained from a survey conducted by the Green Liberal Democrats, an environmental campaign group of the Liberal Democrats party (UK) (Baker, 2004). Although the report disguises the identity of each supermarket, I was able to identify them through the annual turnovers, number of employees, number of stores and other information given. The total number of outlets is obtained from each firm's annual report. The price of electricity, according to the level of usage, is obtained from the department for business enterprise & regulatory reform's website (BERR).

Since the approximated costs are for 2004 only, I inflate and deflate them as appropriate for all other years. I use the average earnings index to inflate the labour cost, and use the average of the electricity and gas price index to inflate the energy cost. All these indices are available on the ONS website. In any case, I believe the approximation error of c_i^r should not have a serious effect on the final results. This is because c_i^r is very small (about 0.89 - 1.82 pence per litre) compared to other cost components.

Table (2.13) reports the summary statistics of variables used for the calculation of c_i^r. The first three rows report values associated with the average earnings index, the electricity price index and the gasoline price index which are used to adjust the labour and electricity costs. The subsequent rows report other values used for the calculation of the labour and electricity costs. On average the electricity useage per outlet is 118.83 kilo watt per square foot per year. The total electricity price is 0.08 pounds per kilo watt hour (kWh). This price is based on the usage of 3300 kWh per year and was observed for 2004 only. I use the electricity price index to inflate or deflate this price for all other years. In terms of number of stores, Tesco has the highest number, followed by Sainsbury's, Safeway, Asda and Morrisons respectively. However, after Morrisons acquired Safeway in 2004, it became the third largest firm in terms of the number of stores. Turning to the store size, Asda's outlets are the largest on average, followed by Morrisons, Sainsbury's and Tesco. In terms of employees, Tesco employed the most people on average, followed by Sainsbury, Asda and Morrisons. Again, after the Morrisons-Safeway acquisition in 2004, the number of Morrisons' employees exceeded that of Asda's.

Table (2.14) reports the labour and energy costs of fresh processed liquid milk. The labour cost is significantly higher than the electricity cost. Asda has the lowest labour cost of about 0.88-0.98 pence per litre of milk, while Sainsbury's has the highest labour cost of about 1.55-1.72 pence per litre. All stores incurred approximately the same electricity cost of about 0.0002-0.0006 pence per litre.

Table 2.13. Summary Statistics of Information for Marginal Costs Estimation

Variable Description	Mean	Standard Deviation	Minimum	Maximum
Average earnings index (retail sector)	113.43	2.98	106.50	118.30
Electricity price index	105.62	10.81	94.56	131.04
Gasoline price index	117.06	23.31	90.50	174.01
Electricity usage (kWh/sq.ft./year)	118.83	25.22	74.00	142.00
Electricity price* (pound/kWh)	0.08		-	
Minimum Wage (pound/hr)	4.41	25.22	4.10	4.85
Number of stores:				
Asda	272.5	16.5	258	296
Morrisons	213.8	189.6	113	498
Safeway*	481.0		-	
Sainsbury's	567.8	117.5	463	727
Tesco	1,591.3	581.4	728	1,779
Average sales area per store (per 1,000 sq.ft.):				
Asda	46.06	1.48	44.40	47.42
Morrisons	33.16	5.42	25.90	36.21
Safeway*	21.86		-	
Sainsbury's	27.68	3.94	22.52	30.99
Tesco	15.71	6.85	11.02	25.85
Number of employees (both full-time and part-time):				
Asda*	120,000		-	
Morrisons	70,756	46,924	42,959	140,901
Safeway*	98,000		-	
Sainsbury	146,450	1,485	145400	147,500
Tesco**	221,000		-	

Notes: * one observation for year 2003 only. ** one observation for year 2004 only.
Sources: Various. Plastic index (ONS website), gasoline index (EIA website), Energy usage (UK Green Lib Dem's report), electricity price (BERR website), minimum wage (BERR website) number of stores and average sales area (CC's report), number of employees (companies' own annual report and website)

Table 2.14. Supermarket's Operational Costs by Year

Firm	Asda		Morrisons		Safeway		Sainsbury's		Tesco	
Year	Labour	Elec.	Labour	Elec.	Labour	Elec.	Labour	Elec.	Labour	Elec.
2002	0.88	0.0003	0.88	0.0003	1.54	0.0005	1.55	0.0005	0.98	0.0003
2003	0.91	0.0003	0.91	0.0002	1.59	0.0005	1.60	0.0005	1.02	0.0003
2004	0.95	0.0003	0.95	0.0003	1.66	0.0006	1.67	0.0006	1.06	0.0003
2005	0.98	0.0003	0.98	0.0005	1.72	0.0010	1.72	0.0010	1.09	0.0006

Notes: All figures are in pence per litre. Labour and electricity cost are from Auther's own approximation. Labour = cost of labour, Ener. = cost of electricity.

2.5.2 Manufacturers' Marginal Costs (MC^w)

Guidelines for the approximation of MC^w are obtained from Smith and Thanassoulis (2008). These authors gathered their information through consultations with the MDC and interviews with dairy processors' executives.

The major dairy processors–i.e. Arla, Dairy Crest, Wiseman and Express–are responsible for acquiring unprocessed milk from dairy farmers, processing them and delivering them to each supermarket outlet. In 2002, their marginal cost could be broken down into: cost of raw milk 0.17-0.20 pounds per litre (ppl), processing (0.04 ppl), bottling (0.048 ppl), transportation (0.045 ppl) and other (0.005 ppl). To make them comparable with the time period in this study, the monthly cost of plastic index is used to inflate the processing, bottling and other costs. On the other hand, the weekly UK gasoline index is used to inflate the transportation cost. The plastic index is obtained from the ONS website, while the UK gasoline index is obtained from the Energy Information Administration, US (EIA).

As for the cost of raw milk paid by dairy processors to farmers, I use the farmgate price estimated by DEFRA and MDC. This is the price farmers receive from selling milk to any buyer at their farm. Therefore, it does not include any transportation cost. The

manufacturers' marginal cost can, therefore, be expressed as follows:

$$MC^w = \text{ cost of raw milk + processing + bottling + transportation + other}$$

Figure 2.13 plots my approximations of MC^w againsts weekly period from week 41 of 2002 to week 40 of 2005. One scatter plot uses the actual farmgate price as the cost of raw milk. Another uses the Actual Milk Price Equivalent (AMPE) as the cost of raw milk. AMPE is estimated by the MDC as the price UK farmers are paid for their raw milk. However, rather than collecting the price information directly from the farmers, the MDC infers it from market prices of butter and skim milk powder (SMP). That is, cost of raw milk = price of butter and SMP - manufacturers' profit - costs not paid to farmers. We can see in figure 2.13 that both measures of MC^w are reasonably close, especially after early 2004. However, since AMPE is obtained through an indirect calculation which is subject to measurement errors of the manufacturers' profit and costs, I prefer using the actual farmgate price.

2.6 Supply

The previously obtained estimates of demand elasticities and marginal costs will be used to generate a menu of outcomes under different vertical arrangements and horizontal competition scenarios. Specifically, I will predict a menu of retailers' prices (P) and manufacturers' wholesale prices (p^w) that are consistent with different competition scenarios. With knowledge of the demand curve, I will also be able to calculate the equilibrium quantity, consumers' welfare, producers' welfare and the deadweight loss.

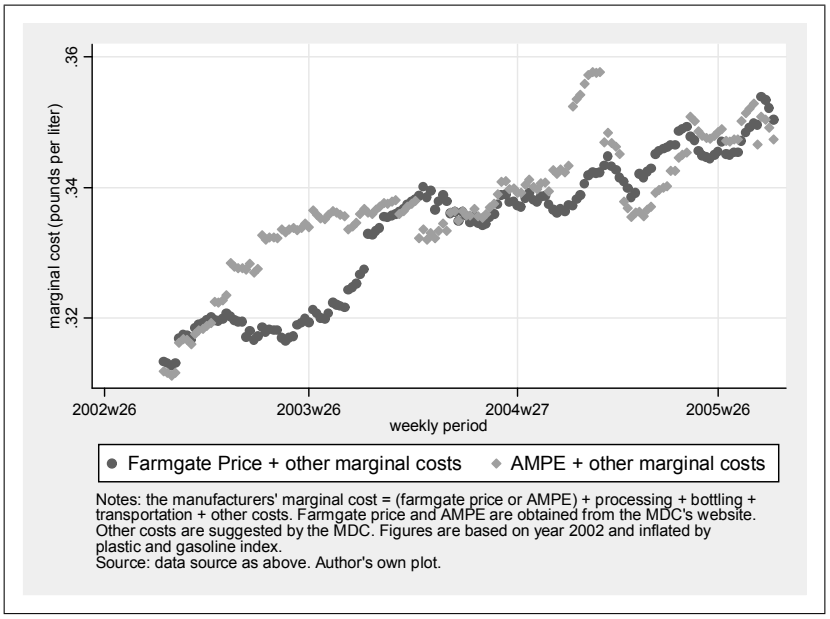


Fig. 2.13. Manufacturers' Average Marginal Cost.

For each scenario I will assume 1) a vertical contract between the manufacturers and the retailers, 2) a horizontal relation among the manufacturers, and 3) a horizontal relation among the retailers. Two types of vertical contracts and three types of horizontal competition will be examined. For vertical contracts, I will examine the linear pricing contract and the efficient non-linear pricing contract. I assume that every pair of manufacturer and retailer adopt the same type of vertical contract. For horizontal competition, I will examine the cases of perfect Bertrand, noncooperative Cournot, and joint-monopoly. The result will, therefore, be a menu of P, p^w and other outcomes under 18 possible vertical-horizontal scenarios.

Since the real P is observed and p^w can be approximated directly, I will use them to compare with different outcomes in the menu and find the closest match. It is worth noting that I assume the linear demand curve for the rest of the chapter. This is because my

observations capture a very small range of price variation–i.e. $P \in [0.44, 0.50]$ ppl. This gives little evidence of what the real shape of the demand curve looks like. However, if it is believed that the demand curve is convex to the origin, a linear demand curve can provide a reasonable lower bound. That is, it gives the lowest possible price for a given quantity level and the minimum deadweight loss under a given competition scheme, etc.

Next, I discuss how the market outcomes are calculated under different vertical and horizontal schemes.

2.6.1 Linear Pricing Vertical Contract

The linear pricing verical contract assumes that the manufacturers first set a price per unit of milk. This price could be equal to or above their marginal cost according to how competitive they are. Then, retailers use this price to decide on how many units they would like to buy. This practice is often called "double marginalization".

Figure (2.14) illustrates the demand curve and the marginal cost faced by the retailers. For simplicity, I assume a constant marginal cost. The figure shows that, because the manufacturers are charging more than their marginal cost, the retailers' marginal cost becomes higher than the manufacturers' marginal cost. As long as the manufacturers' unit price exceeds their marginal cost, or as long as they are not perfectly competitive, the socially desirable quantity (Q_2) will not be achieved.

In what follows, I derive the calculation of the retailers' price (P) and the manufacturers' price (p^w) under the linear vertical arragement and three different horizontal competition schemes–perfect Bertrand, joint-monopoly, and noncooperative Cournot.

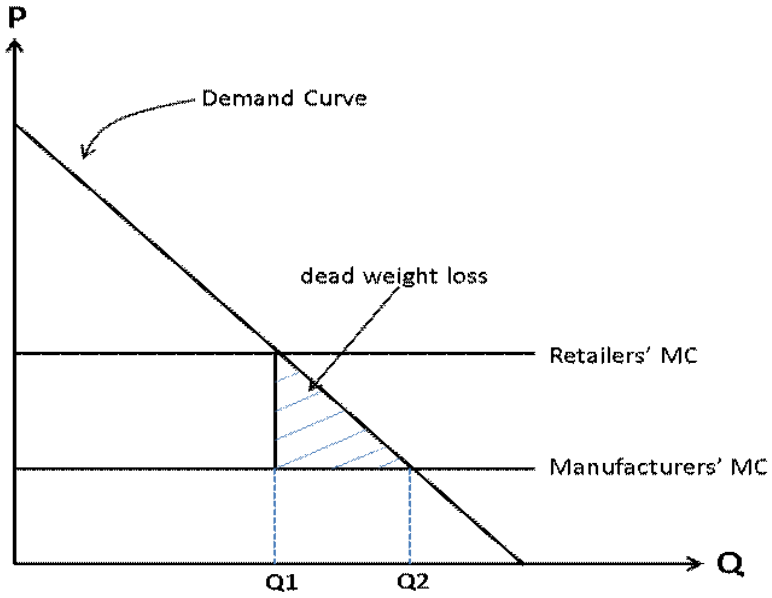


Fig. 2.14. Linear Pricing Illustration

Retailers' Price (P_t)

- Perfect competition (perfect Bertrand): in a perfect Bertrand competition, where no firm has a cost advantage and there is no capacity constraint at the industry level, each firm makes zero profit and charges at a price equal to marginal cost. Therefore, we can write:

$$P_t = MC_t^r = p_t^w + c_t^r. \tag{2.23}$$

- Joint-monopoly: firms first choose the quantity that maximizes their joint profit. They then, set the price accordingly. Given $\partial Q/\partial P$ = the slope of the demand curve,

the joint-monopoly quantity supplied by the retailers would be:

$$\begin{aligned} \arg\max_{Q_t} \Pi_t &= (P_t - MC_t^r)Q_t \\ &\Rightarrow \quad Q_t = -\frac{1}{2}(\alpha - MC_t^r)\frac{\partial Q}{\partial P}, \end{aligned} \tag{2.24}$$

- Noncooperative Cournot competition (or Nash-Cournot): under certain assumptions, the standard Cournot model suggests that the only source of market share asymmetry is the marginal cost asymmetry. Such assumptions are 1) firms are choosing their optimal quantity in a static one-shot game; 2) the cost functions of all firms are common knowledge; 3) there is no capacity constraint at both the firm level and the industry level and, 4) firms produce only one product and compete in one market. Since none of the assumptions is met in this case, it is not necessarily true that the only source of market share asymmetry is the difference in marginal costs. Other factors can also explain the differential in firms' market shares; for example, the number of periods during which firms compete, uncertainty about firms' cost functions, capacity constraints, government regulatory constraints, firms' level of patience (value of their discount rates), etc.

  A few attempts have been made to relax unrealistic assumptions of the model. Puller (2006) allows firms to compete repeatedly. In his paper, firms maximize the discounted value of the stream of profits. Collusion can be sustained if and only if the net present value of cheating exceeds the net present value of colluding. In addition, Puller (2007) incorporates firm-level capacity constraints.

Surprisingly, when the firm-level marginal costs are not observed, both cases suggest the same empirical implimentation, which is to include firm dummies in the supply relation.

In the context of the UK supermarkets, the idea of firm-level capacity constraint provides a means to account for the market share differential. In 2005, Tesco, Asda, Morrisons and Sainsbury's held 28.98, 15.84, 12.26 and 11.78 percent of the market share respectively. These numbers are reasonably proportional with the firms' total sales areas, which were 2,287$\times 10^3$, 1,179$\times 10^3$, 1,158$\times 10^3$ and 1,521$\times 10^3$ sq.ft. respectively (CC, 2008). Therefore, I adopt Puller (2007)'s method to use firm-level capacity constraints in explaining the asymmetry of market shares. Let firm i's best response function be expressed as:

$$\begin{aligned} \arg\max_{q_{it}} \Pi_{it} &= (P_t(Q_t) - MC_t^r)q_{it} \qquad s.t.\ q_i \leq k_i. \\ &\Rightarrow \quad q_{it} = -\frac{1}{2}\left((\alpha - MC_t^r - \hat{\lambda}_{it})\frac{\partial Q}{\partial P} + \sum_{j \neq i} q_{jt}\right). \end{aligned} \tag{2.25}$$

where q_{it} is the quantity supplied by retailer i, MC_t^r is the marginal cost, $\hat{\lambda}_{it}$ is the estimated shadow value of the capacity constraint. I use an iterative process to solve for each firm's best response quantity. In the first round, I assume the symmetric Cournot output for each q_{jt} $(\forall\ j \neq i)$. Then, calculate the best response of firm i (q_{it}) using eq. (2.25). After that, I substitute q_{it} in the best response functions of all other firms $(\forall\ j \neq i)$ and calculate the new values of q_{jt} $(\forall\ j \neq i)$. This process is repeated until no firm has an incentive to change their quantity.

The estimated shadow value $\hat{\lambda}_{it}$ in eq.2.25 is obtained from the following estimation:

$$\arg\max_{q_{it}} \Pi_{it} = (P_t(Q_t) - MC^r_{it})q_{it} \qquad s.t.\ q_i \leq k_i.$$

solving the first order condition yields:

$$\begin{aligned} 0 &= P_t + (q_{it}\frac{\partial P_t}{\partial Q_t}\frac{\partial Q_t}{\partial q_{it}}) - MC^r_{it} - \lambda_i \\ P_t - MC^r_t &= \lambda_i - q_{it}\frac{\partial P_t}{\partial Q_t}\frac{\partial Q_t}{\partial q_{it}} \end{aligned}$$

this gives the following empirical specification of the cournot competition supply relation:

$$P_t - MC^r_t = \Sigma_j \lambda_i firm_i + \zeta q_{it} + \omega_{it} \tag{2.26}$$

where $firm_j$ is a firm dummy which equals one for firm j and zero otherwise, ω_{it} is the error component, λ_i and ζ are parameters to be estimated. The estimation results are reported in table 2.15.

Manufacturers' Price (p^w_t)

- Perfect competition (perfect Bertrand) – assumes price equals marginal cost. Therefore, for manufacturers, we can write:

$$p^w_t = MC^w_t$$

where subscript t denotes time period and MC^w_t is the marginal cost of the manufacturers.

- Joint-monopoly – similar to the retailers' case, the optimality condition of a joint-monopoly manufacturers can be shown as:

$$\begin{aligned} \arg\max_{Q_t^w} \Pi_t &= (p_t^w - MC_t^w)Q_t^w \\ &\Rightarrow \quad Q_t^w = -\frac{1}{2}(\alpha - c_t^r - MC_t^w)\frac{\partial Q}{\partial p^w} \end{aligned}$$

where Q_t^w is the total wholesale quantity supplied by the manufacturers. Here, $\frac{\partial Q}{\partial p^w}$ is the slope of the demand curve faced by the manufacturers. This slope depends on the competitive conduct of the retailers. For perfectly competitive, perfectly collusive and Cournot retailers, $\frac{\partial Q}{\partial p^w}$ can be derived from equation (2.18), (2.19) and (2.20) respectively.

- Cournot competition – the best response function of a manufacturer k in a noncooperative Cournot competition can be derived as follows:

$$\begin{aligned} \arg\max_{q_{kt}^w} \Pi_{kt} &= (p_t^w - MC_t^w)q_{kt}^w \\ &\Rightarrow \quad q_{kt}^w = -\frac{1}{2}\left((\alpha - \bar{c}_t^r - MC_t^w - \lambda_k)\frac{\partial Q}{\partial p^w} + \sum_{l \neq k} q_{lt}^w\right) \end{aligned}$$

where q_{kt}^w is the wholesale quantity supplied by manufacturer k, $\hat{\lambda}_k$ is the estimated shadow value of the capacity constraint. I estimate $\hat{\lambda}_k$ in the same fashion as in the retailers' case (see eq.2.26). As previously discussed, $\frac{\partial Q}{\partial p^w}$ depends on the competitive conduct of the retailers. To solve for the equilibrium values of q_{kt}^w and all q_{lt}^w, I use the same iteration process as that of the retailers.

Table (2.15) reports the estimated value of $\hat{\lambda}_i$ and $\hat{\lambda}_k$ according to equation (2.26). The results suggest that Tesco tends to own the most market share, followed

by Asda, Sainsbury, Safeway and Morrisons. As for the manufacturers, Arla tends to own the most market share, followed by Dairy Crest, Wiseman and Express. Although these results are obvious, the magnitudes of $\hat{\lambda}_i$ and $\hat{\lambda}_k$ helps increase the accuracy of quantity estimation under the noncooperative Cournot scheme.

Table 2.15. Asymmetric Capacity Constraints and Marginal Costs

Retailers					Manufacturers		
Dependent Variable = $MC_t^r - P_t$					Dependent Variable = $MC_t^w - p_t^w$		
vertical contract	Linear		Non-linear			Linear & Non-linear	
merger timing[a]	pre	post	pre	post		pre	post
Explanatory Variable (parameter):					Explanatory Variable (parameter):		
q_{it}	-0.187 (0.002)	-0.219 (0.003)	-0.230 (0.002)	-0.255 (0.003)	q_{kt}	$-.0401$ (0.001)	$-.031$ (0.001)
$firm_i$					$firm_k$		
i = Morrisons	0.058 (0.001)	0.024 (0.007)	0.071 (0.001)	0.028 (0.008)	k = Arla	-0.0073 (0.0006)	0.0031 (0.0010)
i = Safeway	0.050 (0.001)	–	0.062 (0.001)	–	k = DairyC.	-0.0006 (0.0007)	-0.0058 (0.0008)
i = Sainsbury's	0.019 (0.001)	0.023 (0.007)	0.024 (0.001)	0.028 (0.006)	k = Express	0.0009 (0.0007)	–
i = Tesco	-0.054 (0.001)	-0.085 (0.012)	-0.066 (0.002)	-0.099 (0.003)			
i = Asda	(benchmark)				k = Wiseman	(benchmark)	
no. of Obs	400	380	400	380		240	288
adjusted-R^2	0.989	0.780	0.992	0.779		0.940	0.865

Summary statistics:

	pre-merger			post-merger		
	mean (std.dev)	min	max	mean (std.dev)	min	max
MC_t^w	0.351 (0.008)	0.313	0 .340	0.342 (0.004)	0.334	0.354
p_t^w	0.345 (0.008)	0.334	0.351	0.358 (0.006)	0.351	0.367
c_t^r	0.011 (0.003)	0.008	0.017	0.011 (0.003)	0.009	0.017
P_t	0.451 (0.010)	0.436	0.464	0.478 (0.018)	0 .462	0.505
MC_t^r	$= p_t^w + c_t^r$ under linear contract $= MC_t^w + c_t^r$ under non-linear contract					

Notes: These are estimation results for equation 2.26.

[a]For the retailers, Morrisons completed the acquisition of Safeway in February 2004. For the manufacturers, Asda completed the acquisition of Express in October 2003.

2.6.2 Non-Linear Pricing Vertical Contract

In this study, the term "non-linear pricing" refers to an efficient two-part tariff contract. Under this contract, the manufacturers set their price equal to their marginal cost ($p_t^w = MC_t^w$). They then, extract profit as a lump sum transfer from the retailers. Given any horizontal relation among the manufacturers and among the retailers, this two-part tariff vertical arrangement is more efficient than the linear pricing vertical arrangement. It results in a higher quantity supplied, thus consumers' surplus, to the consumers for any given horizontal relation (Tirole, 2003) (Villas-Boas, 2007).

Retailers' Price (P_t)

The retailers' price is calculated in the same way as in the linear pricing case. However, rather than using $MC_t^r = p_t^w + c_t^r$ as the retailers' marginal cost, we use the following:

$$MC_t^r = MC_t^w + c_t^r$$

where MC_t^w is the manufacturers' marginal cost. If the manufacturers are not making a loss, this marginal cost should not exceed the marginal cost under the linear contract case ($MC_t^w \leq p_t^w$). The retailers' prices under perfect Bertrand, joint-monopoly and Cournot competition are calculated using equations (2.23), (2.24) and (2.25) respectively.

Manufacturers' Price (p_t^w)

Under the efficient two-part tariff assumption, however, one cannot evaluate horizontal competition among the manufacturers in the same fashion as under the linear pricing model. This is because the manufacturers do not directly influence the retail price and/or

quantity. The per-unit price the retailers pay to the manufacturers is equal to the manufacturers' marginal cost. The manufacturers gain profit by extracting a lumpsum amount from the retailers.

One way to evaluate the manufacturers' market power is through analyzing the proportion of profit they are able to extract from the retailers. Smith and Thanassoulis (2008) uses a bargaining model to evaluate the relative strength of market power between the dairy farmers vs. manufacturers and between the manufacturers vs. retailers. Analyzing the bargaining power, however, is beyond the focus of this paper.

If we assume that the efficient non-linear vertical arrangement is adopted and that the retail price is set by the supermarkets, the final price and quantity would be independent of the manufacturers' horizontal competition scheme. Therefore, the success of overall collusion can be evaluated by looking at the retailers' price alone.

2.6.3 Perfect Vertical Collusion Model

So far, we have only discussed the non-vertically integrated models. In reality, it is also possible that the manufacturers and retailers are vertically integrated and, altogether, maximize the joint-profit. A successful collusion would yield the same result as that of the non-linear vertical arrangement with joint-monopoly retailers. This is because the vertically-integrated entity would consider the manufacturers' marginal cost as their marginal cost. Then, they would set the price or quantity to maximize a joint-profit. Therefore, the vertically-integrated's marginal cost is:

$$MC_t = MC_t^w + c_t^r$$

where MC_t^w is the manufacturers' marginal cost and c_t^r is the retailers' marginal cost. The vertically-integrated entity maximizes the joint-profit as follows:

$$\begin{aligned} \arg\max_{Q_t} \Pi_t &= (P_t - MC_t)Q_t \\ &\Rightarrow \quad Q_t = -\frac{1}{2}(\alpha - MC_t)\frac{\partial Q}{\partial P}. \end{aligned} \tag{2.27}$$

Again, the corresponding price can be calculated using the demand function eq.2.17.

2.7 Empirical Results

2.7.1 Menu of Outcomes

This section reports the menu of price, quantity, retailers' profit, manufacturers' profit and the lower bound for the deadweight loss under different vertical and horizontal competition scenarios. Variable p_w denotes the average wholesale price of milk. As discussed above, my model can only predict p_w under the linear vertical contract case. Variable P denotes retail price charged by the retailers, Q^{Big4} denotes the total quantity (in million litres per year) supplied by the Big4 supermarkets, Q^{Total} denotes the total quantity supplied by all the retailers in the UK, Π^m denotes the total annual profit earned by the manufacturers, Π^r denotes total annual profit earned by the retailers, DWL denotes the deadweight loss (in pounds per year) according to each scheme. The deadweight loss is zero under the perfect Bertrand retailers with a non-linear vertical arrangement scenario. Finally, the actual (or observed) values of all the variables are reported in the last row.

Pre- and Post-Merger Outcomes

During the observation period, Oct 2002 - Sep 2005, there occurred two mergers. One between two manufacturers, namely Arla and Express in October 2003; another between two retailers, namely Morrisons and Safeway in February 2004. The Arla-Express merger reduced the number of main dairy processors from four to three, while the Morrisons-Safeway merger reduced the number of main supermarkets from five to four. For comparison purposes, I report the pre-merger values in table (2.16) and post-merger values in table (2.17). The value of p_w and Π^m are reported according to the Arla-Express merger timing, whereas the rest is reported according to the Morrisons-Safeway merger timing.

To be concise, let the notation {a, b, c} = {vertical arrangement, manufacturers' horizontal competition, retailers' horizontal competition}. The manufacturers' horizontal competitions cannot be identified under the non-linear vertical arrangement, and therefore are not reported.

In both pre- and post-merger cases, scenario {non-linear, – , Bertrand} yields the lowest retail price (P), which are 0.357 pounds per litre (ppl) pre-merger and 0.371 ppl post-merger. In contrast, the scenario {linear, monopoly, monopoly} yields the highest retail prices, which are 1.297 ppl pre-mergers, and 1.301 ppl post-mergers. The actual prices fall between those predicted by {linear, Bertrand, Bertrand} and {non-linear, – , Cournot}, e.g. $P^{Actual} = 0.451 \in [0.357, 0.562]$ pre-mergers and $P^{Actual} = 0.478 \in [0.371, 0.598]$ post-mergers. The actual prices are lower by far than those predicted schemes with joint-monopoly manufacturers and/or joint-monopoly retailers. In addition, since the

actual post-mergers price still falls within the same range as the pre-merge price, there is no evidence of a more successful collusion after mergers.

Turning to welfare results, the last three columns of tables (2.16) and (2.17) report firms' profits and deadweight loss predicted by the linear demand function. It is shown that the non-linear vertical arrangement consistently gives a lower deadweight loss for any given horizontal competition scheme. The most efficient scenario is {non-linear, – , Bertrand}, where $DWL = 0$. The most inefficient scenario is {linear, monopoly, monopoly}, where $DWL = 1739.20$ million pounds per annum (mpa) pre-merger and $DWL = 1833.33$ mpa post-merger. Post-mergers profits are not necessarily higher for all parties (in nominal terms). The manufacturers' profit decreased from 62.05 to 49.46 mpa, while the retailers' profit increased from 235.01 to 307.38 mpa. The actual deadweight loss (DWL) is 27.04 mpa pre-merger and 28.72 mpa post-merger. The DWL increased by 6.22 percent after mergers.

Pre- and Post-Farmers' Direct Actions Outcomes

Having examined the outcome of mergers, we now turn to the outcome of farmers' direct actions. During the observation period, the price of major liquid milk products underwent two increases. Both were due to farmers' successful direct actions. The first one was in July 2003, and second one was in March 2005. This section investigates whether these farmers' direct actions helped increase the manufacturers' and retailers' profits.

Tables (2.18), (2.19) and (2.20) report the menu of outcomes before first, between the first and the second, and after second farmers' direct action respectively. For purpose of comparison, the last column reports the farmers' price (p^{farm}) estimated by the department

for environment food and rural affairs, UK (DEFRA). The p^{farm} increased from 0.174 ppl to 0.184 ppl (5 percent increase) after the first direct action, but fell to 0.183 (0.5 percent decrease) after the second direct action. From this informal investigation, it does not appear that the farmers benefited from their direct actions.

Similar to the mergers case, the scenario {non-linear, – , Bertrand} predicts the lowest price; while scenario {linear, monopoly, monopoly} predicts the higest price. The actual average price increases every time there is a farmers' direct action. However, it still falls between those predicted by {linear, Bertrand, Bertrand} and {non-linear, – , Cournot}. Despite the increase in price, there is no evidence that farmers' direct actions are associated with more collusive outcomes than those resulting from non-coorporative Cournot competition.

Turning to welfare results, the manufacturers' profit was 60.83 mpa before the first direct action, then fell to 48.91 (19.6 percent decrease) after the first direct action, and increased to 55.44 (13.36 precent gain) mpa after the second direct action. On the other hand, the retailers' profit consistently increased from 216.10 mpa to 269.06 mpa (31.7 percent gain) and to 361.85 (34.5 percent gain) mpa respectively. The DWL increased from 26.82 mpa to 29.33 mpa after the first direct action, then fell to 27.33 after the second direct action. Such a small change in the DWL is mainly due to the inelasticity of demand, $(\eta_{mm}) \approx 0.403$.

Thus, this study finds that the farmers' direct actions in July 2003 and March 2005 were more likely to benefit the manufacturers and the retailers than the farmers. The in-

crease in the price of retail milk, pressured by the farmers' direct actions, did not necessarily pass to the farmers themselves.

Table 2.16. Menu of Outcomes (Pre Mergers)

Linear pricing vertical contract								
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	Π^m	Π^r	DWL
Bertrand	Bertrand	0.320 (0.005)	0.357 (0.008)	2707.78	4237.50	0.00	0.00	1.04
	Cournot	0.320 (0.005)	0.577 (0.007)	2233.23	3411.09	0.00	494.12	111.70
	Monopoly	0.320 (0.005)	0.986 (0.004)	1353.89	1879.79	0.00	853.12	798.76
Cournot	Bertrand	0.579 (0.005)	0.609 (0.033)	2164.32	3291.10	560.65	0.00	142.93
	Cournot	0.579 (0.005)	0.759 (0.022)	1843.69	2732.75	477.60	309.70	338.81
	Monopoly	0.579 (0.005)	1.107 (0.016)	1094.31	1427.74	283.47	565.12	1121.27
Monopoly	Bertrand	0.967 (0.004)	0.978 (0.004)	1371.73	1910.85	887.78	0.00	778.60
	Cournot	0.967 (0.004)	1.094 (0.003)	1121.41	1474.93	725.77	129.71	1085.05
	Monopoly	0.967 (0.004)	1.297 (0.002)	685.87	716.46	443.89	218.23	1739.20
Efficient non-linear pricing vertical contract (two-part tariff)								
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL
n/a	Bertrand	–	0.335 (0.009)	2754.93	4319.61	0.00		0.00
	Cournot	–	0.562 (0.007)	2266.63	3469.26	522.17		97.95
	Monopoly	–	0.975 (0.004)	1377.47	1920.84	886.82		772.17
Successful vertical integration								
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL
Joint-monopoly		0.320 (0.005)	0.975 (0.004)	1377.47	1920.84	886.82		772.17
Actual values		0.345 (0.008)	0.451 (0.010)	2474.94	3867.50	62.05	235.01	27.04

Notes: p^w and P are wholesale and retail prices reported in pounds per year, standard deviations are reported in parentheses. Q^{Big4} is quantity supplied by the Big4, Q^{Total} is total retail quantity (excluding milkman) reported in million litres per year. Π^m is total manufacturers' profit, Π^r is total retailers' profit reported in million pounds per year. DWL is the society's deadweight loss reported in million pounds per year. Source: Author's own calculation.

Table 2.17. Menu of Outcomes (Post Mergers)

Linear pricing vertical contract								
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	Π^m	Π^r	DWL
Bertrand	Bertrand	0.341 (0.005)	0.371 (0.006)	3052.32	4560.11	0.00	0.00	0.66
	Cournot	0.341 (0.005)	0.613 (0.005)	2458.77	3572.11	0.00	640.43	138.01
	Monopoly	0.341 (0.005)	0.993 (0.003)	1526.16	2019.70	0.00	978.13	837.07
Cournot	Bertrand	0.658 (0.004)	0.670 (0.004)	2318.11	3337.96	734.05	0.00	205.63
	Cournot	0.658 (0.004)	0.817 (0.005)	1957.87	2738.32	619.98	290.04	440.05
	Monopoly	0.658 (0.004)	1.137 (0.002)	1173.30	1432.33	371.53	549.33	1255.57
Monopoly	Bertrand	0.974 (0.003)	0.986 (0.003)	1545.17	2051.34	977.62	0.00	816.92
	Cournot	0.974 (0.003)	1,104 (0.003)	1253.95	1566.58	793.37	149.54	1152.46
	Monopoly	0.974 (0.003)	1.301 (0.002)	772.59	765.31	488.81	243.84	1833.33
Efficient non-linear pricing vertical contract (two-part tariff)								
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL
n/a	Bertrand	–	0.353 (0.005)	3095.41	4631.840	0.00		0.00
	Cournot	–	0.598 (0.005)	2495.54	3633.31	612.58		122.55
	Monopoly	–	0.984 (0.003)	1547.71	2055.56	978.33		814.25
Successful vertical integration								
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL
Joint-monopoly		0.341 (0.005)	0.984 (0.003)	1547.71	2055.56	978.33		814.25
Actual values		0.359 (0.006)	0.478 (0.018)	2844.40	4174.00	49.46	307.38	28.72

Notes: p^w and P are wholesale and retail prices reported in pounds per year, standard deviations are reported in parentheses. Q^{Big4} is quantity supplied by the Big4, Q^{Total} is total retail quantity (excluding milkman) reported in million litres per year. Π^m is total manufacturers' profit, Π^r is total retailers' profit reported in million pounds per year. DWL is the society's deadweight loss reported in million pounds per year. Source: Author's own calculation.

Table 2.18. Menu of Outcomes (Before the First Farmers' Direct Action)

Linear pricing vertical contract									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	Π^m	Π^r	DWL	p^{farm}
Bertrand	Bertrand	0.318 (0.002)	0.352 (0.009)	2717.38	4254.90	0.00	0.00	1.03	
	Cournot	0.318 (0.002)	0.574 (0.007)	2240.30	3423.98	0.00	492.07	112.70	
	Monopoly	0.318 (0.002)	0.984 (0.004)	1358.69	1888.49	0.00	855.47	804.16	
Cournot	Bertrand	0.577 (0.002)	0.588 (0.002)	2209.34	3370.05	571.37	0.00	126.32	
	Cournot	0.577 (0.002)	0.774 (0.006)	1873.71	2785.48	484.57	292.46	323.77	
	Monopoly	0.577 (0.002)	1.096 (0.001)	1117.16	1467.81	288.92	567.41	1103.46	
Monopoly	Bertrand	0.966 (0.004)	0.976 (0.004)	1376.69	1919.83	892.16	0.00	783.76	
	Cournot	0.966 (0.004)	1.093 (0.004)	1125.12	1481.67	729.13	129.14	1092.85	
	Monopoly	0.966 (0.004)	1.296 (0.002)	688.34	720.95	446.08	218.82	1751.24	
Efficient non-linear pricing vertical contract (two-part tariff)									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL	
n/a	Bertrand	–	0.328 (0.002)	2770.23	4346.95	0.00		0.00	
	Cournot	–	0.556 (0.003)	2278.47	3490.46	516.01		96.99	
	Monopoly	–	0.972 (0.001)	1385.12	1934.51	889.12		774.29	
Successful vertical integration									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL	p^{farm}
Joint-monopoly		0.318 (0.002)	0.972 (0.001)	1385.12	1934.51	889.12		774.29	
Actual values		0.343 (0.009)	0.442 (0.003)	2474.94	3867.50	60.826	216.10	26.82	0.179

Notes: p^w and P are wholesale and retail prices reported in pounds per year, standard deviations are reported in parentheses. Q^{Big4} is quantity supplied by the Big4, Q^{Total} is total retail quantity (excluding milkman) reported in million litres per year. Π^m is total manufacturers' profit, Π^r is total retailers' profit reported in million pounds per year. DWL is the society's deadweight loss reported in million pounds per year. Source: Author's own calculation.

Table 2.19. Menu of Outcomes (Between the First and the Second Farmers' Direct Action)

Linear pricing vertical contract									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	Π^m	Π^r	DWL	p^{farm}
Bertrand	Bertrand	0.337 (0.006)	0.365 (0.004)	2999.37	4522.94	0.00	0.00	0.61	
	Cournot	0.337 (0.006)	0.597 (0.015)	2441.71	3586.49	0.00	608.30	125.80	
	Monopoly	0.337 (0.006)	0.990 (0.002)	1499.68	2004.60	0.00	962.88	832.05	
Cournot	Bertrand	0.641 (0.029)	0.653 (0.029)	2309.32	3364.18	678.09	0.00	187.52	
	Cournot	0.641 (0.029)	0.798 (0.026)	1961.52	2780.14	575.97	284.10	408.14	
	Monopoly	0.641 (0.029)	1.129 (0.014)	1168.08	1447.76	342.99	555.76	1228.56	
Monopoly	Bertrand	0.973 (0.003)	0.983 (0.003)	1517.65	2034.77	948.34	0.00	812.77	
	Cournot	0.973 (0.003)	1.100 (0.004)	1235.93	1561.69	772.30	143.67	1141.16	
	Monopoly	0.973 (0.003)	1.299 (0.002)	758.83	760.52	474.17	239.22	1824.12	
Efficient non-linear pricing vertical contract (two-part tariff)									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL	
n/a	Bertrand	–	0.347 (0.006)	3042.02	4594.56	0.00		0.00	
	Cournot	–	0.584 (0.014)	2474.91	3642.25	555.19		112.25	
	Monopoly	–	0.981 (0.003)	1521.01	2040.41	946.39		809.19	
Successful vertical integration									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL	p^{farm}
Joint-monopoly		0.329 (0.008)	0.981 (0.003)	1521.01	2040.41	946.39		809.19	
Actual values		0.354 (0.003)	0.463 (0.003)	2749.16	4087.33	48.907	269.06	29.33	0.193

Notes: p^w and P are wholesale and retail prices reported in pounds per year, standard deviations are reported in parentheses. Q^{Big4} is quantity supplied by the Big4, Q^{Total} is total retail quantity (excluding milkman) reported in million litres per year. Π^m is total manufacturers' profit, Π^r is total retailers' profit reported in million pounds per year. DWL is the society's deadweight loss reported in million pounds per year. p^{farm} is the DEFRA's estimates of farmers' price reported in pound per liter.

Source: Author's own calculation.

Table 2.20. Menu of Outcomes (After the Second Farmers' Direct Action)

Linear pricing vertical contract									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	Π^m	Π^r	DWL	p^{farm}
Bertrand	Bertrand	0.348 (0.003)	0.377 (0.001)	3080.98	4569.87	0.00	0.00	0.71	
	Cournot	0.348 (0.003)	0.618 (0.003)	2482.01	3579.38	0.00	642.96	138.36	
	Monopoly	0.348 (0.003)	0.997 (0.001)	1540.49	2022.42	0.00	982.02	836.73	
Cournot	Bertrand	0.662 (0.002)	0.674 (0.002)	2342.43	3348.56	737.24	0.00	204.69	
	Cournot	0.662 (0.002)	0.819 (0.002)	1981.47	2751.66	623.63	288.30	436.28	
	Monopoly	0.662 (0.002)	1.139 (0.001)	1186.08	1436.35	373.30	551.75	1252.24	
Monopoly	Bertrand	0.976 (0.001)	0.987 (0.001)	1563.48	2060.44	983.13	0.00	812.66	
	Cournot	0.976 (0.001)	1.106 (0.001)	1268.93	1573.36	797.92	149.49	1147.63	
	Monopoly	0.976 (0.001)	1.302 (0.001)	781.74	767.71	491.57	245.17	1828.25	
Efficient non-linear pricing vertical contract (two-part tariff)									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL	
n/a	Bertrand	–	0.359 (0.003)	3127.97	4647.57	0.00		0.00	
	Cournot	–	0.602 (0.003)	2521.97	3645.45	612.81		121.75	
	Monopoly	–	0.987 (0.001)	1563.99	2061.28	982.23		812.13	
Successful vertical integration									
Manuf.	Retailers	p_w	P	Q^{Big4}	Q^{Total}	$\Pi^m + \Pi^r$		DWL	p^{farm}
Joint-monopoly		0.348 (0.003)	0.987 (0.001)	1563.99	2061.28	982.23		812.13	
Actual values		0.367 (0.003)	0.501 (0.003)	2935.50	4263.00	55.442	361.85	27.33	-

Notes: p^w and P are wholesale and retail prices reported in pounds per year, standard deviations are reported in parentheses. Q^{Big4} is quantity supplied by the Big4, Q^{Total} is total retail quantity (excluding milkman) reported in million litres per year. Π^m is total manufacturers' profit, Π^r is total retailers' profit reported in million pounds per year. DWL is the society's deadweight loss reported in million pounds per year. p^{farm} is the DEFRA's estimates of farmers' price reported in pound per liter.

Source: Author's own calculation.

2.7.2 Encompassing Test

While the menu of outcomes provides an intuitive but informal assessment of different competition scenarios, the encompassing test provides a formal assessment. The rationale behind the test is that, if one model is said to be correctly specified, it should be able to entirely and exclusively explain the outcome. In reality, it may be difficult to obtain and justify the "most" correct model. However, by employing this test, I hope to be able to rank the competition scenarios used in the menu of price.

I follow an approach employed by Bushnell et al.(2008), which is an adaptation of the J-test proposed by Davidson and MacKinnon (1993, p.386-387). For each pair of rival models, I run the following regression:

$$P_t^{Actual} = \alpha_1 P_t^A + \alpha_2 P_t^B + u \tag{2.28}$$

where P_t^{Actual} is the actual price of milk at time t, P_t^A is the price predicted by scenario A at time t, P_t^B is the price predicted by scenario B at time t, and u is the error component. The null hypotheses of the test are 1) $H_0 : \alpha_1 = 1$ and 2) $H_0 : \alpha_2 = 0$. In other words, only P_t^A explains P_t^{Actual}. Therefore, if the null hypotheses are *not* rejected, we cannot reject scenario A. On the other hand, if the null hypotheses *are* rejected, we cannot accept scenario A. The original Davidson and MacKinnon's version of the J-test suggests P_t^B be the predicted value of P_t^{Actual} under the alternative model. In my case, the predicted prices under different competition scenarios are the models themselves. Thus, there is no need to obtain the predicted value of P_t^B.

As discussed in Bushnell et al.(2008), equation (2.28) is estimated using the OLS estimation with Newey-West standard errors (Newey and West, 1987). This accounts for the presence of heteroskedasticity and autorrelation of the error terms. I allow the errors to be autocorrelated up to 4 lags. Allowing for more prolonged autocorrelation does not change the results.

In any case, both P_t^A and P_t^B are calculated using econometrically estimated variables– e.g. the shadow values of the capacity contraints $\hat{\lambda}_{it}$, the slope and the intercept of the demand curve. The fact that these values already have standard errors themselves makes the usual (including Newey-West) standard error calculation invalid. To this effect, in addition to the Newey-West standard errors, I employ the two-step error correction method proposed by Murphy and Topel (1985). This two-step error correction method is also used in Bushnell et al. (2008). See appendix 2.A for the detail discussion. It is worth noting that the bootstrap errors can be used as a relatively computational-light alternative.

Table (2.21) reports the p-values associated with each pairwise encompassing test. The first column reports the null scenarios (P^A), while the second row reports the alternative scenarios (P^B). For the abbreviations of different scenarios, the first letter represents the manufacturers' competition scheme, while the second letter represents the retailers' competition scheme. For example, "bc" stands for a linear vertical contract between Bertrand manufacturers and Cournot retailers {linear, Bertrand, Cournot}, "cm" stands for a linear vertical contract between Cournot manufacturers and joint-monopoly retailers {linear, Cournot, Monopoly}, etc. The *non-linear* vertical arrangements with Bertrand, Cournot and joint-monopoly retailers are represented by "nl_b", "nl_c" and "nl_m" respec-

tively. Finally, "m" represents the efficient vertically integrated model where, as discussed in section 2.5.3, all the outcomes are identical to those of the "nl_m" case.

There are two p-values for each pair of scenarios being tested (table 2.21). The top values are p-values for the first null hypothesis $H_o : \alpha_1 = 1$, the bottom values are p-values for the second null hypothesis $H_o : \alpha_2 = 0$. I test each hypothesis separately for the sake of informativeness. The last column reports the number of null hypotheses that I fail to reject. There are two numbers. The first number represents the total *single* null hypotheses not rejected. For example, the testing of $H_o : \alpha_1 = 1$ in $P_t^{Actual} = \alpha_1 P_t^{bb} + \alpha_2 P_t^{bc} + u$ would be counted as one; the testing of $H_o : \alpha_2 = 0$ for the same regression would be counted as another one. Therefore, for each P^A scenario, the maximum number of unrejected hypotheses would be 24. On the other hand, the second number represents the total number of *both* null hypotheses not rejected. For each pair of competing models, both $H_o : \alpha_1 = 1$ and $H_o : \alpha_2 = 0$ has to be unrejected to be counted as one. The maximum number of unrejected double hypotheses is, therefore, 12.

Comparing among rows and using the 5 percent significant level, "bc" {linear, Bertrand, Cournot} and "nl_c" {non-linear, – , Cournot} are the two most consistent null hypotheses. The scenario "bc", has the highest number of single and double unrejected null hypotheses. It does not reject 19 out of 24 single null hypotheses, and 8 out of 12 double null hypotheses. When either one or both of the null hypotheses are rejected, the alternative scenario are "bb", "cb", "cc", and "nl_b".

Another most consistent scenario is "nl_c". This scenario does not reject 16 out of 24 single null hypotheses, and 7 out of 12 double null hypotheses. Either one or both of the

null hypotheses are rejected when the alternative scenario are "bb", "cb", "cc", "nl_b", and "bc".

Since there is no scenario P^A that does not reject all of the null hypotheses, I cannot claim that I have found "the model" actually adopted by the manufacturers and the retailers. However, it is possible to claim that firms under this study have achieved no higher profits than the noncooperative static level. This is because both "bc" and "nl_c" belong to the noncooperative static game setting, and are only rejected by noncooperative alternative scenarios (i.e. "bb", "cb", "cc", and "nl_b").

2.8 Conclusions

In 2005, the UK was ranked number 8 in the world in per-capita consumption of liquid milk (Goff, 2006). This was ahead of Australia, Canada, Germany, France, New Zealand and the United States. According to the MDC, total retail milk consumption (excluding the milkman) was 3,814 million litres per year in 2002 and had grown to 4,263 million litres per year by 2005. This means only a couple of pence increase in price could result in a sizable loss of the aggregated consumers' welfare. Nevertheless, in response to the farmers' direct actions, retail milk price was twice adjusted upward during the period of October 2002 to September 2005 alone. On both occasions, the Big4 supermarkets acted almost simultaneously and set identical prices for most major liquid milk products. This practice urged the OFT to put a charge against the major UK dairy processors and supermarkets for their price-fixing behaviour. In response to the case, some parties admitted involvement in

Table 2.21. P-values of the pairwise encompassing test

null scenario* (P^A)	alternative scenarios (P^B)													H_0 not
	bb	bc	bm	cb	cc	cm	mb	mc	mm	nl_b	nl_c	nl_m	m^a	rejected
														single, double
bb	-	0.28 0.09	0.00 0.07	0.57 0.01	0.43 0.04	0.10 0.56	0.01 0.15	0.00 0.07	0.00 0.09	0.00 0.00	0.92 0.27	0.01 0.15	0.01 0.15	12, 3
bc	0.02 0.08	-	0.98 0.50	0.00 0.19	0.02 0.39	0.98 0.52	0.97 0.54	0.93 0.43	1.00 0.46	0.00 0.00	0.37 0.14	0.97 0.54	0.97 0.54	**19, 8**
bm	0.00 0.00	0.00 0.00	-	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.18 0.05	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	2, 1
cb	0.00 0.00	0.00 0.00	0.00 0.00	-	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0, 0
cc	0.00 0.00	0.00 0.00	0.00 0.02	0.97 0.08	-	0.00 0.11	0.00 0.03	0.00 0.05	0.00 0.05	0.00 0.00	0.00 0.00	0.00 0.03	0.00 0.03	5, 1
cm	0.00 0.00	0.00 0.00	0.44 0.16	0.00 0.01	0.00 0.03	-	0.72 0.19	0.71 0.07	0.77 0.10	0.00 0.00	0.00 0.02	0.72 0.19	0.72 0.19	12, 6
mb	0.00 0.00	0.00 0.00	0.25 0.04	0.00 0.00	0.00 0.00	0.00 0.01	-	0.01 0.00	0.00 0.00	0.00 0.00	0.00 0.01	-^{b}	-^{b}	1, 0
mc	0.00 0.00	0.00 0.00	0.00 0.02	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	-	0.17 0.08	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	2, 1
mm	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.01 0.04	-	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0, 0
nl_b	0.56 0.03	0.67 0.01	0.02 0.81	0.00 0.99	0.00 0.49	0.04 0.61	0.02 0.65	0.01 0.68	0.01 0.65	-	0.27 0.55	0.02 0.65	0.02 0.65	13, 1
nl_c	0.00 0.00	0.00 0.00	0.76 0.63	0.00 0.10	0.00 0.14	0.75 0.90	0.83 0.77	0.99 0.43	0.91 0.65	0.00 0.00	-	0.83 0.77	0.83 0.77	**16, 7**
nl_m	0.00 0.00	0.00 0.00	0.25 0.04	0.00 0.00	0.00 0.00	0.00 0.01	-^{b}	0.02 0.01	0.00 0.00	0.00 0.00	0.00 0.01	-	-	2, 0
m^a	0.00 0.00	0.00 0.00	0.25 0.04	0.00 0.00	0.00 0.00	0.00 0.01	-^{b}	0.02 0.01	0.00 0.00	0.00 0.00	0.00 0.01	-	-	2, 0

* the top p-value to test H_o: $\alpha_1 = 1$, the bottom p-value is to test H_o: $\alpha_2 = 0$, a null hypothesis is rejected at the 5 percent level of significance.
^{a}: the vertically integrated outcomes are identical to those of the {non-linear, –, monopoly} scenario.
^{b}: predicted prices under the mb and nl_m scenarios are perfectly correlated

anti-competitive practices–mostly exchange, or facilitating the exchange of commercially sensitive information in respect of the retail price.

Despite the admitting of anti-competitive practices and the exchange of information, the success of collusion is not guaranteed. This chapter evaluates the extent to which the dairy processors and supermarkets succeed in collusion. I adopt the menu approach to generate price, quantity, firms' profits, and the lower bound of the society's deadweight loss under different competitive conducts. Three types of vertical relationships are examined, namely linear (or double marginalization), efficient non-linear, and successful vertical integration. For horizontal competition among the manufacturers and among the retailers, I examine the cases of the perfect Bertrand, noncooperative Cournot and joint-monopoly scenarios. The perfect Bertrand and noncooperative Cournot prices serve as lower and upper bound for prices under non-coorporative competition schemes. The joint-monopoly price serves as the upper bounds for prices under collusive competition schemes.

The menu approach suggests that, although prices and firms' profits have been growing through time, they still fall between the outcomes generated by {linear, Bertrand, Bertrand} and {non-linear, – , Cournot} scenarios. Therefore, the menu approach does not find any evidence of firms achieving more than the noncoorporative Cournot level. In addition to the menu approach, I adopt the encompassing test for a more formal assessment.

The encompassing test suggests that "bc" or {linear, Bertrand, Cournot} is the most consistent with the observed price; while "nl_c" or {non-linear, – , Cournot} is the second most consistent scenario. Thus, the formal encompassing test also suggests no more success in collusion than at the noncoorporative Cournot level.

Apart from examining the overall success of collusion, I assess the outcomes associated with mergers and farmers' direct actions. During the sample period, there occured two merger incidents– Arla acquired Express Milk in 2003 and Morrisons acquired Safeway in 2004. The menu of outcomes shows that mergers resulted in a significant decrease in the manufacturers' profits, and a significant increase in the retailers' profits. Total manufacturers' profits dropped by at least 20.3 percent from 62.05 to 49.46 million pounds per annum (mpa), while total retailers' profits increased by at least 30.8 percent from 235.01 to 307.38 mpa. The deadweight loss increased by at least 6.2 percent from 27.04 to 28.72 mpa. The relatively small magnitude of the deadweight loss is mainly due to the low elasticity of demand.

Turning to the outcomes of the farmers' direct actions, my data show that the retail milk price increased significantly in response to both of the farmers' direct actions. However, data from DEFRA show that these increases in price did not pass to the farmers. The average farmgate price increased by 5 percent after the first direct action and decreased by 0.5 percent after the second direct action. The manufacturers did not seem to benefit much either. Their profit dropped by 19.6 percent after the first direct action, then increased by 13.36 percent after the second direct action. The retailers benefited most from the direct actions. The Big4's total profit was raised by 31.7 percent after the first direct action and 34.5 percent after the second direct action.

In sum, despite the admitting of anti-competitive practices, this study could not find sufficient evidence to support that the dairy processors and supermarkets could achieve more than the equivalent of static noncoorporative profits. The encompassing test suggests

that the {linear, Bertrand, Cournot}and {non-linear, – , Cournot} are the scenarios most consistent with the actual market outcomes. The Arla-Express merger in 2003 did not increase the manufacturers' profit, while the Morrisons-Safeway merger in 2004 increased the retailers' profit. The first direct action did not increase the manufacturers' profit, but increased the retailers' profit. The second direct action increased both the manufacturers' and retailers' profits. However, the significant increase in price after the farmers' direct actions did not pass to the farmers.

Last but not least, it is important for the readers to keep in mind some crucial unrealistic assumptions, which may have affected the results in this paper. Because of limited data, supermarkets in my context are assumed to compete in a single-product market under a static setting. The consideration of the multi-product and the dynamic nature of the problem will be subject of future research.

2.A Sequential two-step error correction

Murphy and Topel (1985)'s Discussion

In a case where one or more variables used for the estimation is itself an estimated parameter from the previous stage, the OLS or a usual method of error calculation would be invalid. To this light, the two-step error correction method can be adopted to obtain asymtotically consistent standard errors. Taking from Murphy and Topel (1985) directly, let the second-step model be expressed as:

$$y = x_2\beta + f(\theta, x_1)\alpha + u,$$

where y is the dependent variable, x_2 is a $1 \times k$ vector of exogenous variable, x_1 is a vector of imputed variables from the first step, β and α are parameters to be estimated, $f(\cdot)$ is a function of the unknown parameter θ appears in the first step. The entire $f(\cdot)$ can represent a vector contains values imputed from the first step estimation. The random component has the following properties:

$$\begin{aligned} E(u|x_1, x_2) &= 0 \\ E(u^2|x_1, x_2) &= \sigma^2 \end{aligned}$$

Murphy and Topel (1985) suggest that the more complicated error structures would not affect the analytical procedure and results. They also assume that x_1 and x_2 are non-stochastic, and that $f(\cdot)$ be twice continuously differentiable in θ. If the estimated value of θ from the first step is denoted $\hat{\theta}$, then, a consistent $\hat{\theta}$ has the following property:

$$\sqrt{n}(\hat{\theta} - \theta) \backsim n(0, V(\hat{\theta})).$$

This implies that $(\hat{\theta} - \theta)$ is asymptotically normal with a covariance matrix $V(\hat{\theta})$.

Now, let X_2 denote the $n \times k$ matrix of observations on x_2 and let F denote the $n\times m$ matrix of $f(\hat{\theta}, x_1)$.The vector of all the second-step regressors can be expressed as $Z = (X_2, F)$. To derive the asymptotically consistent error components, the estimation has to first satisfy the following requirements:

1. Z satisfies $\lim_{n\to\infty} \frac{1}{n} Z'Z = Q_0$, where Q_0 is positive definite and symmetric.

2. $f(\theta, x_1)$ is twice continuously differentiable in θ for each x_1.

3. $\hat{\theta}$ is a consistent estimator of θ. This makes the least square estimator of β and α consistent.

4. The second derivatives of $f(\cdot)$ satisfie $\lim_{n\to\infty} \frac{1}{n} Z'F^* = Q_1$ and $\left|\frac{\partial^2 f(\hat{\theta}, x_1)}{\partial\theta_i \partial\theta_j}\right|^2 \leq M_2(x_1)$ for $\left\|\hat{\theta} - \theta\right\| < \delta$, where $\lim_{n\to\infty} (\frac{1}{n}) \Sigma_{i=1}^{n} M_2(x_{1i}) < \infty$, Q_1 is positive definite and symmetric, and F^* is the Taylor series first-order approximation to F.

5. The sequence of estimates $\hat{\theta}$ is independent of the random component in the second step.

If all the above mentioned conditions are satisfied, then the second-step least square estimator would be consistent with the following property:

$$\sqrt{n}[(Z'Z)^{-1}Z'Y - (\beta', \alpha')'] \backsim n(0, \Sigma),$$

where

$$\Sigma = \sigma^2 Q_0^{-1} + Q_0^{-1} Q_1 V(\hat{\theta}) Q_1' Q_0^{-1}$$

The covariance matrix can be expressed as:

$$\hat{\Sigma} = \hat{\sigma}^2(Z'Z)^{-1} + (Z'Z)^{-1}Z'F^*\hat{V}(\hat{\theta})F^{*\prime}Z(Z'Z)^{-1}$$

where $\hat{\sigma}^2$ is the covariance matrix obtained in the second step. The matrix F^* is the first-order approximation of the matrix F.The element ij of the matrix F^* is defined as $f^*_{ij} = \Sigma^n_{k=1}\hat{\alpha}_k \frac{\partial f_k}{\partial \theta_j}(\hat{\theta}, x_{1i})$. Subscript i indices each parameter in the parameter vector θ, whereas subscript j denotes each regressor in the second step. In the case where there is only one imputed variable and $f(\cdot)$ is linear in θ and x_1, the covariance matrix can be simplified as:

$$\hat{\Sigma} = \hat{\sigma}^2(1 + (\hat{\alpha}^2\frac{\hat{\sigma}^2_\varepsilon}{\hat{\sigma}^2}))(Z'Z)^{-1}.$$

where $\hat{\sigma}^2_\varepsilon$ is the variance of the only imputed regressors in the auxiliary model $f(\cdot)$.

Implementation

To implement the Murphy and Topel (1985) two-step error correction, I adopt the strategy proposed by Bushnell et al.(2008). Let θ be a vector of parameters associate with all imputed regressors in the model, θ_i denotes element i in the vector. I first obtain $f^*_{ij} = \Sigma^n_{k=1}\hat{\alpha}_k \frac{\partial f_k}{\partial \theta_j}(\hat{\theta}, x_{1i})$ using the following steps:

1. For a competition scheme A, let $h^{*A}_i = \partial P^A/\partial\hat{\theta}_i = (P^A(1.001 \times \hat{\theta}_i) - P^A(\hat{\theta}_i))/(1.001 \times \hat{\theta}_i - \hat{\theta}_i), \forall i$.

2. Then, compute $f^*_i = \hat{\alpha}_1 h^{*A}_i + \hat{\alpha}_2 h^{*B}_i, \forall i$, where B denotes the alternative competition model. The parameters $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are those obtained from the pre-correction second-step estimation (In this case, eq.2.28).

   After having obtained f^*_i, proceed to:

3. Regress f_i^* on h_i^{*A} and h_i^{*B}, i.e. $f_i^* = \delta_{i1} h_i^{*A} + \delta_{i2} h_i^{*B}$ and obtain the estimates $\hat{\delta}_{i1}$ and $\hat{\delta}_{i2}$.

4. A matrix with $\hat{\delta}_{ij}$ serves as the element in row i and column j is called matrix $(Z'Z)^{-1}Z'F^*$.

5. A covariance matrix associates with the parameter vector $\hat{\theta}$ (obtained from the first step estimation) is called $\hat{V}(\hat{\theta})$.

6. Calculate $(Z'Z)^{-1}Z'F^*\hat{V}(\hat{\theta})F^{*\prime}Z(Z'Z)^{-1}$ using the products from 4. and 5.

7. Let $\hat{\sigma}^2$ be the pre-correction variance-covariance matrix of estimated parameters obtained from the second step.

8. Calculate $\sqrt{\hat{\sigma}^2(Z'Z)^{-1} + (Z'Z)^{-1}Z'F^*\hat{V}(\hat{\theta})F^{*\prime}Z(Z'Z)^{-1}}$. The diagonal elements are the standard errors.

Chapter 3
Estimating Agglomeration Incentives: The case of the Texas Lodging industry

Abstract

This chapter investigates the agglomeration in quality-type among hotels in Texas. I adopt a discrete choice framework which allows for strategic interactions among firms. That is, a firm's decision could influence other firms' decisions, and vice-versa. After controlling for demand and location-specific effects through area-specific characteristics and area-specific unobserved heterogeneity, the incentive to agglomerate in quality choice still exists. This suggests that hotels of the same quality type enjoy positive spillovers from the presence of one another, while high-quality hotels enjoy more positive spillovers than low-quality hotels counterpart.

3.1 Introduction and Summary

The phenomenon of firms' agglomeration in product type is prevalent in the global economy. Expensive fashion shops, cheap or expensive restaurants, discount electronic stores, etc. are often found in a cluster rather than a stand-alone unit. This fact seems to conflict with the conventional product differentiation wisdom, according to which firms should supply products that consumers find less substitutable in order to soften competition (Shaked and Sutton, 1982). Why do firms still agglomerate? What are the motives that make firms act in opposition to the rational product differentiation behaviour? In response to these questions, this chapter proposes empirical strategies to analyse firms' product type decisions.

To explore the motive of firms' agglomeration in product type, some previous literatures suggest that firms in certain types of industry benefit from proximity to others. As an aggregate unit, they may attract more customers and enjoy demand spillovers if consumers normally bear high search costs (Fischer and Harrington, 1996). They may also benefit from technological spillovers if the industry is subject to high research and development (Krugman, 1991). Finally, they may free-ride on the area development/investment made by others (Enz et al., 2005). These authors claim that, when such benefit from clustering outweighs the loss from intense competition, firms are better off joining a cluster rather than standing alone. This type of motive, the "positive strategic effects", is hereafter called the "strategic effects".

While the strategic effects play a prominent role in some agglomeration cases, they may not exist in others. For example, whether hotels along the Miami Beach, ski resorts

in the Alps, and travel agencies in downtown London benefit from having the same type of competitors around is still questionable. To this effect, some researchers claim that the observed agglomerations in product type are simply driven by favourable location and demand conditions (Hotelling, 1929) (Mazzeo, 2002). I call this type of motive the "location-specific effects". According to Krugman(1991) and Ellison and Glaeser(1997), the geographical concentration of firms can be explained by strategic effects, or location-specific effects, or both effects combined.

This chapter proposes empirical strategies to measure the importance of strategic effects and location-specific effects in determining firms' decisions to cluster in the product type space. The data set is gathered from the Texas lodging industry over the time period 1998-2004. The analysis is performed in two parts. Part one adopts a non-parametric technique to confirm that hotels in Texas are clustered in quality type. Part two develops empirical models to estimate the strength of spillover effects and location-specific effects in determining hotels' decisions to cluster (in the quality type).

I focus on hotels' quality choice taking entry and location as given. This assumption may sound unconventional at first. However, I believe it is the case for a large number of hotels in Texas. First of all, about 57.98 percent of hotels in Texas were not affiliated with a chain (see table 3.24). This type of hotels was usually run by local residences who may have owned the land even before they decided to build a hotel (and choose the quality). Given this, I believe that it is reasonable to assume that hotel owners made their decisions on entry and location prior to quality. Second of all, even if the rest of the hotels were affiliated with a chain, most of them were owned by independent owners. For example,

in 2009, only 16 out of 4,294 InterContinental Hotel Group (IHG)[25] hotels were owned by the company itself. About 3,700 were IHG franchisees, and about 578 hired IHG to manage (IHG, 2009). For these independent owners of branded hotels, it is also possible that their location choices were limited, i.e. close to their home, or on a piece of land that they already owned. Therefore, it is not totally unrealistic to assume that entry and location choices are made prior to quality.

To confirm that hotels in Texas are clustered in quality type, I use a non-parametric testing method called the Multinomial Test for Agglomeration and Dispersion (MTAD) developed by Rysman and Greenstein(2005). In particular, the MTAD is capable of testing whether the observed choices are more agglomerated or dispersed than they would be had the agents made their choices randomly and independently of one another. Despite such capability, the MTAD cannot determine the rationale behind the outcomes. For example, given that there is an agglomeration in quality, the MTAD cannot determine further whether this outcome is driven by strategic effects or simply by demand. In this light, I pursue the second part of the analysis.

The second part of the chapter uses two empirical approaches to analyse determinants of hotels' quality choices. This is where the aforementioned "strategic effects" and "location-specific effects" are quantified. The strategic effects are accounted for by including the number of each type of competitor hotels in the market. Since the geographical distance of competitors is likely to influence the strength of competition, I adopt the distance-band categorisation proposed by Seim (2006). On the other hand, the location-specific ef-

[25] IHG is a hotel development group which operates seven hotel brands, i.e. InterContinental, Crowne Plaza, Hotel Indigo, Holiday Inn, Holiday Inn Express, Staybridge Suites and Candlewood Suites.

fects are accounted for by including relevant area-specific variables. Two empirical strategies are used to implement the estimation.

The first empirical strategy focuses on decisions made by new entrants, taking the pre-entry incumbents' decisions as given. It assumes that firms are short-sighted. That is, any future actions made by future entrants do not matter. The pre-entry information, e.g. decisions made by the incumbents, are assumed to be exogenous. The post-entry actions are assumed to have no effect on firms' choices. Although these assumptions are unrealistic, they make the estimation feasible. As long as the timing of entry is observed, the researcher can employ standard single-equation econometric techniques to analyse firms' choices. Sorensen (2006) uses this method to analyse health insurance choices made by employees of universities in California.

The second empirical strategy uses the static-Bayesian game framework to relax the short-sighted assumption. This framework assumes that firms can form expectations of what other firms would do in the equilibrium. Therefore, what influences a firm's decision is its expectation of all firms' actions in the equilibrium, not just the incumbents' actions. To implement this, I take a snapshot of the market and assume that it is the result (equilibrium) of all firms' simultaneous decisions. In this static-Bayesian game, a firm's action can influence other firms' actions and vice-versa. Therefore, to arrive at consistent estimates, the estimation needs to account for the simultaneity of firms' decisions.

A fair number of papers have suggested solutions to the simultaneity problem. The first strand of solutions writes out the entire system of simultaneously determined decisions. It then, uses a fixed-point algorithm to derive the equilibrium solution(s). This method is

computationally demanding and could result in multiple sets of solutions–in other words, multiple equilibria. Therefore, additional equilibrium selection constraints may have to be imposed. Some examples of the use of the fixed-point algorithm are Seim (2006), Orhun (2006), and Jia (2008).

As an alternative to the fixed-point algorithm, one could adopt the two-step approach. With this approach, the researcher first estimates a policy function to obtain expected values of the strategic variables. Since the policy function uses only exogenous variables to estimate the strategic variables, the estimated strategic variables would not suffer the simultaneity bias problem. In the second step, the researcher replaces the strategic variables with the estimated values obtained from the first step. This method is more computationally tractable, and also gives consistent estimates (Bajari et al., 2006). Moreover, by construction, this method studies only one equilibrium. Therefore, the researcher does not need to specify equilibrium selection constraints. Examples of the two-step approach are Bajari et al. (2006) and Ellickson and Misra (2008). This chapter employs the two-step approach for its computational tractability.

Apart from the simultaneity bias, my estimation may be subject to the unobserved heterogeneity problem. This is because I may not be able to obtain all the exogenous variables that actually influence firms' quality choice. Such variables are, for example, building regulations that prevent the establishment of large luxurious hotels, a run-down environment that is not attractive for an investment of high-quality hotels, etc. If these variables are omitted, and are correlated with variables of interest, my estimation results would be inconsistent. To this effect, my empirical models include proxy variables for

the market-specific unobserved heterogeneity. It also include a market-specific random intercept.

To summarise the empirical findings, the MTAD indicates that lodging establishments in Texas are agglomerated in quality. After controlling for location-specific effects (i.e. area-specific characteristics and area-specific unobserved heterogeneity), the incentive to agglomerate in quality choice still exists. This suggests that hotels of the same quality type enjoy positive spillovers from the presence of one another. Such positive spillovers could come from the reduction in consumers' search cost (Fischer and Harrington, 1996), firms' learning about the market-specific unobserved profitability (Toivanen and Waterson, 2005), for example. More specifically, my results are consistent with those of Kalnins and Chung (2004) and Enz et al. (2005), where they found positive spillovers from high-quality hotels; and with Baum and Haveman (1997), where they found positive spillovers among hotels within the same price range. In any case, identifying the exact sources of the spilloves is beyond the focus of this chapter.

To discuss the benefits from this study, I highlight that analysis of the location-specific effects versus the strategic effects of product agglomeration has welfare and policy implications. In particular, it can help policymakers analyse whether the current provision of product variety is optimal. If firms' decisions are largely based on the strategic and not the location-specific effects, the agglomeration incentive may result in too few product varieties. The overall welfare can be improved by encouraging more product differentiation. An exception, however, applies to situations where positive spillovers include better information about the demand.

Lastly, it is important to stress that this study takes entry and location choice as given. While it is more realistic to allow firms to make joint decisions on entry-location-quality, it is unlikely that the qualitative findings on quality choice would change. On the other hand, such a comprehensive model would demand a highly informative data set and complex computational techniques. To the best of my knowledge, no work has been done on firms choosing entry, location and quality at the same time. Quite a few papers focus on the joint entry-location issue, while a much smaller number analyse the joint entry-quality decision. Holmes (2008) uses a dynamic model to analyse joint entry-location decision of Wal-Mart stores. Then, employs the moment inquality technique to analyse the data. Jia (2008) uses a two-period model to analyse joint entry-location decision of Wal-Mart and Kmart stores. For more examples on joint entry-location analysis, see Toivanen and Waterson (2005) and Seim (2006). To my best knowledge, the only work on joint entry-quality decision is Mazzeo (2002). This paper analyses the entry-quality decision of hotels along the US highway-exits using a static setting.

The remaining of this chapter proceeds as follows. Section 3.2 discusses related literature. Section 3.3 contains data sources and the industry background. Section 3.4 uses the MTAD to confirm the agglomeration among hotels of the same quality type. Section 3.5 further analyses whether the agglomeration is due to "strategic effects" or "location-specific effects". Finally, Section 3.6 concludes.

3.2 Literature Review

Empirical Test of Product Agglomeration and Dispersion

Three non-parametric methods have been widely used to test for independent-random choice decisions of agents: the Pearson's Chi-squared test of goodness of fit, the Dartboard index, and the Multinomial Test of Agglomeration and Dispersion (MTAD).

The first technique is the Pearson's Chi-squared test for goodness of fit. This method is one of a variety of chi-squared tests and was proposed by Pearson (1900). The Pearson's Chi-squared test generates a test statistic which compares frequencies of the observed events and that of a specified random distribution benchmark. It then evaluates whether the observed events are randomly distributed by reference to the chi-squared distribution. In general, the test statistic takes the following form:

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

where O is an observed frequency and E is an expected (random benchmark) frequency.

Existing literature applies the Pearson's Chi-squared test in various research contexts. For example, Belaire-Franch and Contreras (2002) use it to test for symmetry of the Spanish business cycle; while Edwards (2004) uses it to test for the independence of current account deficits reversals across regions in the world economy. Here, despite its popularity, the Pearson's Chi-squared test has not been adopted to test for product differentiation and agglomeration in recent industrial organization literature. This is because, while being able to test for non-random events, the Pearson's Chi-squared test cannot further suggest whether the non-random choices are driven by agglomerative or dispersive force. For this reason, the later two tests were developed.

Alternative to the Pearson's Chi-squared test, Ellison and Glaeser (1997) developed the Dartboard Index (γ^k), an index that increases with the degree of industry concentration. The formula of the dartboard index of industry k is:

$$\gamma^k \equiv \frac{\sum_{i=1}^{M}(s_i - x_i)^2 - \left(1 - \sum_{i=1}^{M} x_i^2\right)^2 H^k}{\left(1 - \sum_{i=1}^{M} x_i^2\right)(1 - H^k)},$$

where for each M geographical area (market), s_i is the employment share of industry k, x_i is the total employment, and H^k is the Herfindahl index $H^k = \sum_{j=1}^{N} z_j^2$ of the industry. Essentially, s_i and x_i can be calculated based on information other than employment share, so long as it compares the share of industry's business activities to that of the entire market.

Since the Dartboard index gives more flexibility in terms of measuring the degree of agglomeration, it has been used in much of the recent empirical industrial organization literature. For instance, Ellison and Glaeser (1997) use the Dartboard index to compare concentration across different industries in the US, while Ellickson (2004) uses the Dartboard index to find evidence against spatial differentiation in the US supermarket industry. Despite its flexibility relative to the Pearson's Chi-squared test, the Dartboard index is still limited to only the case where non-random events are driven by the agglomerative force. As argued by Rysman and Greenstein (2005), if the non-random outcome happens to be a result of the dispersive force, the Dartboard index cannot be used to analyse degree of dispersion.

To combine the merits of both the Pearson's Chi-squared test and the Dartboard index, Rysman and Greenstein (2005) developed the Multinomial Test of Agglomeration and Dispersion (MTAD). In addition to being able to both test the null hypothesis of

independent-random choice and measure the degree of agglomeration, the MTAD can also measure the degree of dispersion at its presence. This makes the MTAD the most complete method available for the test of agglomeration and dispersion to date. Due to all its merits, I will use the MTAD to test for quality and size agglomeration and dispersion among the lodging establishments in my study. The MTAD test statistic is given by:

$$t(\mathbf{X}, \mathbf{n}, \mathbf{p}) = l(\mathbf{X}, \mathbf{n}, \mathbf{p}) - E[l(f, \mathbf{p})],$$

where $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ is the observed likelihood of the events and $E[l(f, \mathbf{p})]$ is the expected likelihood of independent-random events. A comprehensive discussion on the MTAD can be found in section 3.4.

Since the MTAD is very new, there has been only one formal work so far to use this technique. In their analysis, Augereau et al. (2004) use MTAD to analyse the adoption pattern of 56K modems[26] by Internet Service Providers (ISPs) in the US. They find ISPs to be differentiated across the technologies rather than agglomerated in one technology or the other. This clear example of dispersive force can neither be measured by the Pearson's Chi-squared test nor the Dartboard Index.

Discrete Choice Studies with Strategic Decisions

When firms' decisions are influenced by other firms' decisions and vice-versa, this layer of complication has to be accounted for in the model. Draganska et al.(2008) gives a thorough summary of relevant issues on discrete choice models with strategic decisions. Their article suggests that theoretical models with strategic decisions have been presented

[26] There are two incompatible internet technologies which are available to ISPs. Each ISP can choose to adopt neither, only one, or both technologies.

in various fashions and can be classified into three dimensions: first, by the information context (complete vs. private); second, by the temporal context (static vs. dynamic); and third, by the timing of moves (simultaneous and sequential).

Much of the earlier work assumes complete information due to computational tractability. In a game with complete information, firms know the action of their competitors with certainly. Thus, standard discrete choice estimation techniques can be used to analyse the problem. Examples of this type of model can be seen in Bresnahan and Reiss (1991), Mazzeo (2002) and Dranove et al. (2003). As the complete information setting is unrealistic in many cases, later developments have allowed for private information. When firms have private information, their competitors' actions can no longer be known with certainty. To this effect, firms can only form their beliefs about their competitors' actions. Seim (2006) uses this strategy but assumes symmetric uncertainty across firms. Other literatures that also allow for private information are that of Orhun (2006), Ellickson and Misra (2008) and Vitorino (2007), for instance.

Apart from the information context, one has to specify the temporal context. Although in many cases firms clearly compete repeatedly and a dynamic assumption would be most appropriate, dynamic estimation still remains a great challenge to econometricians. To make estimation feasible, most cases assume static interactions. That is, firms are forward-looking and make decisions in the same period. Although this static assumption is restrictive, it can be justified if it reasonably approximate the steady state outcomes. Some of a large number of static game literature includes that of Mazzeo (2002), Seim (2006), Orhun (2006), Ellickson and Misra (2008), and Bajari et al. (2006). (Some of a few re-

cent developments on the dynamic setting are shown in Aguirrebabiria and Mira (2007) and Bajari et al. (2007))

Lastly, the timing of moves is a crucial element of a model. When players move simultaneously, the model assumes no updating of the information set. Thus, players do not learn from the revealed actions of others. However, when the players move sequentially, their information set can be updated by the observed actions of others. Sequential moves models usually require empirical information on the timing of players' actions, and are usually used in the context of social learning. For example Sorensen (2008) and Toivanen and Waterson (2005). Without information on the timing of players' actions, other work assumes simultaneous moves.

Along with model development, estimation techniques have been extensively advanced during recent years. This is due to the increased availability of the data, as well as the the advances in computational technology. The estimation approaches can be grouped into two categories–namely the nested fixed point method, and the two-step estimation method.

The nested fixed point method was developed first as an attempt to find solutions to a system of simultaneously determined equations, either in a static or dynamic setting. This approach is computationally intensive, especially in discrete choice cases where closed form solutions do not exist. Based on either the maximum likelihood estimation or the method of moments estimation, solutions have to be obtained through an iterative process. Often, additional equilibrium selection constraint(s) have to be used due to the existence of multiple equilibria. Applications of this method can be seen in Seim (2006) which

analyses entry in the US video rental industry; and Orhun (2006) which analyses entry in the US supermarket industry.

The two-step method was developed as a more computationally tractable alternative to the nested fixed point method. It was first introduced by Bajari et al. (2006). Rather than estimating the entire system of simultaneously determined equations, this method performs the estimation in two steps. The first step uses a policy function to obtain consistent estimates of firms' beliefs about their competitors' actions. Since these beliefs are estimated by exogenous variables only, they are not subject to the simultaneity bias. In the second step, these estimated beliefs are plugged in, and a standard discrete-choice estimation can be performed. Variations of the two-step estimation method has been adopted in many cases since it was introduced. Bajari et al. (2006) apply this method to a static game of analysts' stock recommendations. In addition, when the two-step estimation is performed continously until there is no improvement in the estimation criteria (e.g. when the value of the likelihood function cannot be increased), the estimation can be called the nested pseudo likelihood method. This estimation method is adopted, for example, by Judd and Su (2007), Vitorino (2007), Ellickson and Misra (2008).

In this study, I conduct and compare results from two models. The first model adopts the sequential decisions approach. Every new hotel makes a quality choice based on the observed choices made by the incumbents. I assume that all firms are myopic and thus do not consider subsequent entrants. The second model adopts the static game setting where, conditional on entry, firms simultaneously choose quality. Since firms' decisions are influenced by other firms' decisions and vice-versa, I use the two-step approach to remedy

the simultaneity bias. In both cases, the estimations are still subject to the unobserved heterogeneity problem. Therefore, econometric remedies have to be adopted for consistent estimation.

Hotel Agglomeration

In terms of the context, my work is the closet to that of Kalnins and Chung (2004). These authors use a data set of 570 new hotels in Texas during the time period 1992-1999. They employ the conditional logit model (McFadden, 1974) to analyse location choices of different types of new hotels–branded upscale, branded economy, and unbranded. Their assumption regarding the information available to firms is similar to the one I apply in the new hotels approach. That is, all firms are short-sighted and only decisions made by the incumbents matter. The empirical estimations use only county-level dummy variables to control for the unobserved heterogeneity, and do not allow for any strategic interactions among firms' choices.

Kalnins and Chung (2004) find that the presence of branded upscale incumbent hotels increases the probability to enter by all types of new hotels. However, it increases the probability to enter by the branded upscale type (the highest quality type) the most. In contrast, the presence of branded economy incumbent hotels decreases the probability to enter by branded upscale hotels, but has no significant effect on the probability to enter by branded economy and unbranded hotels. Finally, the presence of unbranded hotels has no significant effect on entry decisions of any type of new hotels. Hence, they conclude that branded upscale hotels create positive spillovers to all types of hotel; while the new en-

trants of the same quality-type benefit the most from these spillovers. In contrast, branded economy hotels create negative spillovers to new entrants of a better quality.

Using a data set containing 14,995 hotels across the US for the year 2000, Enz et al.(2005) arrive at a similar conclusion. The authors examine how the presence of different types of competitors affect hotels' revenue. They also discuss that high-quality hotels are generally differentiated in terms of products and services. Low-quality hotels, on the other hand, are less differentiated among themselves. Their results suggest that high-quality hotels generate positive spillovers and help increase the revenue of hotels within the area. In contrast, low-quality hotels generate negative impacts on the revenue of hotels nearby.

Regarding the empirical strategy, the authors adopt the feasible generalized least squares (FGLS) procedure, and regress revenue per available room (RevPAR) on a variety of explanatory variables. They use the "proportion of hotels in other segments", or of other quality type, to capture the spillover effects. In any case, the paper does not account for the potential simultaneity between RevPAR and "proportion of hotels in other segments"; hence, the "proportion of hotels in other segments" variable could be endogenous.

Baum and Haveman (1997) studies multi-dimensional differentiation of 614 hotels in Manhattan during the period 1989-1990. They use the three-stage least squares (3SLS) to analyse a system of simultaneously determined characteristics of each hotel–e.g. distance to other hotels, price range, and size. They find that hotels with similar price tend to locate close to each other, while hotels with comparable size tend to avoid being close to each other. In other words, the benefits from co-locating with hotels of similar price outweight

the costs. This, however, is not the case for size. The results imply that, if there is any benefit from co-locating with hotels of similar size, it would not outweight the costs.

In sum, the abovementioned literatures consistently found that both the location-specific and the strategic effects have significant impacts on hotels' entry and/or quality choice. In particular, Kalnin and Chung (2004) and Enz et al.(2005) found that high-quality hotels created positive spillovers to co-locating hotels. Baum and Haveman (1997) found that firms with similar price range enjoyed positive spillovers from one another.

My contribution to this body of literature is to incorporate more realistic assumptions in the analysis. Firstly, although strategic interactions among hotels' choices claim a credible role in explaining the outcomes, i.e. a firm's choice can influence its competitors' choices and vice-versa, the issue was not taken into consideration by any of the previous studies on hotels. In section 3.5.4, I adopt the static Bayesian approach to allow for strategic interactions among firms' quality choices. Secondly, none of these previous studies systematically accounted for market-specific unobserved heterogeneity. To this effect, I use market-specific proxy variables and a random intercept to account for it.

3.3 Data and Industry Background

3.3.1 The Texas Lodging Industry Overview

In 2003, the total visitor spending in Texas amounted to US$ 33.5 billion; this figure ranked as the third highest (6 percent) among all states in the US[27]. About 17 percent of it came from the lodging industry.

The huge amount of visitor spending in Texas creates almost half a million new jobs as well as thousands of millions of dollars in tax revenues every year. In terms of travellers' origins, the majority come from within the US, while about 10 percent come from overseas. Table (3.22) shows the amounts of spending, employment, and tax receipts generated by each type of traveller.

With a large number of travellers drawn to Texas every year, it is not surprising that they exhibit a variety preferences. To accommodate a wide range of preferences, there appears to be a large selection– ranging from the most basic to the most luxurious–of lodging services. According to the hotel ratings of the American Automobile Association (AAA), the quality of each establishment can be ranked from one-diamond to five-diamond: one-diamond represents those providing minimum-requirement service; two-diamond represents those providing more than the basic service but still at a comfortable price; three-diamond represents those providing a comprehensive service; four-diamond represents those that are aggressively more refined and stylish; and five-diamond represents those providing ultimate luxury, world-class service. The average proportions of 1-, 2-, 3-, 4-, and 5-diamond establishments operating during period 1998 - 2004 were about 60, 45, 25, 1, and 0.001 percent respectively. Table (3.23) reports the number of hotels operating dur-

[27] Florida had the highest share (12 percent), while California had the second highest (11 percent).

Table 3.22. Origins of Travelers, 2003 (predicted)

Origin	Spending ($Billion)	Employment (Thousand)	Total Tax Receipts ($Million)
Texas	17.1	211	2,222
Other US.	13.3	177	1,789
International	3.2	39	403
Total	33.5	427	2,414

Source: Texas Tourism Office of the Governor

ing the period 1998 - 2004 by their diamond assignment. Reported in parentheses are the percentages of brand-affiliated hotels.

The proportion of affiliated hotels has been growing throughout the past decade. In particular, about 39.5 percent of all hotels were brand-affiliated in year 1998; while this number grew into 42.02 percent in 2004. According to *Hotel Occupancy Tax Receipts* and the hotel brand report from Smith Travel Research[28], there appears to be 89 different hotel brands operating in Texas during the year 2004. Out of these large number of brands, none owned more than five percent of the market share. The top five most popular brands were Best Western® (4.06%), Holiday Inn® (3.77%), Comfort Inn® (3.07%), Super 8® (2.52%), and Days Inn® (2.38%)–with their market shares in brackets. See table (3.24) for more details about major hotel brands in Texas.

3.3.2 Hotel Location and Quality Choice

In light of the availability of many quality choices, it would be interesting to examine the quality-type configuration in each market, i.e. whether it is common to see all kinds of hotels *disperse* across markets or whether it is common to see the same quality-type of

[28] These data sources will be discussed in the Hotel Data section.

Table 3.23. Hotel quality breakdown by year

Year/Diamond	1	2	3	4	5	total
1998	2,390	577	938	40	7	3,952
	(6.36)	(83.02)	(94.88)	(85.00)	(85.71)	(39.50)
1999	2,438	617	991	40	8	4,098
	(6.56)	(84.12)	(95.66)	(82.50)	(87.50)	(40.78)
2000	2,461	627	1,047	40	8	4,183
	(6.58)	(83.89)	(95.89)	(82.50)	(87.50)	(41.41)
2001	2,512	638	1,073	42	5	4,270
	(6.81)	(83.70)	(95.43)	(88.10)	(80.00)	(41.45)
2002	2,580	662	1,110	43	5	4,400
	(6.71)	(83.53)	(94.95)	(86.05)	(80.00)	(41.39)
2003	2,555	683	1,123	43	5	4,409
	(6.81)	(84.04)	(94.84)	(88.37)	(80.00)	(42.07)
2004	2,639	699	1,147	41	5	4,531
	(7.01)	(83.69)	(95.47)	(85.37)	(80.00)	(42.02)

Notes: Author's tabulation and calculation based on Hotel Occupancy Tax Receipts data from Texas Comptroller's Office. Number in parentheses represents percentage of independent establishments (%).

hotels *agglomerate* in each market. However, since such an analysis involves using more advanced techniques than simply plotting graphs and constructing tables, I will postpone it to section 3.4 and section 3.5. For now, I suggest some simplification to facilitate the procedure.

The two simplifications to the Texas lodging data are: first, hotel quality is discretised into two groups; and second, a geographical market is defined as a 5-digit zip code area.

Regarding quality discretisation, I categorise lodging establishments in two quality groups – *low* and *high*. As such, the establishments with one and two diamonds are grouped as low-quality; while those with three, four, and five diamonds are grouped as high-quality. Table 3.25 displays two types of hotel rating–chain scale and diamond–from which I justify this categorisation. I obtained the chain scale information from Smith

Table 3.24. Major hotel chains and number of affiliated hotels in Texas, 2004

Chain name	Number in Texas	Percent of Total(%)	Average Star
Extended Stay America	20	0.44	1.95
Travelodge	21	0.46	1.95
Marriott	24	0.53	3.08
Country Inn & Suites	25	0.55	2.56
Intown Suites	27	0.60	2.00
Residence Inn	27	0.60	2.85
Executive Inn	29	0.64	1.17
Courtyard	30	0.66	3.00
Quality Inn	30	0.66	2.83
Best Value	31	0.68	1.94
Howard Johnson	31	0.68	2.00
Fairfield Inn	33	0.73	3.00
Econolodge	36	0.79	1.92
Hampton Inn	63	1.39	2.97
Ramada	70	1.54	2.66
Motel 6	80	1.77	1.23
La Quinta Inn	94	2.07	3.00
Days Inn	108	2.38	2.03
Super 8	114	2.52	2.07
Comfort Inn	139	3.07	2.90
Holiday Inn	171	3.77	2.91
Best Western	184	4.06	2.83
Other Chains	517	11.37	–
Independents	2627	57.98	1.09
Total	4,531	100	–

Sources: Author's tabulation based on the Hotel Occupancy Tax Receipts (2004) obtained from Texas Comptroller's Office and Hotel Brand report obtained from Smith Travel Research.

Table 3.25. Hotel rating comparison

Chain Scale	Diamond	Quality	Chain Examples
Independent	1,2	low	–
Economy Chains	1,2	low	Days Inn, Motel 8, Econolodge
Midscale w/o F&B Chains	3	high	La Quinta Inns, Fairfield Inn
Midscale w/ F&B Chains	3	high	Holiday Inn, Best Western
Upscale Chains	3	high	Residence Inn, Amerisuites
Upper Upscale Chains	3,4	high	Marriott, Westin, Hilton
Luxury Chains	4,5	high	Inter-Continental, Four Seasons

Sources: Tabulation based on Smith Travel Research's data for Chain Scale and chain example, AAA TourBook (2004 edition) for Diamond, and author's own categorisation for Quality.

Travel Research, while I acquired the diamond ratings from the AAA Texas TourBook (2004)[29]. I believe that the one- and two-diamond hotels are reasonably close substitutes because they fall into the same chain scales. On the other hand, I believe that the three-, four-, and five-diamond hotels are reasonably close substitutes because there are some overlaps between their chain scales. Within the context of two quality-group discretisation, I claim that the proposed strategy results in minimum within-group heterogeneity[30].

As for the market definition, I define a market as a 5-digit zip code area. Ideally, I prefer a market to be an area where only the competing firms are located. Due to uneven area sizes and the possibility of overlapping markets, however, pursuing such a definition involves careful sample selection. In earlier studies, Bresnahan and Reiss (1991) select only small isolated towns to rule out the possibility of overlapping markets, while Mazzeo (2002) selects only motel clusters located at highway exits for the same reason. Given that

[29] See the next section for more information on these two data sources.

[30] Using a different dataset of motels along the U.S. highway exits, Mazzeo (2002) also adopts the same categorization strategy.

the markets in my data set are not isolated, it is not possible to adopt the same market definition in this study.

Figure (3.15) presents quality configurations in 5-digit zip code markets in 2004. The horizontal axis represents the number of low-quality establishments, the vertical axis represents the number of high-quality establishments, while the larger bubble means the higher frequency of such quality configuration. The plot shows that many of the markets contain both type of hotels; while the proportion of low-quality to high-quality hotels varies from location to location. Many markets contain only low-quality hotels. Finally, none of the markets contain only high-quality hotels.

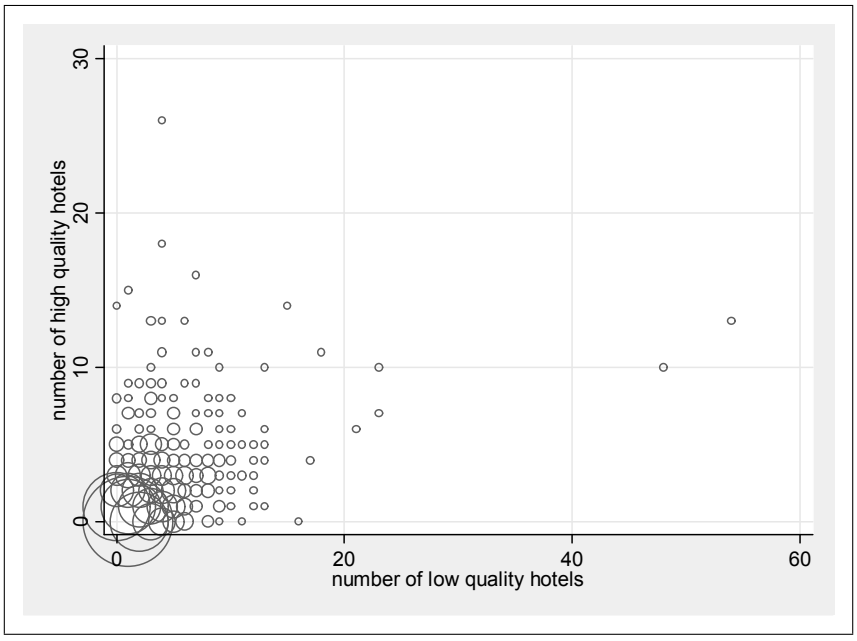


Fig. 3.15. Quality configurations in 5-digit zip code market, the entire Texas state, 2004.

Figure (3.16) portrays a geographical mapping of hotel establishments operating during 2004. The flags represent low-quality establishments, while the hollow circles represent high-quality establishments.

The geographical mapping (figure 3.16) suggests that low-quality hotels are more evenly spread across Texas, while high-quality hotels are more localised in metropolitan areas. However, no formal conclusion can be drawn from merely informal observation. Section 3.4 therefore uses the Multinomial Test of Agglomeration and Dispersion (MTAD) to formally analyse the degree of aggromeration/dispersion among hotels of the same quality. Section 3.5 then uses econometrics techniques to formally analyse the motives underlying the MTAD results. Before proceeding to these sections, the next section first discusses the data set used in the analyses.

3.3.3 The Data

Since the ensuing analysis on the MTAD (section 3.4) and hotels' quality choice (section 3.5) will require knowledge about each individual hotel's characteristics as well as the location demography, I divide the discussion on data into two parts. The first part is about the sources of individual hotel characteristics, while the second provides information on the sources of area-specific data.

Hotel Data

I assembled detailed information about each individual hotel's characteristics from four different data sources–the State of Texas Comptroller's Office, the AAA Texas Tour-

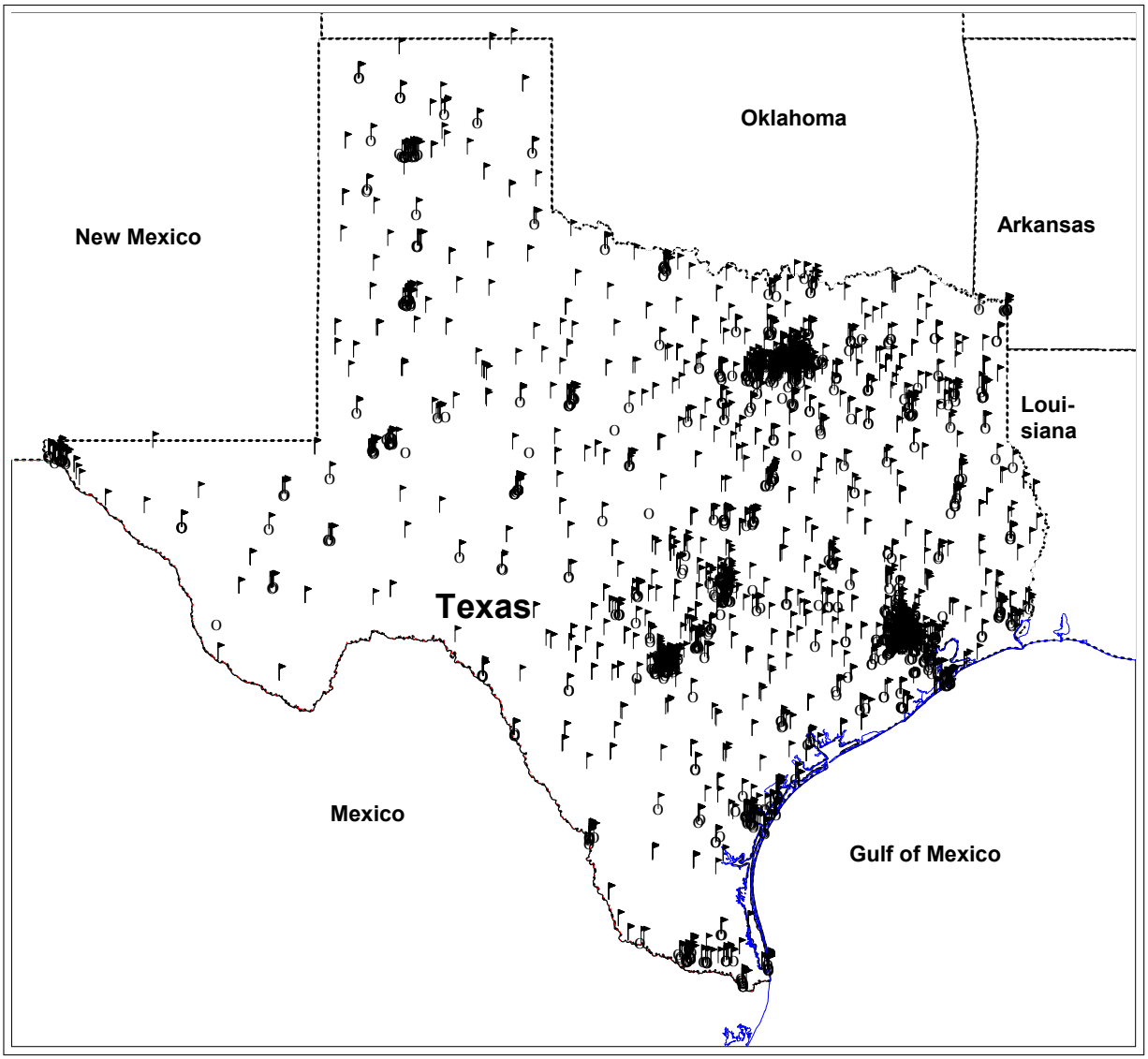


Fig. 3.16. Geographic location of all hotels in Texas, 2004. The flags represent low-quality hotels. The hollow circles represent high-quality hotels.

Book (2004 edition), Smith Travel Research, and Source Strategies Incorporated. The last two sources are private consulting companies specialising in the lodging industry.

The State of Texas Comptroller's Office discloses to the public the *Hotel Occupancy Tax Receipts* of all lodging properties in Texas. From their website and upon request, I gathered information on the quarterly revenue, the number of rooms, and the addresses of all properties operating during the period 1998-2004. Since some lodging properties are not usually perceived as competitors by others, I discard them from my sample. Such properties include the ones that have less than seven rooms or earn less than 10,000 US dollars per year or are university dormitories.

Since the precise number of operating firms in the market is crucial to the analysis, I verify the status of properties that were missing from or were added into the State of Texas Comptroller's Office's quarterly record using their actual opening and closing dates. This information, along with the hotel *chain scale* and some 9-digit zip codes, were obtained from Smith Travel Research. Smith Travel Research's *chain scale* measures the service scale of each hotel chain. There exist about 89 different hotel chains in Texas with only 21 major chains owning more than 20 properties in the entire state. Other information I acquired from Smith Travel Research was the property's 9-digit zip code. The 9-digit zip code, also termed ZIP+4®, has four more digits than the original 5-digit zip code (here after zip5) commonly used in the US postal system. The four additional digits are augmented to the original zip5 for more geographical precision.

Turning to the AAA Texas TourBook (2004 edition); from this, I obtained the information on each individual hotel rating. The AAA's rating ranges from one to five diamonds

with five diamonds being the top. Properties that do not meet the AAA's standards are not listed in the TourBook. As such, only 16.66% of the hotels with a lower rating than two diamonds are listed, while 91.37% of the hotels with a higher rating than three diamonds are listed[31]. Since it is almost certain that the missing high-quality chain hotels are not included in the TourBook because they were overlooked and not because they did not meet the AAA's standards, I assume that their diamond rating equals the average brand's diamond rating. For the sake of convention, I will use the term "star" instead of "diamond" from now on.

Note that, for the simplicity of estimation, the empirical models in section 3.5 use both Smith Travel Research's chain scale and the AAA's diamond rating to categorise hotels into two quality groups–low quality and high-quality. See table (3.25) for comparison between Smith Travel Research's chain scale and the AAA's diamond rating.

Lastly, the information on the annual average occupancy rate for properties operating during 2002 were obtained from the *Texas Hotel Performance Factbook* published by Source Strategy Incorporated. As will be seen in the ensuing analysis, the average occupancy rate will be used to measure the relative popularity of each quality-type of hotels across markets.

Area-specific Data

Apart from characteristics of the competitors, area-specific variations can also influence hotels' quality choice. For example, *ceteris paribus*, a new Holiday Inn® may choose

[31] These figures come from my own calculation which compares the number of properties acquired from the State of Texas Comptroller's (year 2004) and the number of properties listed in the AAA Texas TourBook (2004 edition).

to locate in Pharr rather than in Houston because the wage rate of labour working in Pharr is lower, and thus, their variable cost would be lower; *ceteris paribus,* a new Hilton® may choose to locate in downtown Dallas rather than elsewhere because the area has more entertainment establishments, and thus, is likely to attract more customers. As such, the failure to account for these area-specific variations could lead to misleading or biased results. I then try to control for these area-specific variations by obtaining the data from various sources.

To control for the market size, I obtained zip5 population, zip5 area, and zip5 urban area from a geographic correspondence engine, *MABLE/ Geocorr* version 2K, provided on the University of Missouri's website. The *MABLE/ Geocorr* 2K based its output on the US decennial Census 2000, the latest version of US Census available to the public. I calculated the zip5 population density by dividing zip5 population by zip5 geographical area.

To control for business activities, I obtained zip5 expenditure per square mile (psm) and zip5 employment psm from the "ETI Purchasing Power, Business Activity, and Workforce Density Profiles for All Residential ZIP Codes in U.S.", webpage provided by the University of Wisconsin-Milwaukee. The ETI (Employment and Training Institute) compiled the expenditure data from the Bureau of Labour Statistics Consumer Expenditure Surveys, 2002. Their employment information came from US Census Bureau ZIP Code Business Patterns, 2001. Other business activities such as the numbers of retail, professional, educational, art and recreational, administration and beauty salon establishments were obtained from the US Census Bureau website (guide to the 2002 economic census)[32].

[32] These data are available for 1997 and 2002. Thus, the 2002 dataset is more relevant to this study.

From the same website, I also obtained the zip5 real estate tax and zip5 average household size (2000 census of population and housing)[33]

Information about crime is only available at the city level. I obtained the crime rate index from the FBI Uniform Crime Reports (UCR) website. These data are available from 1996-2008. Finally, median income and population by race are only available at the county level. I obtained the number of population by race from the US Census Bureau website, while obtained median income from the Small Area Income & Poverty Estimates (SAIPE) website.

In order to take into account the physical proximity between hotels, I compiled the required information for a distance computation from various sources. First of all, I gathered 9-digit zip codes from Smith Travel Research's data base and the United States Postal Service's website[34]. I then, used the US. geographical data base,Topological Integrated Geographic Encoding and Referencing (TIGER)/ZIP+4 file, to obtain latitude and longitude coordinates of each 9-digit zip code. The TIGER/ZIP+4 file was developed by the United States Postal Service and the US Census Bureau. It is detailed enough to identify physical distance between two places located less than 0.5 of a mile apart. However, as shown in table (3.26), 4.88% of hotels in my data set were not assigned a 9-digit zip code. I identify their location by assuming that they are located in the middle of their 5-digit zip code area.

---

[33] Only information for the year 2000 is available under this survey topic. To approximate the tax values in all other years, I inflate and deflate the 2000 values using the CPI.

[34] On their website, the United States Postal Service provides a 9-digit zip code search for any given US mailing address through a search engine called "ZIP code Lookup". From here, I was able to obtain the 9-digit zip code (one-by-one) of most of the 4,415 hotels in my records.

Table 3.26. Number of hotels with and without 9-digit zip code, 2004

Star	9-digit zip assigned	9-digit zip not assigned	Percent not assigned (%)
1	1,560	125	8.01
2	614	13	2.12
3	1,037	20	1.93
4	40	1	2.50
5	5	0	0
Total	3,256	159	4.88

Notes: Author's tabulation based on Hotel Occupancy Tax Receipts data from Texas Comptroller's Office. "9-digit zip assigned" means 9-digit zip code assigned by US. Postal Service.

Finally, the geographic distance between each pair of hotels was calculated by using the Haversine formula[35]. This formula is also used to calculate geographical distances between firms in Seim (2006) and Orhun (2006). It takes into account the curvature of the earth and gives highly accurate distance approximation. See Sinnott (1984) for the complete discussion on the formula.

3.4 Test of Product Agglomeration and Dispersion (MTAD)

I use the Multinomial test of agglomeration and dispersion (MTAD) to systematically test whether the hotels in each market are more agglomerated or dispersed in quality and size

[35] The distance between two points on the surface of the earth is calculated using the Haversine formula. Suppose the latitude and longitude of location "A" and location "B" are given by$(lat_A, lon_A)$and (lat_B, lon_B) respectively. The distance between them $(d_{A,B})$ is calculated as follows:

$$d_{A,B} = 2R \arcsin\left[\min\left\{1, \left[\sin\left(\frac{lat_B - lat_A}{2}\right)\right]^2 + \cos(lat_A)\cos(lat_B)\left[\sin\left(\frac{lon_B - lon_A}{2}\right)\right]^2\right)^{\frac{1}{2}}\right\}\right]$$

where $R = 6367$ km = 3956 mile is the average radius of the earth.

than the independent and random benchmark. The MTAD is an econometric method first introduced in Rysman and Greenstein (2005). It constructs a test statistic which exploits the property of combinatoric statistic embedded in the multinomial distribution. To explain the intuition behind this test method, I adopt an example similar to that mentioned in Rysman and Greenstein (2005).

Suppose there exists an industry with four firms $(i = 1, 2, 3, 4)$ who are free to locate in either one of two cells $(c = 1, 2)$. As such, the possible number of firms per market ranges from zero to four. The maximum agglomeration happens when all four firms choose to locate in the same cell; in which case, the combinatoric statistic equals $\binom{4}{4} = \frac{4!}{(4-4)!-4!} = 1$. In contrast, the maximum dispersion occurs when two firms choose to locate in each of the two cells; in which case, the combinatoric statistic equals $\binom{4}{2} = \frac{4!}{(4-2)!2!} = 6$. In other words, the more the firms disperse, the higher the combinatoric statistic. The MTAD exploits this combinatoric property embedded in the multinomial's probability mass function (pmf)[36] to construct its test statistic.

The MTAD's test statistic is constructed using two sets of information–the *observed likelihood* $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ and the *expected likelihood* $E[l(f, \mathbf{p})]$. The observed likelihood $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ is the log-likelihood of the multinomial probability mass function (pmf) calculated using observed information. Since the multinomial pmf takes the form:

$$L(x_m, n_m, p) = \binom{n_m}{x_m^1, ..., x_m^C} p_1^{x_m^1} ... p_C^{x_m^C},$$

[36] The joint pmf of the vector $(\mathbf{X}_1, ..., \mathbf{X}_n)$ is given by the multinomial form $f(x_1, ..., x_n) = \frac{m!}{x_{1!} \cdot ... \cdot x_n!} p_1^{x_1} \cdot ... \cdot p_n^{x_n} = m! \Pi_{i=1}^n \frac{p_i^{x_i}}{x_i!}$. The combinatoric term $\frac{m!}{x_{1!} \cdot ... \cdot x_n!}$ is called the *multinomial coefficient* (Casella and Berger, 1990, p.173).

where $\mathbf{x}_m$ is a vector of observing choice $x_m^1, ..., x_m^C$ in market m; n_m is the number of hotels in market m; and $\mathbf{p}$ is the vector of probabilities $p_1, ..., p_C$. Treating the state of Texas as the entire population, I derive probabilities $p_1, ..., p_C$ from their proportions in Texas. The *observed likelihood* $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$, then, takes the form:

$$l(\mathbf{X}, \mathbf{n}, \mathbf{p}) = \frac{1}{M} \sum_{m=1}^{m} \ln \left(\binom{n_m}{x_m^1, ..., x_m^C} \right) + x_m^1 \ln(p_1) + ... + x_m^C \ln(p_C),$$

where $\mathbf{X}$ is an $M \times C$ matrix of choices of agents in all markets and $\mathbf{n}$ is an $M \times 1$ vector of the number of hotels in each market. The observed likelihood $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ is thus the outcome of agents' actual choices.

To construct the random and independent choice benchmark, I find the expected value of the log-likelihood function $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ and call it the *expected likelihood* $E[l(f, \mathbf{p})]$. Its functional form is as follows:

$$E[l(f,p)] = \sum_{n_m=\underline{n}}^{\overline{n}} \sum_{\mathbf{z} \in \Theta(n_m)} \begin{pmatrix} \ln \left(\binom{n_m}{z^1, ..., z^C} \right) \\ +z^1 \ln(p_1) + ... + z^C \ln(p_c) \end{pmatrix} L(x_m, n_m, p) f(n_m),$$

where $\Theta(n_m)$ is the set of all possible combinations of choices made by agents in market m; $f(n_m)$ is the empirical distribution of n_m(the number of hotels in market m). Finally, to test whether agents' choices are made randomly and independently of one another, Rysman and Greenstein (2005) propose a test statistic $t(\mathbf{X}, \mathbf{n}, \mathbf{p})$ which compares the observed likelihood $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ with the expected likelihood $E[l(f, \mathbf{p})]$ benchmark. The functional form of $t(\mathbf{X}, \mathbf{n}, \mathbf{p})$ is given by:

$$t(\mathbf{X}, \mathbf{n}, \mathbf{p}) = l(\mathbf{X}, \mathbf{n}, \mathbf{p}) - E[l(f, \mathbf{p})],$$

where Rysman and Greenstein (2005) provide a proof that the distribution of $t(\mathbf{X}, \mathbf{n}, \mathbf{p})$ is asymtotically normal with zero mean and finite variance – $\sigma^2 = \mathbf{V}[l(f, \mathbf{p})]$.

To interpret the test statistic result, I use the combinatoric statistic intuition as follows:

- If the observed likelihood is significantly smaller than the expected likelihood, or $t(\mathbf{X}, \mathbf{n}, \mathbf{p})$ is significantly less than zero, there is evidence of firms' agglomeration.
- If the observed likelihood is significantly larger than the expected likelihood, or $t(\mathbf{X}, \mathbf{n}, \mathbf{p})$ is significantly greater than zero, there is evidence of firms' dispersion.
- If the observed likelihood does not significantly differ from the expected likelihood, or $t(\mathbf{X}, \mathbf{n}, \mathbf{p})$ does not significantly differ from zero, there is evidence of independent random choice.

3.4.1 Application

In this section, I apply the MTAD method to the data set which includes all lodging properties operating in Texas during the year 2002. For simplicity, the word "hotels" hereby refers to all types of lodging properties in Texas including bed and breakfasts, motels, and hotels. I will analyse the MTAD twice, one with respect to hotel quality choice, another with respect to hotel size choice. In parallel with the aforementioned example, the number of firms is replaced by the number of hotels and the number of cells is replaced by the number of choices.

Regarding hotel quality, I discretise the quality into two groups. The 1-diamond and 2-diamond hotels are categorized as *low quality*, while the 3-diamond, 4-diamond, and 5-diamond hotels are categorized as *high-quality*. I find this categorization more justified than

Table 3.27. Characteristics of all hotels in Texas, 2002

Quality/Size	Small	Big	Total	Percent of total (%)
Low	2,785	436	3,221	73.62
High	562	592	1,154	26.38
Total	3,347	1,028	4,375	100
Percent of Total(%)	76.50	23.50	100	

Notes: Author's own tabulation. The information on size is acquired from the Texas Comptroller's Office. The information on quality is acquired from the Texas Tourbook, 2004 and Smith Travel Research

otherwise because it results in a minimum within-group heterogeneity. See table (3.25) in section 3.3.3 for more information on hotel ratings.

Regarding hotel size, I categorize hotels with no more than 100 rooms as *small* and hotels with more than 100 rooms as *big*. Again, this discretisation results in a minimum within-group heterogeneity in my case. Table 3.27 presents the breakdowns in quality and size of all hotels in Texas that were operating during the year 2002.

From now on, although they are not technically the same, I will use the terms *binomial* and *multinomial* interchangeably because hotels only have to make decisions between two choices–'low or high' and 'small or big'–due to the discretisation in my case. Treating all hotels in Texas as the entire industry, table 3.27 suggests that the (population) probabilities of the binomial distribution of quality are 0.7362 for low and 0.2638 for high[37]; while the probabilities of the binomial distribution of size are 0.7650 for small and 0.2350 for

[37] According to table 3.27, the industry's probability of observing a low-quality hotel equals (number of low-quality hotel) ÷ (total number of hotels) = 3221 ÷ 4375 = 0.7362. Since there are two alternatives, the probability of observing a high-quality hotel equals 1 - (probability of observing a low-quality hotel) = 1 - 0.7362 = 0.2638. The probabilities of the binomial distribution of size is calculated in the same manner.

Table 3.28. MTAD result for quality and size agglomerations

Choice	Market Definition	Observed Likelihood	Expected Likelihood	Std. Dev.	Markets
Quality	5-digit zip code	-1.4228	-1.1038	0.0175	954
	County	-1.5346	-1.4798	0.0351	284
Size	5-digit zip code	-1.4066	-1.0622	0.0181	954
	County	-1.9054	-1.4369	0.0360	284

big. I use these binomial probabilities to construct the $\mathbf{p}$ vector appeared in the *observed likelihood* $l(\mathbf{X}, \mathbf{n}, \mathbf{p})$ and the *expected likelihood* $E[l(f, \mathbf{p})]$ functions.

Two market definitions are adopted for the purpose of results comparison. The first one is '5-digit zip code' and the second one is 'county'. Since a 5-digit zip code area is geographically smaller than a county area, there are 954 5-digit zip code markets and 284 county markets. Comparing the MTAD results from different market definitions gives more insight into the scale of agglomeration or dispersion. In particular, one should expect the MTAD result obtained from a larger market definition to support a higher degree of independent random choice. For example, the observed likelihood would equal to the expected likelihood benchmark $-\ l(\mathbf{X}, \mathbf{n}, \mathbf{p}) = E[l(f, \mathbf{p})]-$ if the whole industry is treated as one market. According to this ad hoc market division, the MTAD would imply independent random decisions.

3.4.2 Results

Table (3.28) presents the observed likelihoods, the expected likelihoods, and the associated standard errors[38] for quality choice and size choice. Since the observed likelihoods are sig-

[38] Following Augereau et al. (2005), I draw the matrices $\mathbf{X}^s$ for 500 times and calculate $l(\mathbf{X}^s, \mathbf{n}, \mathbf{p})$ for each

nificantly smaller than the expected likelihoods, the results suggest that firms agglomerate in quality and in size. The agglomerative degree, although somewhat weakened, does not disappear when the markets' geographical area expands to the county level.

To illustrate the difference between agglomerative outcome and random outcome, I simulate random choices of hotels in each market and compare them with the true observations. Figures 3.17 and 3.18, show frequency histograms of true vs. simulated proportion of low-quality hotels; while figures 3.19 and 3.20 show frequency histograms of true vs. simulated proportion of small hotels. Histogram bars are divided into five bars: 0-19%, 20-39%, 40-59%, 60-79%, and 80-100%.

If the hotels make their quality choices randomly, the probability of drawing a low-quality hotel from *each market* should equal the probability of drawing a low-quality hotel from the *entire industry pool*. According to table(3.27), the probability of drawing a low-quality hotel from the entire industry pool is 0.7362, I use this probability to simulate the random choice benchmark. Since the probability of drawing a low-quality hotel (prob = 0.7362) more than doubles the probability of drawing a high-quality hotel (prob = 0.2638), the histograms are left skewed.

Having examined the nature of the histograms, their meanings can now be interpreted. Figures 3.17 and 3.18 strongly support low-quality hotel agglomeration. It can be seen that the actual number of markets with 80-100 percent low-quality hotels (the last pair of bars) is much higher than the simulated one. In other words, markets with 80-100 per-

---

of the new draws. Then, the standard deviation is calculated by Std. Dev. = $\sqrt{\frac{1}{499}\sum_{i=1}^{500}(l_i(\mathbf{X}^s,\mathbf{n},\mathbf{p}) - E[l(f,\mathbf{p})])^2}$.

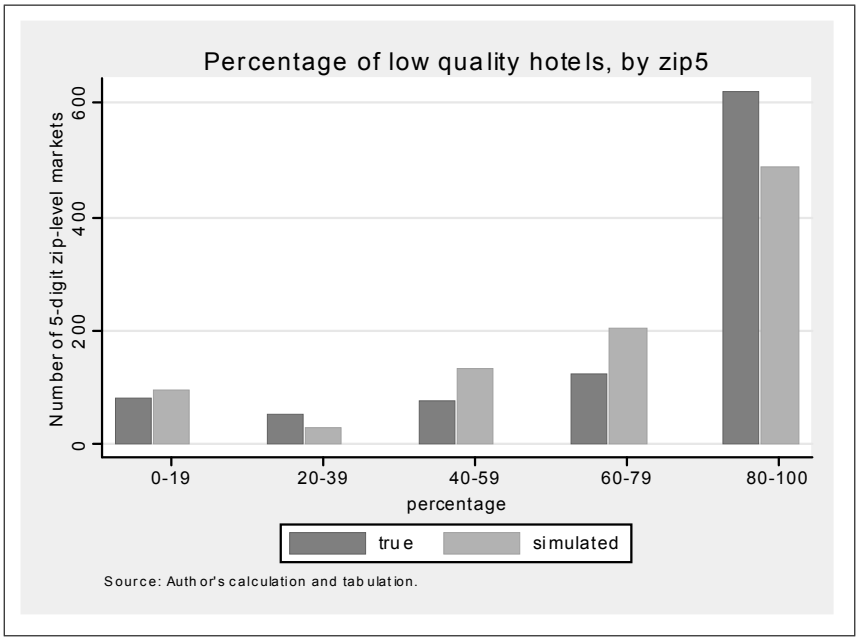


Fig. 3.17. Percentage of low-quality hotels in 5-digit zip markets.

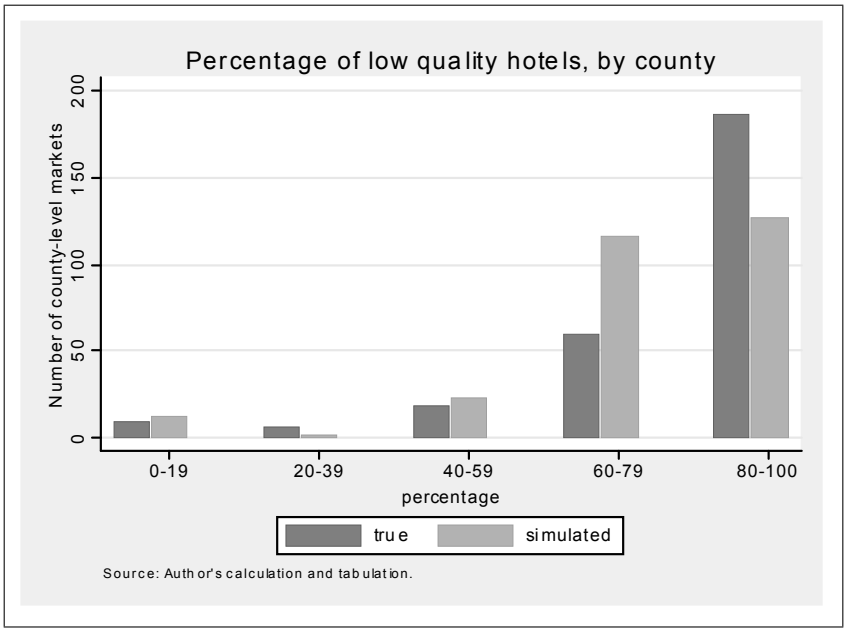


Fig. 3.18. Percentage of low-quality hotels in county markets.

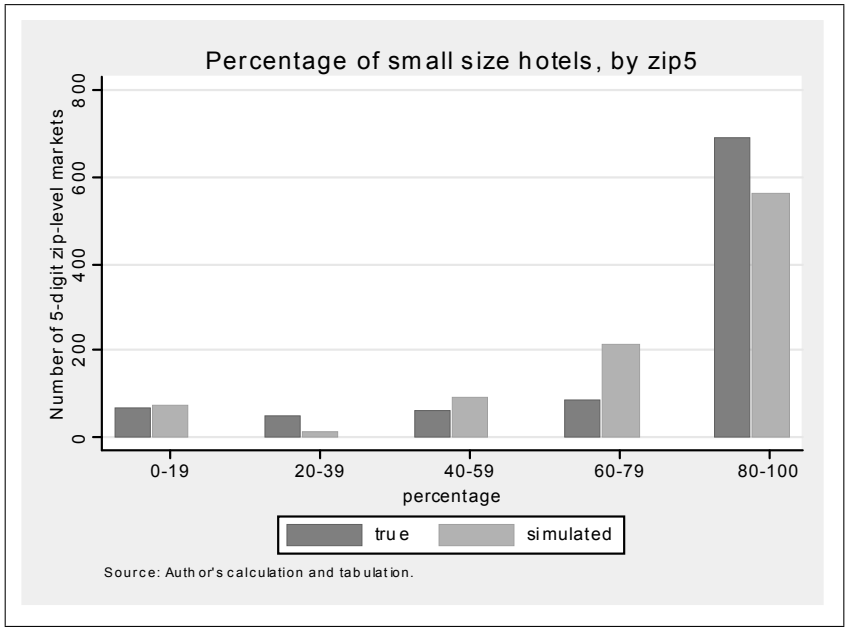


Fig. 3.19. Percentage of small hotels in 5-digit zip markets.

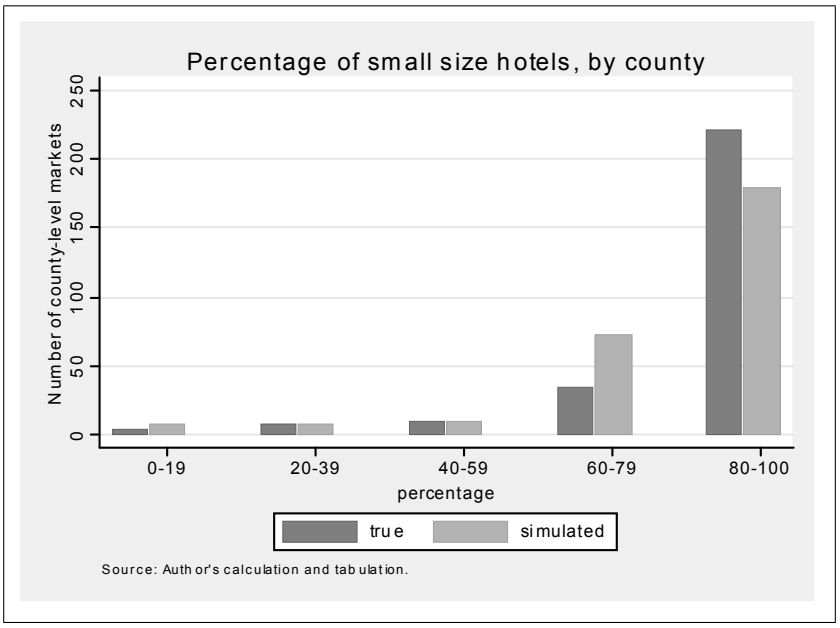


Fig. 3.20. Percentage of small hotels in county markets.

cent low-quality hotels are observed *more frequently* than it would be the case, had the hotels made their quality choices randomly and independently of one another.

Figures 3.17 and 3.18 suggest a contrasting implication for high-quality hotels. At both zip code and country levels, the actual number of markets with 0-19 percent low-quality hotels (the first pair of bars), which is the same as the number of markets with 81-100 percent high-quality hotels, is *lower* than the simulated one. In other words, markets with 81-100 percent high-quality hotels are observed slightly *less frequently* than it would be the case, had the hotels made their quality choices randomly and independently from one another. These results suggest a slight dispersion of high-quality hotels.

The histograms of size agglomeration across markets can be interpreted in the same fashion. Figures 3.19 and 3.20 strongly suggest an agglomeration of small hotels. It can be seen that the actual number of markets with 80-100 percent small hotels (the last pair of bars) is much higher than the simulated one. This result implies that markets with 80-100 percent of small hotels are observed more often than they would be the case, had the hotels made their decisions on size randomly and independently of one another. These results suggest an agglomeration of small hotels.

In contrast, figures 3.19 and 3.20 suggest that big hotels are slightly more dispersed than the random benchmark. Notice that the number of markets with 0-19 percent small hotels (the first pair of bars), which is equivalent to the number of markets with 81-100 percent big hotels, is lower than the simulated one. In other words, markets with 81-100 percent big hotels are observed less often than they would be the case, had the hotels made their decision on size randomly and independently of one another.

In summary, the MTAD test statistics in table 3.28 give strong evidences of hotel agglomeration in both quality and size; while the histograms further suggest that these results are mainly due to the high degree of agglomeration among low-quality and small hotels. High-quality and big hotels, on the other hand, are slightly dispersed.

Before moving on, note that this study will sidestep agglomeration in size and only focus on agglomeration on quality. This is not because the analysis on size is not interesting. One may, of course, further analyse firms' decisions on size in the context of limiting capacity, returns to scale, etc. Nonetheless, the agglomeration in quality is more relevant to my primary focuses, which is not only the incentive to agglomerate but also the incentive to differentiate in product type. To this effect, quality choice is a more appealing subject than size choice, i.e. while consumers may not view hotels of different sizes as less substitutable, they would surely view hotels of different qualities as less substitutable.

As such, the ensuing analysis will examine the determinants of hotels' quality choices in a more systematic manner. An econometric framework will be used to analyse whether the agglomeration suggested by the MTAD is driven by location-specific effects or strategic effects or both effects combined.

3.5 The Determinants of Quality Agglomeration

Having used the MTAD to confirm hotels' agglomeration in quality, I now proceed to investigate the underlying incentive of this phenomenon. Although various explanations can be given as to why firms of like-quality would want to locate near each other, I group those explanations into two main categories. The first is called the location-specific effects,

while the second is called the strategic effects. Geographical and demand advantages, when exist, usually play an important role in determining a firm's location and product-type choice (Ellison and Glaeser, 1997 and Mazzeo, 2002). For example, we see vineyards agglomerating in the Mediterranean because the climate is ideal for growing grapes; we see hotels agglomerate along beaches because those locations are appealing to tourists. As such, only geographical and/or demand advantages alone are already sufficient to explain some agglomeration phenomena.

In any case, some industries may agglomerate because firms enjoy a positive benefit from the presence of others. For example, antique shops usually cluster together because, as an agglomerated unit, they reduce shoppers' search difficulties and thus attract more shoppers; firms in Silicon Valley agglomerate to enjoy technological spillovers, transportation cost savings, and etc.

This section attempts to measure the strength of location-specific and strategic effects in determining hotels' quality choice. As previously mentioned, the difficulty of such a task is twofold. Firstly, the researcher may not observe all location-specific attributes that influence a firms' decision. If the observed variables are correlated with such unobserved attributes, their coefficients will be biased. Secondly, hotels' choices could be simultanously determined. That is, a given firm's quality choice depends upon other firms' quality choice and vice-versa. Therefore, a standard empirical estimation which treats the observed choices of other firms as exogenous would be invalid.

I adopt two approaches to remedy the above problems. The first approach selects only new entrant hotels and treats the quality choice made by incumbents as exogenous. It

also assumes that the incumbents are short-sighted. Therefore, they did not expect future entrants at the time they made their decisions. These assumptions rule out the simultaneity problem. I call this approach the *new hotels approach.* The second approach treats a snapshot of the market as an equilibrium of firms' quality choice. It then, uses the static Bayesian equilibrium concept to analyse the outcome. I name this second approach the *static Bayesian game approach.*

In what follows, I first set up the underlying quality-choice model. I then, use the new hotels approach and the static Bayesian game approach to analyse the data.

3.5.1 The Model

I assume that any prospective hotel first decides whether to enter a particular location. Then, conditional on its decision to enter, the market characteristics, and the decisions made by other hotels, a prospective hotel chooses the quality which maximizes its profit. The reduced form profit function of a prospective hotel j with quality choice $g \in \{lo, hi\}$ in market m takes the form:

$$\pi_j^m = \Pi_j^m(\mathbf{x}_j, f(n_g^m), g_j, \xi^m) + \varepsilon_j^m, \tag{3.29}$$

where Π_j^m is the deterministic part of the profit function, ε_j^m is the individual hotel's error component. Other variables are defined as follows:

$\mathbf{x}_j$	is a vector of observed exogenous variables determining firm j's profit. This includes things that would affect demand, consumers' willingness to pay, and firms' costs.
$f(n_g^m)$	a function of the number of g-quality rivals of firm j.
g_j	quality of firm j.
ξ^m	market-specific unobserved heterogeneity (by the researcher). This includes demand and cost shifters that are not observed by the researcher.

My assumption on demand will be refected through variables in $\mathbf{x}_j$, the vector of exogenous variables. These are variables that could indicate the level of demand for each type of hotel rooms. For example, the number of professional establishments and the number of educational establishments. In general, one could expect high-quality hotels to be located close to professional establishments because it would imply a high demand from corporate customers. On the other hand, one could expect high-quality hotels to avoid being close to educational establishments, because this would imply that the area is a residential area. Unobserved demand shifters will also be accounted for through ξ^m, the market-specific unobserved heterogeneity. In any case, the correlation between each demand variable and the hotels' quality choice will be discussed in the result section.

As hotels are profit maximizers, I assume that they would choose the quality that maximizes profit given their decision to enter, the market characteristics, and the decisions of other hotels. Suppose hotel j in market m chooses $g = high$, the choice probability can be given by:

$$\begin{aligned} P_j^m(g_j &= hi) = \Pr(\Pi_{j,hi}^m + \varepsilon_{j,hi}^m > \Pi_{j,lo}^m + \varepsilon_{j,lo}^m) \\ &= \int I(\Pi_{j,hi}^m + \varepsilon_{j,hi}^m > \Pi_{j,lo}^m + \varepsilon_{j,lo}^m)\, f(\varepsilon_j^m)\, d\varepsilon_j^m \end{aligned} \tag{3.30}$$

where $I(\cdot)$ is an indicator function equal to 1 when the statement in parentheses is true and 0 otherwise, $\Pi^m_{j,hi}$ is hotel's profit when choosing $q = hi$, $\Pi^m_{j,lo}$ is hotel's profit when choosing $q = lo$, $\varepsilon^m_{j,hi}$ is the error component when $q = hi$ is chosen, and $\varepsilon^m_{j,lo}$ is the error component when $q = lo$ is chosen. The probability of choosing quality choice $g = hi$ is then the integral of the statement over the distribution of the error component $f(\varepsilon^m_j)$.Without knowing the true distribution of ε^m_j, I arbitarily assume that it is normally distributed[39] with zero mean and a covariance matrix Ω. Let $\phi(\cdot)$ be the normal distribution function:

$$\phi(\varepsilon^m_j) = \frac{1}{2\pi|\Omega|^{1/2}} \exp(-\frac{1}{2}\varepsilon^{m\prime}_j \Omega^{-1} \varepsilon^m_j),$$

then, we can express(3.30) as:

$$\begin{aligned} P^m_j(g_j &= hi) = \int I(\Pi^m_{j,hi} + \varepsilon^m_{j,hi} > \Pi^m_{j,lo} + \varepsilon^m_{j,lo})\ \phi(\varepsilon^m_j)\ d\varepsilon^m_j \\ &= \int I((\Pi^m_{j,lo} - \Pi^m_{j,hi}) + (\varepsilon^m_{j,lo} - \varepsilon^m_{j,hi}) < 0)\ \phi(\varepsilon^m_{j,lo} - \varepsilon^m_{j,hi})\ d(\varepsilon^m_{j,lo} - \varepsilon^m_{j,hi}) \\ &= \int I(\widetilde{\Pi}^m_j + \widetilde{\varepsilon}^m_j < 0)\ \phi(\widetilde{\varepsilon}^m_j)\ d\widetilde{\varepsilon}^m_j \qquad (3.31) \\ &= \int_{\widetilde{\varepsilon}^m_j \in B_j} \phi(\widetilde{\varepsilon}^m_j)\ d\widetilde{\varepsilon}^m_j \qquad (3.32) \end{aligned}$$

Adopting Train (2003)'s expression, let B_j represent the set of error terms $\widetilde{\varepsilon}^m_j$ when the hotel chooses $g_j = hi$, i.e. $(B_j = \widetilde{\varepsilon}^m_j\ s.t. \widetilde{\Pi}^m_j + \widetilde{\varepsilon}^m_j < 0)$. For identification, profit level from choosing low-quality $(g = lo)$ is normalized to zero, $\Pi^m_{j,lo} = 0$. The probability expression(3.32) then becomes:

$$\begin{aligned} P^m_j(g_j &= hi) = \Phi(\Pi^m_{j,hi}) = \Phi(\Pi^m_j) \qquad (3.33) \\ P^m_j(g_j &= lo) = 1 - \Phi(\Pi^m_j) \end{aligned}$$

[39] Another popular distribution for ε^m_j is the iid extreme value, which would give the logit model.

where Φ is the cumulative distribution function of the standard normal distribution.

To construct the empirical model, I assume linear relations between profit and its determinants. At this point, it is crucial to discuss how the market boundary m is defined. Ideally, I would like it to be the boundary which contains all of the firms that compete for the same customers. Bresnahan and Reiss(1991) and Mazzeo(2002) select geographically isolated markets where the areas are likely to contain all (and only) the firms and consumers that interact with each other. When markets in the data set are not geographically isolated, such as in my case, an arbitrary market definition is usually adopted. Often, the researcher uses a market definition upon which the area-specific data are available–e.g. zip code, census block, census track, city, county, state, etc. One problem associated with this strategy is that arbitary geographical markets may capture too many or too few firms and consumers; or they may draw boundaries right in the middle of markets. Therefore, in addition to using an arbitary geographical definition, physical proximity to competitors should be accounted for. For instance, Seim(2006) uses city as the market definition, but classifies effects from competitors by their distance. I adopt a similar strategy by defining a market as a 5-digit zip code area (zip5). I then, account for the competitors' effects (strategic effects) by their proximities to the prospective hotel. As such, the reduced form of a hotel's profit function becomes:

$$\Pi_j^m = \underbrace{\xi^m + \sum_k x_k \beta_k}_{\text{(location-specific effects)}} + \underbrace{\sum_g \sum_d n_{g,d} \theta_{g,d}}_{\text{(strategic effects)}}, \tag{3.34}$$

Variable ξ^m is the unobserved (by the researcher) zip5-specific preference for quality. The vector x_k includes zip5-level population density, expenditure per square mile (psm), employment psm, retail establishments psm, professional establishments psm, educational

establishments psm, art and entertainment establishments psm, restaurant establishments psm, beauty salon establishments psm, dummy variables for big cities, and dummy variables for each semi-annual period. Variable $n_{g,d}$ is the number of other hotels with quality g who locate in the d-mile concentric circle around hotel j. In this study, *the parameters associated with $n_{g,d}$ are particularly of interest because they quantify the strategic effects created by proximities to other hotels.* See Table (3.29), (3.30), (3.32), (3.33) for the summary statistics of each variable.

Before moving on, it is worth nothing that I assume all hotels located in the same distance band to have the same impact on the hotel located at the centre of the circle, regardless of their actual distance from the centre. For example, I assume all hotels in $d = 0 - 2$ (between 0 to 2 mile) to have the same impact on hotel j that is located at the centre of the circle; and $d = 2 - 5$ (between 2 to 5 miles) to have the same impact on hotel j that is located at the centre of the circle. Figure 3.21 illustrates the 0-2 and 2-5 mile distance bands in a sample market. A flag represents a low-quality hotel; while a hollow circle represents a high-quality hotel.

3.5.2 Problems with Empirical Estimation

Simultaneity Bias

Suppose we use quality choices of all the 4,531 existing hotels to estimate the probit model (3.33), (3.34). This will be equivalent to assuming that all choices are made simultaneously in a static setting. Such a setting would be subject to simultaneity bias because, while each hotel's quality choice g_j depends on the number of g-quality hotels in the area

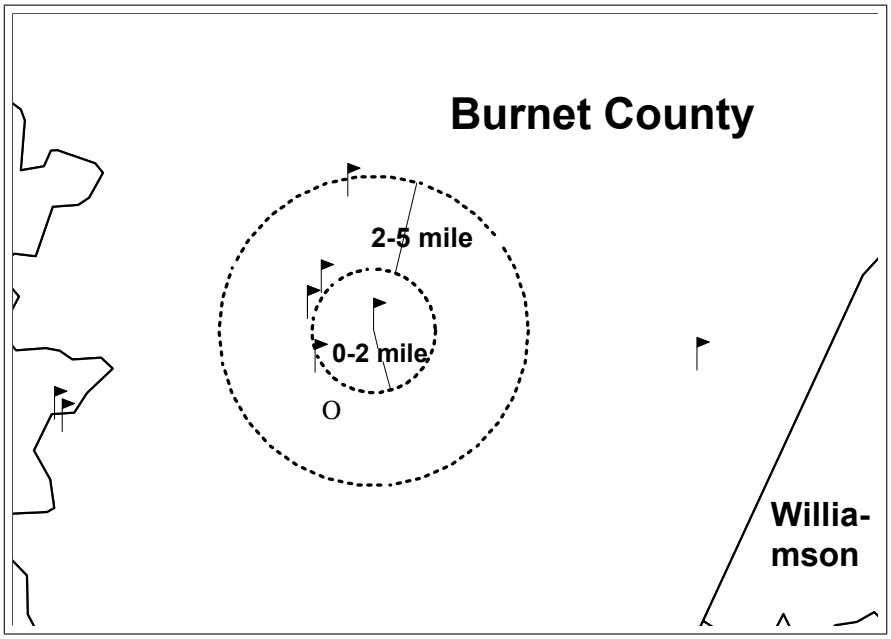


Fig. 3.21. Illustration of distance bands in a sample market – Burnet county. The Flags represent low-quality hotels, hollow circles present high-quality hotels.

$n_{g,d}$, the number of g-quality hotels in the area $n_{g,d}$ also depends on hotels' quality choices g_j. Thus, the coefficient on $n_{g,d}$ measures not only the *direct* effects from $n_{g,d}$ itself, but also the *indirect* effects from the dependent variable g_j. In econometrics, we can say that the explanatory variable $n_{g,d}$ would be correlated with the error term; and the estimate of its coefficient would be biased.

Under the circumstance where this simultaneity bias is significant, using all existing hotels' choices to estimate the probit model (3.33), (3.34) would overstate the positive spillovers. Here, since $n_{low,d}$ is likely to be negatively correlated with the error term ε_j, the estimate of $\theta_{low,d}$ tends to be downward-biased. On the other hand, since $n_{high,d}$ is

likely to be positively correlated with the error term ε_j, the estimate of $\theta_{high,d}$ tends to be upward-biased[40].

In order to remedy this simultaneity bias problem, I consider two approaches. The first approach looks only at the new entrant hotels and treats the incumbents' quality choices as *exogenous*.The second strategy follows the static Bayesian game literature (Seim, 2006) and uses the two-step estimation approach to derive consistent estimates (Bajari et al., 2006). In short, the two-step approach suggests that one first derives the "expected $n_{g,d} = \hat{n}_{g,d}$" through a policy function. Then, given that $\hat{n}_{g,d}$ is no longer subject to the simultaneity bias, one can employ standard econometrics techniques to derive consistent estimates in the second step (Bajari et al., 2006)(Ellickson and Misra, 2008). Both strategies will be discussed in detail in the next section.

Unobserved Heterogeneity (ξ^m)

The term "unobserved heterogeneity" refers to attributes that are not observed by the econometrician, but are observed by the hotels and influence their decisions. For example, building regulations, parking amenities, and other factors that could increase the profit of one type of hotel than the other. The researcher, on the other hand, cannot observe such information. If the unobservables are not accounted for, but are correlated with variables in the regression, the estimated coefficients would be biased.

[40] For example, the higher number of low-quality hotels $n_{low,d}$ tends to induce other firms in the area to choose $g = 0$ (low quality), thus the low value of ε_j results. So, $n_{low,d}$ is negatively correlated with ε_j. On the other hand, the higher number of high-quality hotels $n_{high,d}$ tends to induce other firms in the area to choose $g = 1$ (high quality), thus the high value of ε_j results. Therefore, $n_{high,d}$ is positively correlated with ε_j.

In my case, the unobserved heterogeneity ξ^mcould be regarded as the market-specific preference for quality. I account for it through two means. First, I include a proxy variable for ξ^m; and second, I incorporate a market-specific random intercept.

To discuss properties of a valid proxy variable, consider again the profit function:

$$\Pi_j^m = \xi^m + \sum_k x_k \beta_k + \sum_g \sum_d n_{g,d} \theta_{g,d} \tag{3.35}$$

where all the variables are as defined in eq. 3.34. Now, let z_1 be a proxy variable. The relation between the market-specific unobservables and z can be written as:

$$\xi^m = \varphi_0 + \varphi_1 z + r$$

where r is the error component. For z to be a valid proxy for ξ^m, r has to be uncorrelated with z and with all the regressors in eq. 3.35 (Wooldridge, 2002, p.105). In practice, the restriction of no correlation between r and all the regressors in eq. 3.35 can be reduced to no correlation between r and the regressors of interest in eq.3.35. In this study, the variables of interests are $n_{g,d}$, the number of g quality competitors located within the d-mile distance band. They can generally be regarded as the supply of hotels. A valid proxy variable (z), therefore, should not have shocks (r) that are correlated with the supply of hotels in the area.

To this effect, I construct two proxy variables for ξ^m using the demand-side information as follows:

$$z_1 = \frac{high_room_sold^c}{all_room_sold^c} \tag{3.36}$$

$$z_2 = \frac{high_room_sold^{zip} - low_room_sold^{zip}}{all_room_sold^{zip}} \tag{3.37}$$

where superscript c denotes city, superscript zip denotes zip5 area, $high_room_sold$ is the total number of high-quality rooms sold in 2002, low_room_sold is the total number of low-quality rooms sold in 2002, and all_room_sold is the total number of rooms sold in 2002. I assume that all rooms in a high-quality hotel are high-quality rooms, while all rooms in a low-quality hotel are low-quality rooms.

Both proxy variables should be highly correlated with the market-specific preference for high quality. Moreover, the shocks to both z_1 and z_2, which are demand-side shocks should not be correlated with the supply of hotels in the market. Therefore, I believe that both proxy variables are valid. But since they are very similar, using only one of the two may be sufficient. I will explore this issue in the estimation.

In addition to proxy variables, I use a random intercept to allow for market-specific common unobservables. Within the same market, the intercepts take the same value; across markets, they are independently and normally distributed. To incorporate both proxy variables and a random intercept in the model, one can rewrite the profit function as:

$$\Pi_j^m = \eta^m + \sum_l z_l \tau + \sum_k x_k \beta_k + \sum_g \sum_d n_{g,d} \theta_{g,d} \tag{3.38}$$

where η^m is the independently and normally distributed random intercept for market m, and z_l represents proxy variables defined in eq.3.36 and 3.37. Since η^m has a distribution, the probability of choosing high quality (defined in eq.3.33) then becomes:

$$\begin{aligned} P_j^m(g_j &= hi) = \int \Phi(\Pi_j^m(\mathbf{z}, \mathbf{x}, \mathbf{n}; \boldsymbol{\theta}, \eta^m))\ \phi(\eta^m)\ d\eta^m \\ P_j^m(g_j &= lo) = 1 - \int \Phi(\Pi_j^m(\mathbf{z}, \mathbf{x}, \mathbf{n}; \boldsymbol{\theta}, \eta^m))\ \phi(\eta^m)\ d\eta^m \end{aligned} \tag{3.39}$$

where $P_j^m(g_j = hi)$ is the probability of hotel j in market m choosing high quality, Π_j^m is the profit of hotel j in market m (defined in eq.3.38), $\mathbf{z}$ is a vector of proxy variables, $\mathbf{x}$ is a vector of all the area-specific variables x_k^m, $\mathbf{n}$ is a vector of all the strategic variables $n_{g,d}$, $\boldsymbol{\theta}$ is a vector of all the parameters in the model, $\Phi(\bullet)$ is the cumulative distribution function of the standard normal distribution, and $\phi(\bullet)$ is the density function of the standard normal distribution. With a random intercept, $P_j^m(g_j = hi)$ is a weighted average of the probit probabilities evaluated at different values of η^m.

3.5.3 Estimation Strategy 1: New Hotels Approach

The new hotels approach alleviates the simultaneity bias by only looking at the decisions made by new entrant hotels. I treat existing market structures (the incumbent's entry and quality choice) at the time of entry as exogenous. By this, I also have to assume that all the incumbents were short-sighted because their decisions would not be influenced by the expectation of the new entrants' actions. I also assume that it takes an entrant six months to build a new hotel. This time lag should be sufficient for an investor to plan and execute an entry and quality choice decisions. However, it should not be long enough to cause significant change in fundamental cost and demand factors.

The data is, therefore, divided into 14 semi-annual periods from January 1998- December 2004. For each period, I keep only the newly opened hotel. Then, I collect the explanatory variables $\mathbf{x}$ and $\mathbf{n}$ from the previous semi-annual period. However, most area-specific variables are only available for the end of year 2002 due to the US Census Bureau's survey timing. This lack of time-varying data would be a problem if they did not change

proportionately (relative to one another) over time. However, if the changes were proportionate (so the relative values of variables do not change much), they can be used to control for variations across locations. Variable definitions and their summary statistics can be found in tables 3.29 and 3.30. The estimated model is similar to the one previously discussed in eq.3.38. However, time superscripts are added to clarify that a hotels' profit is determined by explanatory variables observed in the previous period.

$$\Pi^{m,t}_{h(new_hotel)} = \Phi(\ \eta^{m,t-1} + \sum_{l} z_l^{m,t-1}\tau + \sum_{k} x_k^{t-1}\beta_k + \sum_{g}\sum_{d} n_{g,,d}^{t-1}\theta_{g,d}),\ \forall(h \in H) \quad (3.40)$$

where t indicates different semi-annual time periods, $h = 1, ..., H$ indicates new hotels. All other variables are as defined in eq.(3.34). Under the assumption of short-sighted incumbents, eq.(3.40) no longer suffers the simultaneity bias. As discussed in the previous section, the market-specific unobserved heterogeneity $\xi^{m,t-1}$ is accounted for by a proxy variable $z_l^{m,t-1}$ and a random intercept $\eta^{m,t-1}$.

The estimation of eq.3.40 is implemented by a maximum likelihood technique, where the estimated parameters are those that maximize the likelihood or log-likelihood function. In this case, the probit likelihood function with a random intercept can be defined as

$$\begin{aligned} L &= \prod_{h=1}^{H} P_h^m(g_h = hi)^{y_{h,hi}} \times P_h^m(g_h = lo)^{y_{h,lo}} \\ \log L &= \sum_{h=1}^{H} y_{h,hi} \log P_h^m(g_h = hi) + \sum_{h=1}^{H} y_{h,lo} \log P_h^m(g_h = lo) \end{aligned}$$

where L is the likelihood function, $P_h^m(g_h = hi)$ is the probability that the new hotel h in market m will choose high quality, $P_h^m(g_h = lo)$ is the probability that the new hotel h in market m will choose low quality, $y_{h,hi} = 1$ if high quality is chosen and $= 0$ if high quality is not chosen, $y_{h,lo} = 1$ if low quality is chosen and $= 0$ otherwise. Substitute

for $P_h^m(g_j = hi)$ with an expression which incorporates both the proxy variable and the random intercept (eq.3.39), and use $\Pi^{m,t}_{h(new_hotel)}$ as defined in eq.3.40, we get:

$$\begin{aligned} \log L \;=\; & \sum_{h=1}^{H} y_{h,hi} \log \int \Phi(\Pi^m_{h(new_hotel)}(\mathbf{z}, \mathbf{x}, \mathbf{n}; \boldsymbol{\theta}, \eta^m))\; \phi(\eta^m)\; d\eta^m + \\ & \sum_{h=1}^{H} y_{h,lo} \log \int (1 - \; \Phi(\Pi^m_{h(new_hotel)}(\mathbf{z}, \mathbf{x}, \mathbf{n}; \boldsymbol{\theta}, \eta^m)))\; \phi(\eta^m)\; d\eta^m \end{aligned} \tag{3.41}$$

Since there is no closed form solution to this maximisation problem, the solution has to be obtained through an iteration process. In particular, I use a Newton-Raphson algorithm to search for the set of parameters which maximize eq.3.41. This maximisation routine can be implemented through the gllamm command available in STATA.

Results (New Hotels Approach)

Tables 3.29 and 3.30 report the definition and summary statistics of each variable used for estimation. Strategic effects are accounted for by $n_{lo,0-2}, n_{lo,2-5}, n_{hi,0-2}$ and $n_{hi,2-5}$. The variable $n_{lo,0-2}$, for example, denotes the number of low-quality competitors located within the zero to two miles distance band. Location-specific effects are accounted for by several location-specific variables. At the zip5 level, I include expenditure per square mile (psm), employment psm, population density psm, retail establishments psm, professional establishments psm, educational establishments psm, art and entertaimment establishments psm, food and restaurant establishments psm, salon establishments psm, administration establishments psm, and total population. At the city level, I include median house tax, average house size, the crime index, and a dummy variable for each city with more than 20 new hotels establishments–i.e. Austin, Dallas, Houston, Fort Worth, San Antonio and

Irving. At the county level, I include proportion of black residents, proportion of urban area, and median income.

Table 3.31 sets out the probit estimation results. Specification 1 does not include any proxy variable, specification 2 includes only the first proxy variable ($z_1 = \frac{high_room_sold^c}{all_room_sold^c}$ where c denotes city), specification 3 includes only the second proxy variable ($z_2 = \frac{high_room_sold^{zip} - high_room_sold^{zip}}{all_room_sold^{zip}}$ where zip denotes zip5 area), specification 4 includes both proxy variables. In all specifications, I allow the zip5-specific intercept to be randomly distributed with a standard normal distribution. This is to account for the zip5-specific unobserved heterogeneity. The null hypotheses are rejected at the 5 percent level of significance, unless stated otherwise. I drop hotels with less than 20 rooms and cities with less than 10 hotels. This is to avoid cases with no strategic interactions. This reduces the sample size from 791 hotels to 562 hotels. In the robustness check section 3.5.5, I will show that including all hotels would not give different qualitative results. However, some coefficients may become insignificant.

Results from all specifications generally support the existence of positive spillovers from the *same* quality-type of hotels. In other words, high-quality (low-quality) hotels benefit more from locating near other high-quality (low-quality) hotels than low-quality (high-quality) hotels. Including proxy variables to account for ξ^m (the area-specific unobserved attractiveness for high-quality hotels) in specifications 2, 3 and 4 reduces the strength of positive spillovers, but does not change the qualitative results. That is, the inclusion of proxy variables neither change the sign of coefficients on strategic variables, nor cause them to become insignificant.

The negative and significant coefficient on $n_{lo,0-2}$ implies that when there are more low-quality incumbents in the area, the probability of choosing high-quality, or the expected profit from choosing high-quality, decreases. In other words, the profit from choosing low-quality increases. After controlling for the unobserved heterogeneity through proxy variables in specifications 2, 3 and 4, the value of the coefficient on $n_{lo,0-2}$ becomes less negative from -0.021 to -0.017, -0.015 and -0.016 respectively. This suggests that, even after controlling for the zip5-specific unobservables, low-quality hotels still create positive spillovers to competitors of the same quality-type. However, evidently from the negative but insignificant coefficient on $n_{lo,2-5}$, the positive spillovers fade away with distance.

Similarly, the positive and significant coefficient on $n_{hi,0-2}$ suggests that the presence of high-quality hotels in the area increases the expected profit from choosing high-quality. After controlling for the unobserved heterogeneity through proxy variables in specifications 2, 3 and 4, the value of the coefficient on $n_{hi,0-2}$ becomes less positive from 0.037 to 0.033, 0.026, and 0.027 respectively. From this, we can infer that the proxy variables do account for some unobserved location-specific effects (unobserved heterogeneity). However, the positive spillovers among high-quality hotels are still there. In any case, since the coefficient on $n_{hi,2-5}$ becomes insignificant, it implies that these positive spillovers fade away with distance. My estimation results suggest that the magnitude of positive spillovers from high-quality hotels is greater than low-quality hotels.

Turning to the location-specific effects, all specifications still give similar results. However, since specification 4 is the most consistent, I will discuss the results based

on coefficient estimates in specification 4. I find that high-quality hotels benefit more in areas where the household size is smaller ($\beta_{house\,size} = -0.318$), where there are more retail establishments psm ($\beta_{retail} = 0.024$), more professional establishments psm ($\beta_{professional} = 0.017$), less educational establishments psm ($\beta_{educational} = -0.287$), and less art and recreational establishment psm ($\beta_{art} = -0.257$). By assumption of the probit model, the effects are the opposite for the low-quality hotels.

All other coefficients on location-specific variables are insignificant. However, since this may have been caused by my small sample size, I decide to include them in my specifications for the purpose of consistency.

3.5.4 Estimation Strategy 2: Static Bayesian Game Approach

In the previous section, I remedy the simultaneity bias by only analyzing decisions of the new entrants hotels. In this section, I propose an alternative by using the static Bayesian game approach. In many cases where the sequence of agents' decisions is not observed, this approach renders feasibility of the analysis. It relaxes the assumption made by the new hotels approach that agents are short-sighted. On the other hand, it assumes that firms make simultaneous decisions based on their expectations of what other firms would do in equilibrium. This approach is adopted by (Seim, 2006), (Orhun, 2006), (Ellickson and Misra, 2008), (Jia, 2008), for example.

The static Bayesian game approach treats every firm's action as endogenous. It takes a snapshot of the market as an equilibrium, which results from agents' simultaneous actions. The action made by one firm can influence other firms' actions and vice-versa.

Table 3.29. Variable Definitions and Sample Characteristics

Variable Name	Definition	Mean and (Standard Deviation)
z_1	The first proxy variable defined in eq.3.37, $z_1 = \frac{high_room_sold^c}{all_room_sold^c}$	0.588 (0.139)
z_2	The second proxy variable defined in eq.3.38, $z_2 = \frac{high_room_sold^{zip} - low_room_sold^{zip}}{all_room_sold^{zip}}$	0.067 (0.348)
$n_{lo,0-2}$	predicted number of low-quality incumbents located within 0-2 mile distance band	10.287 (11.602)
$n_{lo,2-5}$	predicted number of low-quality incumbents located within 2-5 mile distance band	17.171 (16.357)
$n_{hi,0-2}$	predicted number of high-quality incumbents located within 0-2 mile distance band	10.573 (11.540)
$n_{hi,2-5}$	predicted number of high-quality incumbents located within 2-5 mile distance band	15.061 (17.192)
Expenditure (million $US.)	total expenditure (of all activities) per sq.mile in the zip5 area	9.848 (7.943)
Employment (1,000 people)	number of people employed (in all sectors) per sq.mile in the zip5 area	0.013 (0.034)
Pop density (1,000 people)	number of residents per sq.mile in the zip5 area	2.515 (1.913)
Population (million people)	total number if residents in the zip5 area	0.282 (0.158)
House tax (10,000 $US/year)	median annual real-estate tax in the zip5 area	0.210 (0.131)
Household size (thousand people)	average number of household members (adult & children) in the zip5 area	2.559 (0.526)
Crime Index	The sum of seven types of offenses known to law enforcement, namely, murder, forcible rape, robbery, aggravated assault, burglary, larceny-theft and motor vehicle theft.	10.018 (16.382)
Retail (no./sq.mile)	number of retail establishments per sq.mile in the zip5 area	12.854 (18.299)
Professional (no./sq.mile)	number of professional establishments per sq. mile in the zip5 area	27.855 (114.676)

Table 3.30. Variable Definitions and Sample Characteristics (continued)

Variable Name	Definition	Mean and (Standard Deviation)
Educational (no./sq.mile)	number of schools, educational establishment per sq. mile in the zip5 area	0.803 (1.487)
Entertainment (no./sq.mile)	number of entertainment, art establisment per sq. mile in the zip5 area	1.279 (2.805)
Restaurant (no./sq.mile)	number of restaurant, food establishment per sq. mile in the zip5 area	8.961 (21.985)
Salon (no./sq.mile)	number of beauty salon establishment per sq. mile in the zip5 area	1.075 (1.658)
Admin (no./sq.mile)	number of administrative establishment per sq. mile in the zip5 area	7.532 (17.508)
Urban	proportion of urban area per total county area	0.731 (0.345)
Black	proportion of black population in the county	0.138 (0.074)
Income (10,000 $US/year)	median annual income of residents in the county	4.242 (0.910)
City Dummies (for 6 popular cities)	0-1 dummy variables for Austin, Houston, San Antonio, Fort Worth, Dallas, Irving	–
Year dummies (for 7 years)	0-1 dummy variables for each half year period (e.g.199801-200402)	–

Table 3.31. Probit Estimation Result: New Hotel Approach

	(1)	(2)	(3)	(4)
constant	0.392 (0.758)	0.102 (0.773)	0.936 (0.780)	0.993 (0.803)
Proxy variables				
z_1	–	1.278**(0.619)	–	-0.200 (0.677)
z_2	–	–	1.361**(0.203)	1.381**(0.215)
Strategic effects				
$n_{lo,0-2}$	-0.021**(0.006)	-0.017**(0.006)	-0.015**(0.006)	-0.016**(0.006)
$n_{lo,2-5}$	0.000 (0.005)	0.003 (0.006)	0.001 (0.006)	0.001 (0.006)
$n_{hi,0-2}$	0.037**(0.008)	0.033**(0.008)	0.026**(0.008)	0.027**(0.008)
$n_{hi,2-5}$	0.004 (0.006)	0.001 (0.006)	0.002 (0.006)	0.003 (0.006)
Zip5 characteristics				
Expenditure	0.028 (0.037)	0.032 (0.037)	0.036 (0.038)	0.035 (0.038)
Employment	-22.815(33.547)	-19.982(33.319)	-14.878(33.311)	-15.316(33.360)
Pop density	-0.158 (0.143)	-0.164 (0.143)	-0.136 (0.148)	-0.135 (0.148)
Population	1.733 (1.235)	1.370 (1.248)	1.397 (1.265)	1.442 (1.274)
House tax	0.720 (0.724)	0.636 (0.717)	-0.001 (0.745)	0.001 (0.745)
House size	-0.304* (0.178)	-0.295* (0.178)	-0.316* (0.183)	-0.318* (0.183)
Retail	0.025* (0.013)	0.024* (0.013)	0.024* (0.014)	0.024* (0.014)
Professional	0.017**(0.008)	0.016* (0.008)	0.017** (0.008)	0.017**(0.008)
Education	-0.316* (0.166)	-0.305* (0.165)	-0.286* (0.149)	-0.287* (0.149)
Art	-0.276* (0.146)	-0.266* (0.143)	-0.255* (0.141)	-0.257* (0.141)
Restaurant	0.004 (0.025)	0.004 (0.025)	-0.007 (0.023)	-0.007 (0.023)
Salon	-0.181 (0.122)	-0.176 (0.121)	-0.141 (0.119)	-0.142 (0.119)
Admin	0.031 (0.034)	0.028 (0.034)	0.007 (0.033)	0.008 (0.033)
City characteristics				
Crime index	-0.137 (0.136)	-0.171 (0.136)	-0.104 (0.140)	-0.098 (0.141)
County characteristics				
Black	-1.246 (0.996)	-1.606 (1.015)	-1.635 (1.035)	-1.580 (1.050)
Urban	0.109 (0.297)	0.050 (0.298)	-0.071 (0.314)	-0.063 (0.315)
Income	0.028 (0.078)	-0.014 (0.080)	-0.018 (0.082)	-0.013 (0.084)
City dummies	yes	yes	yes	yes
No. of obs.	562	562	562	562
log-likelihood	-335.888	-333.709	-311.539	-311.496

Notes: ** significant at 5%, * significant at 10%. All specifications allow for a random intercept.

Because of this, we have a system of simultaneously determined equations to estimate. Existing literature has suggested two main estimation strategies. The first strategy is called the "nested fixed point" (Seim, 2006) (Orhun, 2006). This strategy searches for a set of parameters which solves the system of simultaneously determined equations. If multiple solutions are found, the researcher needs to specify an equilibrium selection constraint. This approach is computationally intensive when applied to a discrete-choice model because, in most cases, the closed form solution does not exist; hence, one has to find the solution through an iterative process.

The second strategy is called the "two-step approach" (Bajari et al., 2006) (Ellickson, 2008). I adopt this strategy for its ease of implementation. The two-step approach remedies the computation burden of imposed by the nested fixed point method. It also does not require any equilibrium selection constraint because it assumes that there is one equilibrium. In the first step, firms' beliefs about their competitors' actions are estimated through a policy function. In the second step, firms' actions are estimated conditioning on the set of beliefs estimated in the first step.

If all the exogenous variables used for the estimation are market-level variables, the estimation would suffer from the multi-colinearity problem. This is because all the regressors used to estimate beliefs in the first-step would be a subset of the regressors used in the second step. Bajari et al.(2006) suggest that, for an estimation to be identified, the exclusion restriction has to be satisfied. That is, at least one exogenous variable has to appear only in the first step. To be consistent with the model setup, this variable(s) should be observed by all firm, but only affects one's own firm's profit, not its rivals' profits.

On the other hand, if the exogenous variables used for the estimation are specific to firm and/or to other markets, the exclusion restriction is no longer needed. This is because the *values* of the regressors used to estimate beliefs in the first step would not be a subset of the regressors used in the second step. This type of identification is acheived by Ellickson and Misra(2008). My study falls into this case where the exogenous variables used in the first step contain both market-specific and other-market-specific variables. For example, the variable $\hat{n}_{g,d}$ will be estimated in the first step using each competitor hotel's market-specific variables. However, the set of market-specific variables used to estimate the number of rival $\hat{n}_{g,d}$ may not be the same as that used to estimate the own-hotel quality choice.

The econometric specification is defined in a similar fashion to the new hotels approach. I also include the same proxy variable and a random intercept to account for unobserved heterogeneity. However, since all firms are assumed to execute their quality choice simultaneously, time superscript are no longer required. The profit function of hotel j in market m is, then, defined as:

$$\Pi^{m}_{j(static_Bayesian)} = \Phi(\ \eta^{m} + \sum_{l} z^{m}_{l}\tau + \sum_{k} x_{k}\beta_{k} + \sum_{g}\sum_{d} n_{g,d}\theta_{g,d}),\ \forall(j \in J)$$

where subscript $j = 1, ..., J$ denotes all hotels in Texas at the end of 2002. All other variables are as defined previously. According to eq.(3.39), the probit probability of choose each quality-type can be expressed as:

$$\begin{aligned} P^{m}_{j}(g_{j} &= hi) = \int (\Phi(\eta^{m} + \sum_{l} z^{m}_{l}\tau + \sum_{k} x_{k}\beta_{k} + \sum_{g}\sum_{d} n_{g,d}\theta_{g,d}))\ \phi(\eta^{m})\ d\eta^{m} \quad (3.42) \\ P^{m}_{j}(g_{j} &= lo) = 1 - P^{m}_{j}(g_{j} = hi). \end{aligned}$$

For notational conciseness, I will use P_j^m interchangeably with $P_j^m(g_j = hi)$. This renders $P_j^m(g_j = lo) \equiv 1 - P_j^m$. The probability P_j^m of all hotels are simultaneously determined by assumption of the static Bayesian game approach. Therefore, all the $n_{g,,d}$ are endogenous because they are functions of P_j^m.

To overcome the endogeneity (of $n_{g,,d}$) problem, I first estimate the *expected* value of P_j^m using all the state (exogenous) variables $(x_k, k = 1, ..., K)$. A consistent estimate of P_j^m is called $\hat{P}_j^m$ and can be obtained through various estimation techniques. For example, Bajari et al.(2006) use the semi-parametric sieve estimator, while Ellickson and Misra(2008) use the same parametric specification as the main model but include interactions and squares to improve the fit. When possible, the nonparametric or semi-parametric approach such as sieve, kernel and spline estimation are preferred because the predicted beliefs $(\hat{P}_j^m)$ would be less sensitive to misspefication. In my case, however, the nonparametric estimation is not feasible due to a limited number of observations (4,360 observations) and the high dimensionality of exogenous variables (over 30, including dummy variables for popular cities). I, therefore, adopt Ellickson and Misra(2008)'s solution which uses the same parametric specification as the estimating model (eq.3.42) but exclude the strategic variables $(n_{g,d})$. The goodness of fit is also improved by including many interaction terms and squares.

- The solution to the first-step (policy function) estimation is obtained through the maximum likelihood estimation. The log-likelihood function can be defined as:

$$\log L = \sum_{j=1}^{J} y_{j,hi} \log \tilde{P}_j^m + \sum_{j=1}^{J} y_{j,hi} \log(1 - \tilde{P}_j^m) \tag{3.43}$$

where $\tilde{P}_j^m$ is the first-step probability, or the policy function which contains all the exogenous variables as follows:

$$\tilde{P}_j^m = \tilde{P}_j^m(g_j = hi) = \int \Phi(\eta^m + \sum_l z_l^m \tau + \sum_k x_k \beta_k + f(x_1, ..., x_k))\ \phi(\eta^m)\ d\eta^m$$

where $f(x_1, ..., x_k)$ indicates a function of interactions and squares of x_k,the exogenous variables. I use the same maximisation routine as in the new hotel approach to obtain parameter values which maximize the log-likelihood function (eq.3.43). I then, calculate the predicted probability of choosing high quality by each hotel $(\hat{P}_j^m)$. These predicted probabilities are used to calculate $\widehat{n}_{g,,d}$– the expected number of g-quality rival hotels within the d-mile distance band.

$$\hat{P}_j^m(g_j = hi) = \int \Phi(\hat{\eta}^m + \sum_l z_l^m \hat{\tau} + \sum_k x_k \hat{\beta}_k + f(\widehat{x_1, ..., x_k}))\ \phi(\hat{\eta}^m)\ d\hat{\eta}^m \quad (3.44)$$

$$\widehat{n}_{hi,1} = \sum_{l \in l_1} \hat{P}_l^m \quad , \forall\ l \neq j \quad (3.45)$$

$$\widehat{n}_{lo,1} = \sum_{l \in l_1} (1 - \hat{P}_j^m)\ , \forall\ l \neq j \quad (3.46)$$

$$\widehat{n}_{hi,d} = \sum_{l \in l_d} \hat{P}_l^m \quad , \forall\ l \text{ and } d > 1 \quad (3.47)$$

$$\widehat{n}_{lo,d} = \sum_{l \in l_d} (1 - \hat{P}_l^m)\ , \forall\ l \text{ and } d > 1 \quad (3.48)$$

where $\widehat{n}_{g,d}$ is a consistent estimate of $n_{g,d}$, $\widehat{n}_{hi,1}$ is a consistent estimate of the number of high-quality competitors with in 1-mile concentric distance band. $\widehat{n}_{lo,1}$ is a consistent estimate of the number of low-quality competitors with in 1-mile concentric distance band. $\widehat{n}_{hi,d}$ is a consistent estimate of the number of high-quality hotels located in the d-mile concentric distance band $(d > 1$ mile), and $\widehat{n}_{lo,d}$ is a consistent estimate of the number of low-quality hotels located in the d-mile concentric distance band. Since I only need to calculate the expected number of

"competitors", $\widehat{n}_{hi,1}$ and $\widehat{n}_{lo,1}$ in eq. 3.45 and eq. 3.46 do not count one's own hotel $(\forall\, l \neq j)$.

- In the second step, I replace $n_{g,,d}$ with $\widehat{n}_{g,,d}$ in the main specification (3.42). Then, I use the same maximisation routine to obtain the parameter estimates. In this case $\widehat{n}_{g,,d}$, the strategic variables, is no longer subject to the simultaneity bias. The probit probability can be defined as:

$$P_j^m(g_j = hi) = \int (\Phi(\eta^m + \sum_l z_l^m \tau + \sum_k x_k \beta_k + \sum_g \sum_d \hat{n}_{g,,d}\theta_{g,d}))\phi(\eta^m)d\eta^m.$$

Results (Static Bayesian Game Approach)

Table(3.32) and (3.30) report variable definition and summary statistics; while table(3.34) sets out the two-step probit estimation results. For consistency, I use the same set of explanatory variables as in the new hotels approach.

The simultaneity bias of the strategic variables, $n_{lo,0-2}$, $n_{lo,2-5}$, $n_{hi,0-2}$ and $n_{hi,2-5}$ is remedied by using a policy function to estimate these variables in the first step. The estimated variables are denoted $\hat{n}_{lo,0-2}$, $\hat{n}_{lo,2-5}$, $\hat{n}_{hi,0-2}$ and $\hat{n}_{hi,2-5}$. I use the same proxy variables to account for the zip5-specific unobserved attractiveness for high-quality hotels. In addition to the proxy variables, I use a random intercept to also account for the zip5-specific unobserved heterogeneity. The proxy variables and a random intercept are also used in the first step estimation. In all specifications, I bootstrap the coefficients' standard errors because $\widehat{n}_{g,d}$ are estimated variables from the first step. In which case, a standard

calculation of the error terms–i.e. those given by the maximum likelihood estimation–would be incorrect.

Turning to the estimation results, specification 1 does not include any proxy variable. Specification 2 includes only the first proxy variable z_1, specification 3 includes only the second proxy variable z_2, while specification 4 includes both z_1 and z_2. The null hypotheses are rejected at the 5 percent level of significance, unless stated otherwise. Similar to the new hotel approach, I drop hotels with less than 20 rooms and cities with less than 10 hotels. This is to avoid cases with no strategic interactions.

Overall, the estimation results are consistent with those obtained from the new hotels approach. Without including proxies for the unobserved heterogeneity, results in specification 1 suggest a stronger positive spillover effects. Hotels appear to benefit from the presence of competitors of the same quality-type. After using proxy variable(s) to control for the area-specific attractiveness for high-quality hotels, these positive spillovers become weaker, but remain highly significant.

Since specification 4 have controlled for the most unobserved heterogeneity through both proxy variables, and would give the most consistent results, the ensuing discussion will be based on specification 4.

The negative and significant coefficients on $n_{lo,0-2}$ and $n_{lo,2-5}$ suggest that an expectation of one additional low-quality competitor encourages a hotel to choose low quality rather than high quality. However, since $\theta_{lo,0-2} = -0.020$ is more negative than $\theta_{lo,2-5} = -0.016$, one can say that closer competitors are more influential than distance competitors.

The positive and significant coefficient on $n_{hi,0-2}$ implies that an expectation of one additional high-quality competitor encourages a hotel to choose high quality. This impact fades away and becomes insignificant with distance, i.e. the coefficient on $n_{hi,2-5}$ is not statistically significant. In any case, the results suggest that high-quality hotels generate more positive spillovers than low-quality hotels counterpart.

Turning to the location-specific effects, all the significant coefficients take the same signs as those obtained from the new hotels approach. As for the insignificant coefficients, most of them take the same sign under both approaches. Looking only at the significant coefficients, the results suggest that high quality hotels benefit more in areas where the household size is smaller ($\beta_{household\ size} = -0.318$), where there are more retail establishments psm ($\beta_{retail} = 0.024$), more professional establishments psm ($\beta_{professional} = 0.017$), less educational establishments psm ($\beta_{education} = -0.287$), and less art and recreational establishments psm ($\beta_{art} = -0.257$).

3.5.5 Robustness Check

The results from both the new hotels approach and the static Bayesian approach suggest that hotels of the same quality-type enjoy more positive spillovers from presence of one another. In both estimations, the strategic effects (or spillover effects) are accounted for through the number of hotels within the zero to two-mile distance band ($n_{g,0-2}$), and the number of hotels within the two to five-mile distance band ($n_{g,2-5}$). Are these variables reasonable? What if it is the case that other variables better represent what firms care, such as the number high- and low-quality rooms within some certain distance bands? To

Table 3.32. Variable Definitions and Sample Characteristics

Variable Name	Definition	Mean and (Standard Deviation)
z_1	The first proxy variable defined in eq.3.37, $z_1 = \frac{high_room_sold^c}{all_room_sold^c}$	0.562 (0.154)
z_2	The second proxy variable defined in eq.3.38, $z_2 = \frac{high_room_sold^{zip} - low_room_sold^{zip}}{all_room_sold^{zip}}$	-0.041 (0.586)
$\widehat{n}_{lo,0-2}$	predicted number of low-quality incumbents located within 0-2 mile distance band	7.934 (5.731)
$\widehat{n}_{lo,2-5}$	predicted number of low-quality incumbents located within 2-5 mile distance band	13.204 (11.185)
$\widehat{n}_{hi,0-2}$	predicted number of high-quality incumbents located within 0-2 mile distance band	6.038 (6.159)
$\widehat{n}_{hi,2-5}$	predicted number of high-quality incumbents located within 2-5 mile distance band	0.959 (3.335)
Expenditure (million $US.)	total expenditure (of all activities) per sq.mile in the zip5 area	9.722 (7.159)
Employment (1,000 people)	number of people employed (in all sectors) per sq.mile in the zip5 area	0.122 (0.218)
Pop density (1,000 people)	number of residents per sq.mile in the zip5 area	2.654 (2.150)
Population (million people)	total number if residents in the zip5 area	0.282 (0.158)
House tax (10,000 $US/year)	median annual real-estate tax in the zip5 area	0.171 (0.120)
Household size (thousand people)	average number of household members (adult & children) in the zip5 area	2.585 (0.515)
Crime index	The sum of seven types of offenses known to law enforcement, namely, murder, forcible rape, robbery, aggravated assault, burglary, larceny-theft and motor vehicle theft.	5.409 (5.639)
Retail (no./sq.mile)	number of retail establishments per sq.mile in the zip5 area	14.699 (23.533)
Professional (no./sq.mile)	number of professional establishments per sq. mile in the zip5 area	28.415 (228.683)

Table 3.33. Variable Definitions and Sample Characteristics (continued)

Variable Name	Definition	Mean and (Standard Deviation)
Educational (no./sq.mile)	number of schools, educational establishment per sq. mile in the zip5 area	0.875 (4.761)
Entertainment (no./sq.mile)	number of entertainment, art establisment per sq. mile in the zip5 area	1.615 (11.027)
Restaurant (no./sq.mile)	number of restaurant, food establishment per sq. mile in the zip5 area	10.004 (33.803)
Salon (no./sq.mile)	number of beauty salon establishment per sq. mile in the zip5 area	1.168 (2.403)
Admin (no./sq.mile)	number of administrative establishment per sq. mile in the zip5 area	7.652 (53.291)
Urban	proportion of urban area per total county area	0.718 (0.363)
Black	proportion of black population in the county	0.124 (0.074)
Income (10,000 $US/year)	median annual income of residents in the county	4.062 (0.885)
City Dummies (for 6 popular cities)	0-1 dummy variables for Austin, Houston, San Antonio, Fort Worth, Dallas, Irving	–

Table 3.34. Probit Estimation Result: Static Baysian Game Approach

	(1)	(2)	(3)	(4)
constant	-0.357 (0.330)	-0.644**(0.357)	-0.329 (0.288)	-0.201 (0.317)
Proxy variables				
z_1	–	0.773* (0.362)	–	-0.299 (0.330)
z_2	–	–	0.751**(0.066)	0.008**(0.001)
Strategic effects				
$\hat{n}_{lo,0-2}$	-0.046**(0.009)	-0.039**(0.010)	-0.017**(0.008)	-0.020**(0.008)
$\hat{n}_{lo,2-5}$	-0.023**(0.007)	-0.021**(0.007)	-0.015**(0.006)	-0.016**(0.006)
$\hat{n}_{hi,0-2}$	0.083 **(0.010)	0.077**(0.010)	0.044**(0.009)	0.046**(0.009)
$\hat{n}_{hi,2-5}$	0.015**(0.007)	0.012* (0.007)	0.007 (0.006)	0.008 (0.006)
Zip5 characteristics				
Expenditure	-0.024**(0.010)	-0.023**(0.010)	-0.012 (0.009)	-0.013 (0.010)
Employment	0.269 (0.350)	0.267 (0.359)	0.239 (0.321)	0.237 (0.316)
Pop density	-0.046 (0.036)	-0.043 (0.036)	-0.021 (0.034)	-0.021 (0.035)
Population	0.237 (0.281)	0.252 (0.282)	0.325 (0.252)	0.327 (0.254)
House tax	1.489**(0.407)	1.559**(0.409)	0.880**(0.365)	0.848**(0.371)
House size	-0.141 (0.098)	-0.139 (0.099)	-0.129 (0.086)	-0.131 (0.087)
Retail	0.014**(0.006)	0.014**(0.006)	0.016**(0.005)	0.016**(0.005)
Professional	0.003 (0.004)	0.003 (0.004)	0.006* (0.003)	0.007* (0.003)
Education	0.067 (0.073)	0.065 (0.073)	0.039 (0.066)	0.037 (0.065)
Art	-0.124**(0.056)	-0.125**(0.056)	-0.076 (0.050)	-0.074 (0.050)
Restaurant	0.002 (0.010)	0.002 (0.010)	-0.014 (0.009)	-0.014 (0.009)
Salon	-0.132**(0.049)	-0.128**(0.049)	-0.112**(0.044)	-0.112**(0.044)
Admin	0.009 (0.012)	0.010 (0.012)	-0.006 (0.011)	-0.007 (0.010)
City characteristics				
Crime index	0.035 (0.072)	0.012 (0.073)	0.038 (0.062)	0.045 (0.064)
County characteristics				
Black	0.214 (0.609)	-0.055 (0.624)	-0.300 (0.543)	-0.218 (0.553)
Urban	0.205 (0.172)	0.150 (0.174)	0.104 (0.157)	0.134 (0.160)
Income	0.046 (0.047)	0.024 (0.048)	0.041 (0.041)	0.046 (0.042)
City dummies	yes	yes	yes	yes
No. of obs.	2127	2127	2127	2127
log-like	-1235.421	-1233.069	-1182.292	-1169.655

Notes: ** significant at 5%, * significant at 10%, all the standard errors were obtained through a bootstrapping process. All specifications allow for a random intercept.

this effect, I conduct a robustness check which replaces $n_{g,0-2}$ and $n_{g,2-5}$ with $r_{g,0-2}$ and $r_{g,2-5}$. The variable $r_{g,0-2}$ denotes the number of g-quality *rooms* within the zero to two-mile distance band, while the variable $r_{g,2-5}$ denotes the number of *rooms* within the two to five-mile distance band.

The results are reported in table 3.35, the first column uses the new hotel approach, while the third column uses the static Bayesian approach. These two specifications are almost identical to specification 4 in table 3.31 and table 3.34, except the variables $n_{g,0-2}$ and $n_{g,2-5}$ are replaced by $r_{g,0-2}$ and $r_{g,2-5}$. Here after, I will call specification 4 in table 3.31 and table 3.34 my "main specification". As was done in my main specification, I include both proxy variables and a zip5-specific random intercept.

The results do not contradict with my main specifications. However, some coefficients on the strategic variables become insignificant. Under the new hotel approach, the coefficients on the number of low-quality rooms within the zero to two-mile distance band $(n_{lo,0-2})$ and two to five-mile distance band $(n_{lo,2-5})$ are not statistically significant. This may have been because the sample size is small, or the "number of rooms" do not matter to a hotels' quality choice as much as the "number of hotels". Turning to the static Bayesian approach, the coefficients on $\hat{r}_{lo,2-5}$ and $\hat{r}_{hi,0-2}$ are significant, and have the same sign as those in the main specification. The coefficients on $\hat{r}_{lo,0-2}$ and $\hat{r}_{hi,2-5}$, however, are not statistically significant.

As for the location-specific effects, the coefficients that are statistically significant have the same sign as those in the main specification. This suggests that using the number

of rooms, instead of the number of hotels, do not change my results. However, it would not improve the empirical specification either.

In addition to the first robustness check, which examines the definition of my strategic variables, the second robustness check examines the sample selection issue. For the main specifications of both the new hotels approach and the static Bayesian approach, I drop hotels with less than 20 rooms and cities which contain less than 10 hotels. This is to ensure that all hotels in my sample are subject to the strategic interactions (there are some hotels located within a close proximity). To examine whether the results can generally be applied to small cities with less than 10 hotels, I add them back to my robustness check's specification. In any case, small hotels with less than 20 rooms are still discarded. This is because I believe that they are too small to create any positive or negative spillovers to other hotels.

The results are reported in the first and the fourth column of table 3.35. The sample size increases from 562 hotels to 758 hotels under the new hotels approach, and from 2,127 hotels to 2,723 hotels under the static Bayesian approach. The inclusion of small cities does not give contradicting results to my main specifications. I still find that a given hotel creates more positive spillovers to other hotels of the same quality-type.

One caveat should be noted. That is, even after my best attempts to account for the unobserved heterogeneity through proxy variables and a market-specific random intercept, I still cannot completely rule out the possibility of the unobserved heterogeneity bias. While many additional effective remedies are available, the nature of my data set does not allow for the utilization of them. For example, the fixed-effect estimation, which subtracts out

all the time-invariant area-specific heterogeneity, cannot be employed. This is because, although I observe hotels' actions for 14 half-yearly periods, almost all hotels made their quality decision only once and never switch. Since my data set does not have a panel structure, panel estimation techniques cannot be used.

3.6 Conclusion

This study explores empirical strategies to measure the strength of location-specific and strategic effects in determining quality choice of lodging establishments in Texas during the period 1998-2004. Its contribution is twofold. First, it compares results from two different solutions to the simultaneity of strategic actions problem–namely, the new hotels approach, and the static Bayesian game approach. Second, it shows that the quality agglomeration of lodging establishments in Texas is due to both the location-specific effects and the positively spillovers effects.

By using the Multinomial Test for Agglomeration and Dispersion (MTAD) in section 3.3.4, I find that hotels are more agglomerated in quality-type than they would had they made their choices randomly and independently of one another. After having confirmed the agglomeration in quality-type phenomenon by MTAD, section 3.3.5 examines the determinants of this agglomeration outcome.

In doing so, I divide the determinants of hotels' quality choices into "location-specific effects" and "strategic effects" (or positive spillovers effects). Location-specific effects are accounted for by observed and unobserved attributes of each market. I include proxy variables and a random intercept for each zip5 market to account for market-specific unob-

Table 3.35. Robustness Check

	New Hotels Approach		Static Bayesian Approach	
	(room)	(all sample)	(room)	(all sample)
constant	0.665 (0.780)	1.317**(0.660)	-0.555* (0.308)	-0.451* (0.270)
Proxy variables				
z_1	0.155 (0.658)	0.139 (0.337)	0.038 (0.325)	-0.057 (0.207)
z_2	1.414**(0.213)	1.343**(0.202)	0.847**(0.068)	0.802**(0.064)
Strategic effects:				
$n_{lo,0-2}$ or $\hat{n}_{lo,0-2}$	–	-0.027**(0.006)	–	-0.013* (0.008)
$n_{lo,2-5}$ or $\hat{n}_{lo,2-5}$	–	0.003 (0.005)	–	-0.002 (0.007)
$n_{hi,0-2}$ or $\hat{n}_{hi,0-2}$	–	0.034**(0.007)	–	0.025**(0.009)
$n_{hi,2-5}$ or $\hat{n}_{hi,2-5}$	–	-0.001 (0.005)	–	-0.009 (0.007)
$r_{lo,0-2}$ or $\hat{r}_{lo,0-2}$	-0.488 (0.616)	–	0.923 (0.833)	–
$r_{lo,2-5}$ or $\hat{r}_{lo,2-5}$	0.175 (0.568)	–	-1.337**(0.611)	–
$r_{hi,0-2}$ or $\hat{r}_{hi,0-2}$	0.846**(0.402)	–	0.637**(0.311)	–
$r_{hi,2-5}$ or $\hat{r}_{hi,2-5}$	0.075 (0.260)	–	0.225 (0.182)	–
Zip5 characteristics				
Expenditure	0.038 (0.037)	0.049 (0.036)	-0.012 (0.010)	-0.005 (0.009)
Employment	-20.444(33.892)	-8.332(32.915)	-0.197 (0.532)	-0.526 (0.494)
Pop density	-0.139 (0.147)	-0.197 (0.144)	-0.033 (0.036)	-0.033 (0.034)
Population	1.309 (1.262)	0.171 (1.022)	0.527**(0.258)	0.324 (0.214)
House tax	0.135 (0.748)	0.118 (0.697)	0.986**(0.382)	1.245 (0.343)
House size	-0.252 (0.180)	-0.189 (0.160)	-0.104 (0.094)	-0.049 (0.083)
Retail	0.024*(0.014)	0.021 (0.013)	0.019**(0.005)	0.013**(0.005)
Professional	0.018**(0.008)	0.012 (0.008)	0.008**(0.003)	0.005 (0.003)
Education	-0.261*(0.148)	-0.265* (0.147)	0.117 (0.072)	0.099 (0.069)
Art	-0.306**(0.143)	-0.246* (0.137)	-0.115**(0.055)	-0.086* (0.051)
Restaurant	-0.001 (0.024)	0.008 (0.023)	-0.019 (0.010)	-0.015* (0.009)
Salon	-0.129 (0.117)	-0.161 (0.114)	-0.116**(0.045)	-0.092* (0.043)
Admin	0.012 (0.033)	0.010 (0.033)	-0.001 (0.013)	0.014 (0.012)
City characteristics				
Crime index	-0.103 (0.143)	-0.024 (0.126)	0.008 (0.065)	-0.011 (0.059)
County characteristics				
Black	-1.397 (1.043)	-1.628**(0.827)	-0.328 (0.566)	-0.674 (0.455)
Urban	-0.084 (0.316)	-0.355 (0.266)	0.143 (0.162)	0.131 (0.142)
Income	-0.017 (0.084)	-0.054 (0.063)	0.051 (0.043)	0.026 (0.034)
City dummies	yes	yes	yes	yes
No. of obs.	562	758	2127	2723
log-like	-315.720	-413.255	-1176.021	-1522.234

Notes: ** significant at 5%, * significant at 10%, all the standard errors under the standard Bayesian approach were obtained through a bootstrapping process. All specifications allow for a random intercept.

served heterogeneity. Since my data does not have a panel structure, it was not possible to employ the fixed-effect estimation technique. On the other hand, strategic effects are accounted for by actions or expected actions of competitors. Since firms actions could influence the actions of their competitors, and vice-versa, my strategic variables suffer simultaneity bias. This problem was mitigated through two estimation strategies–namely the new hotels approach, and the strategic Bayesian game approach with the two-step estimation technique.

My results from both approaches suggest that firms make their quality choices based on both location-specific and positive spillover effects. Discussing on the ground of location-specific effects, high-quality hotels benefit more from areas with a smaller average household size, more retail establishments, and more professional establishments. Compared to low-quality hotels, they hurt less from a high real estate tax rate. In general, the characteristics that favour high-quality hotels are more likely to be found in downtown and/or business areas. Low quality hotels, on the other hand, are more attracted to areas with a lower real estate tax rate, a higher average household size, more art and recreational establishments, and more educational establishments. These characteristics are usually found in less business-oriented areas and/or outdoor recreational attractions such as theme parks, national parks, for example.

Turning to the strategic effects, my results imply that firms with the same quality-type benefit more from the presence of one another. This finding is consistent with other studies on firms' strategic interaction which support the existence of positive spillovers. For example, Toivanen and Waterson (2005) and Ellickson and Misra (2008). It is also

consistent with other studies on hotel agglomeration such as Enz et al.(2005), Kalnins and Chung (2004), and Baum and Haveman (1997).

References

Aguirregabiria, V. and P. Mira (2007). "Sequential Estimation of Dynamic Discrete Games, ." Econometrica 75(1): 1-53.

American Automobile Association (AAA) (2004). Texas TourBook. Heathrow, Florida, AAA Publishing.

Asche, F. and C. R. Wessells (1997). "On Price Indices in the Almost Ideal Demand System." American Journal of Agricultural Economics 79(4): 1182-1185.

Augereau, A., S. Greenstein, et al. (2004). "Coordination vs. Differentiation in a Standards War: 56K Modems." NBER Working Paper Series (No.10334).

Bajari, P., L. Benkard, et al. (2007). "Estimating Dynamic Models of Imperfect Competition." Econometrica 75(5): 1331-1370.

Bajari, P., H. Hong, et al. (2006). "Estimating Static Models of Strategic Interaction." NBER Working Paper Series(No. W12013).

Bajari, P. and L. Ye (2003). "Deciding Between Competition and Collusion." Review of Economics and Statistics 85(4): 971-989.

Baker, N. (2004). How Green is Your Supermarket? Green Liberal Democrats Party. London, UK.

Baldwin, L. H., R. C. Marshall, et al. (1997). "Bidder Collusion at Forest Service Timber Auctions." Journal of Political Economy 105(4): 657-699.

Banerjee, A., J. Dolado, et al. (1993). Co-integration, Error Correction, and the Econometric Analysis of Non-stationary Data. Oxford, Oxford University Press.

Banerji, A. and J. V. Meenakshi (2004) Millers, Commission Agents and Collusion in Grain Auction Markets: Evidence from Basmati Auctions in North India: Working Paper No.129

Bayer, P. and C. Timmins (2007). "Estimating Equilibrium Models of Sorting Across Locations." Economic Journal 117(518): 353-374.

Belaire-Franch, J. and C. D. (2002). "A Pearson's Test for Symmetry with an Application to the Spanish Business Cycle." Spanish Economic Review 4: 221-238.

Berry, S. (1994). "Estimating Discrete-Choice Models of Product Differentiation." The RAND Journal of Economics 25(2): 242-262.

Berry, S., J. Levinsohn, et al. (1995). "Automobile Prices in Market Equilibrium." Econometrica 60(4): 889-917.

Berry, S., J. Levinsohn, et al. (2004). "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market." Journal of Political Economy 112(1): 68-104.

Berry, S. and J. Waldfogel (1999). "Free Entry and Social Inefficiency in Radio Broadcasting." RAND Journal of Economics 30(3): 397-420.

Blackburn, M. and D. Neumark (1992). "Unobserved Ability, Efficiency Wages, and Interindustry Wage Differentials." Quarterly Journal of Economics 107(4): 1421-1236.

Bliss, C. (1988). "A Theory of Retail Pricing." The Journal of Industrial Economics 36(4): 375-391.

Block, M. K., F. C. Nold, et al. (1981). "The Deterrent Effect of Antitrust Enforcement." Journal of Political Economy 89(3): 429-445.

Blundell, R., P. Pashardes, et al. (1993). "What do we Learn About Consumer Demand Patterns from Micro Data?" The American Economic Review 83(3): 570-597.

Bonnet, C. and P. Dubois (2007). Non Linear Contracting and Endogenous Buyer Power between Manufacturers and Retailers: Identification and Estimation on Differentiated Products. mimeo, Toulouse School of Economics.

Bresnahan, T. F. (1989). Empirical Studies of Industries with Market Power. The Handbook of Industrial Organization. R. Schmalensee and R. D. Willig. New York, North-Holland.

Bresnahan, T. F. and P. Reiss (1991). "Entry and Competition in Concentrated Markets." Journal of Political Economy 99(51): 977-1009.

Breusch, T. S. and A. R. Pagan (1979). "A Simple Test for Heteroscedasticity and Random Coefficient Variation." Econometrica 47(5): 1287-1294.

Brock, W. and S. Durlauf (2000). "Interactions-Based Models." NBER Working Paper Series(No. 258).

Buam, J. A. C. and H. A. Haveman (1997). "Love Thy Neighbor? Differentiation and Agglomeration in the Manhattan Hotel Industry, 1898-1990." Administrative Science Quarterly 42(2): 304-338.

Bushnell, J. B., E. T. Mansur, et al. (2008). "Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets." The American Economic Review 98(1): 237-266.

Cameron, A. C. and P. K. Trivedi (2005). Microeconometrics: Methods and Applications. New York, Cambridge University Press.

Casella, G. and R. Berger (1990). Statistical Inference. Belmont, CA, Duxbury Press.

Cho, E. (2007). Store Brand and National Brand Demand Characteristics: The Case of Over-The-Counter Pain Relievers. mimeo. Cambridge, Massachusett, Department of Economics, MIT.

Clay, K. and W. Troesken (2003). "Further Tests of Static Oligopoly Models: Whiskey, 1882-1898." The Journal of Industrial Economics LI(2): 151-166.

Competition Commission (CC). (2000). "Chapter 4: Structure of the Markets and Participants." Scottish Milk: A report on the supply of fresh processed milk to middle-ground retailers in Scotland Retrieved 16 June, 2008, from http://www.competition-commission.org.uk/inquiries/completed/2000/index.htm#supply.

Competition Commission (CC). (2008). "Appendix 3.1: Overview of UK Grocery Retailers." Groceries Market Investigation Retrieved 22 May, 2008, from http://www.competition-commission.org.uk/rep_pub/reports/2008/fulltext/538_3_1.pdf.

Competition Commission (CC). (2007). "Appendix 8.3: Milk Supply Chain Profitability." Groceries Market Investigation: Provisional Findings Report Retrieved 12 May, 2008, from http://www.competition-commission.org.uk/inquiries/ref2006/grocery/prov_findings/prov_find_app_8_3.pdf.

Competition Commission (CC). (2003a). "Chapter 3: The fresh processed milk industry." Arla Foods amba and Express Dairies plc: A report on the proposed merger Retrieved 1 June, 2008, from http://www.competition-commission.org.uk/inquiries/completed/2003/arla/index.htm.

Competition Commission (CC). (2003b). "Chapter 5: The market for fresh processed milk." Arla Foods amba and Express Dairies plc: A report on the proposed merger Retrieved 1 June, 2008, from http://www.competition-commission.org.uk/inquiries/completed/2003/arla/index.htm.

Corts, K. S. (1999). "Conduct Parameters and the Measurement of Market Power." Journal of Econometrics 88: 227-250.

Dairy Co (2009). Dairy Statistics: An Insider's Guide 2009. Kenilworth, DairyCo, Agriculture and Horticulture Development Board.

Davidson, R. and J. G. MacKinnon (1981). "Several Tests for Model Specification in the Presence of Alternative Hypotheses." Econometrica 49(3): 781-793.

Davidson, R. and J. G. MacKinnon (1993). Estimation and Inference in Econometrics. Oxford, Oxford University Press.

Deaton, A. and J. Muellbauer (1980a). Economics and Consumer Behavior. Cambridge, Cambridge University Press.

Deaton, A. and J. Muellbauer (1980b). "An Almost Ideal Demand System." The American Economic Review 70(3): 312-326.

Department for Environment Food and Rural Affairs (DEFRA) (2007). The Environmental, Social and Economic Impacts Associated with Liquid Milk Consumption in the UK and its Production: A Review of Literature and Evidence. London, DEFRA.

Draganska, M. (2008). "Discrete Choice Models of Firms' Strategic Decisions." Marketing Letters 19(3): 399-416.

Dranove, D., A. Gron, et al. (2003). "Differentiation and Competition in HMP Markets." The Journal of Industrial Economics LI(4): 433-453.

Eaton, B. C. L., R. G. (1975). "The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition." Review of Economic Studies 42(1): 27-49.

Edwards, S. (2004): "Thirty Years of Current Account Imbalances, Current Account Reversals and Sudden Stops," NBER Working Paper Series (No.10276).

Ellickson, P. B. (2004). Supermarket as a Natural Oligopoly. Mimeo. Durham, NC, Duke University.

Ellickson, P. B. and S. Misra (2008). "Supermarket Pricing Strategies." Marketing Science 27(5): 811-828.

Ellison, G. and E. L. Glaeser (1997). "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." Journal of Political Economy 105(889-927).

Employment and Training Institute (ETI) (2004). ETI Purchasing Power, Business Activity, and Workforce Density Profiles for All Residential ZIP Codes in U.S., Employment and Training Institute, University of Wisconsin Milwaukee.

Empson, J. (1998). "The History of the Milk Marketing Board, 1933-1994: British farmers' greatest commercial enterprise." International Journal of Dairy Technology 51(3): 77-85.

Engle, R. F. (1982). "Autoregressive conditional heteroscedasticity, with estimates of the variance of United Kingdom inflation." Econometrica 50: 987-1007.

Enz, C. A., L. Canina, et al. (2005). How Co-location Helps Some Hotels and Hurts Others. The Agglomeration Conundrum. The Center for Hospitality Research, Cornell University.

Evans, W. N., W. E. Oates, et al. (1992). "Measuring Peer Group Effects: A Study of Teenage Behavior." Journal of Political Economy 100(5): 966-991.

Fischer, J. H. and J. Harrington, J. E. (1996). "Product Variety and Firm Agglomeration." The RAND Journal of Economics 27(2): 281-309.

Franks, J. (2001). "Developments in Milk Marketing in England and Wales during the 1990s." British Food Journal 102(9): 631-643.

Funderburk, D. R. (1974). "Price Fixing in the Liquid-Asphalt Industry: Economic Analysis versus the 'Hot Document'." Antitrust Law and Economics Review 7: 61-74.

Genesove, D. and W. P. Mullin (1998). "Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914." The RAND Journal of Economics 29(2): 355-377.

Godfrey, L. G. (1978). "Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables." Econometrica 46: 1303-1313.

Goff, D. (2006). "Introduction to Dairy Science and Technology: Milk History, Consumption, Production, and Composition." Retrieved 20 May, 2008.

Gorman, W. M. (1959). "Separable Utility and Aggregation." Econometrica 27(469-481).

Graddy, K. (1995). "Testing for Imperfect Competition at the Fulton Fish Market." The RAND Journal of Economics 26(1): 75-92.

Green, E. J. and R. H. Porter (1984). "Noncooperative Collusion under Imperfect Price Information." Econometrica 52(1): 87-100.

Haltiwanger, J. and J. E., Jr. Harrington (1991). "The Impact of Cyclical Demand Movements on Collusive Behavior." The RAND Journal of Economics 22(1): 89-106.

Harrington, J. E., Jr. (1987). "Collusion in Multiproduct Oligopoly Games under a Finite Horizon." International Economic Review 28(1): 1-14.

Harrington, J. E., Jr. (2004). "Cartel Pricing Dynamics in the Presence of an Antitrust Authority." The RAND Journal of Economics 35(4): 651-673.

Harrington, J. E., Jr. (2005). Detecting Cartels. Baltimore, Department of Economics, Johns Hopkins University.

Harrington, J. E., Jr. (2006). Modelling the Birth and Death of Cartels with an Application to Evaluating Antitrust Policy. Economics Working Paper Archive 532, The Johns Hopkins University, Department of Economics.

Hausman, J. A. (1997). "Valuing the Effect of Regulation on New Services in Telecommunications." Brookings Papers on Economic Activity. Microeconomics 1997: 1-54.

Hausman, J. A. and E. Leibtag (2005). Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart. NBER Working Paper No.11809.

Hay, D. A. (1976). "Sequential Entry and Entry-Deterring Strategies in Spatial Competition." Oxford Economic Paper 28(2): 240-257.

Hoderlein, S. G. N. and A. Lewbel (2007). Price Dimension Reduction in Demand Systems With Many Goods, Mannheim University and Boston College.

Holmes, T. J. (2008). "The Diffusion of Wal-Mart and Economies of Density." NBER Working Paper Series (No. 13783).

Hotelling, H. (1929). "Stability in Competition." Economic Journal 39(41-57).

Huber, P. J. (1967). The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability. California, University of California Press. 1: 221-233.

InterContinental Hotel Group (IHG). (2009). "Company Overview." Retrieved 10 October 2009, from http://www.ihgplc.com/index.asp?pageid=16.

Jia, P. (2008). "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry." Econometrica 76(6): 1263-1316.

Johnston, J. and J. DiNardo (1997). Econometric Methods. Singapore, McGraw-Hill.

Judd, K. and C.-L. Su (2007). Constrained Optimization Approaches to Estimation of Structural Models. Working Paper, Hoover Institution.

Kalnins, A. and W. Chung (2004). "Resource-Seeking Agglomeration : A Study of Market Entry in the Lodging Industry " Strategic Management Journal 25: 689-699.

Kim, D.-W. and C. R. Knittel (2006). "Biases in Static Oligopoly Models? Evidence From the California Electricity Market." The Journal of Industrial Economics LIV(4): 451-470.

Klemperer, P. D. and M. A. Meyer (1989). "Supply Function Equilibria in Oligopoly Under Uncertainty." Econometrica 57(6): 1243-1277.

Krugman, P. (1991). "Increasing Returns and Economic Geography." Journal of Political Economy 99(3): 483-499.

Lechene, V. (2000). Section 6: Income and Price Elasticities of Demand for Food Consumed in the Home. Report of the National Food Survey Committee - 2000/01. The National Food Survey (NFS). UK, the Department for Environment, Food and Rural Affairs (DEFRA).

Lehmann-Grube, U. (1997). "Strategic Choice of Quality When Quality is Costly: The Persistence of the High-Quality Advantage." RAND Journal of Economics 28(2): 372-384.

Levinsohn, J. and A. Patrin (2003). "Estimating Production Functions Using Inputs to Control for Unobservables." Review of Economic Studies 70(317-341).

Lewbel, A. (1996). "Aggregation Without Separability: A Generalized Composite Commodity Theorem." The American Economic Review 86(3): 524-543.

Mas-Collel, A., M. D. Whinston, et al. (1995). Microeconomic Theory, Oxford University Press.

Mazzeo, M. J. (2002). "Product Choice and Oligopoly Market Structure." The RAND Journal of Economics 33(2): 221-242.

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. New York, Academic Press.

Milk Development Council (MDC) (2005). Dairy Facts and Figures 2004. Cirencester, Milk Development Council, UK.

Milk Development Council (MDC). (2008a). "Producer Numbers." Farm Data & Prices Retrieved 10 May, 2008, from http://www.mdcdatum.org.uk/FarmDataPrices/producernumbers.html.

Milk Development Council (MDC). (2008b). "Retail and Doorstep." Liquid Milk Expenditure and Volume Retrieved 10 May, 2008, from http://www.mdcdatum.org.uk/RetailerDataPrices/tnsliquidmilkdoorstepandretail.html.

Miller, S. (2004). Prices and Local Market Structure: Evidence from the Lodging Industry. Economics. Oxford, University of Oxford. MPhil.

Missauri Census Data Center (2000). MABLE/Geocorr2K: Geographic Correspondence Engine with Census 2000 Geography, Missauri Census Data Center.

Murphy, K. M. and R. H. Topel (1985). "Estimation and Inference in Two-Step Econometric Models." Journal of Business & Economic Statistics 3(4): 370-379.

Nevo, A. (1998). "Identification of the oligopoly solution concept in a differentiated-products industry." Economics Letters 59: 391-395.

Nevo, A. (2001). "Measuring Market Power in the Ready-to-Eat Cereal Industry." Econometrica 69(2): 307-342.

Newey, W. K. and K. D. West (1987). "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Coveriance Matrix." Econometrica 55(3): 703-708.

Office of Fair Trading (OFT). (2007a). "OFT Issues Provisional Decision Against Supermarkets and Dairies Over Price Fixing." Retrieved 21 May, 2008, from http://www.oft.gov.uk/news/press/2007/134-07.

Office of Fair Trading (OFT). (2007b). "OFT Welcomes Early Resolution Agreements and Agrees Over £116m penalties." Retrieved 21 May, 2008, from http://www.oft.gov.uk/news/press/2007/170-07.

Olley, S. and A. Pakes (1996). "The Dynamics of Productivity in the Telecommunications Equipment Industry." Econometrica 64(6): 1263-1298.

Orhun, Y. (2006). Spatial Differentiation in the Supermarket Industry. Working Paper, GSB University of Chicago.

Pakes, A., M. Ostrovsky, et al. (2004). "Simple Estimators for the Parameters of Discrete Dynamic Games (with Entry/Exit Samples)." NBER Working Paper Series (No.10506).

Pearson, K. (1900). "On the criterion that a Given System of Deviations from the Probable if the case of a Correlated System Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling." The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 50: 157-175.

Pesendorfer, M. (2000). "A Study of Collusion in First-Price Auctions." Review of Economic Studies 67(3): 381-411.

Pesendorfer, M. and P. Schmidt-Dengler (2003). "Identification and Estimation of Dynamic Games." NBER Working Paper Series (No. 9726).

Pinkse, J. and E. M. Slade (2004). "Mergers, brand competition, and the price of a pint." European Economic Review 48: 617-643.

Porter, R. H. and J. D. Zona (1993). "Detection of Bid Rigging in Procurement Auctions." The Journal of Political Economy 101(3): 518-538.

Porter, R. H. and J. D. Zona (1999). "Ohio School Milk Markets: An Analysis of Bidding." The RAND Journal of Economics 30(2): 263-288.

Puller, S. L. (2006). Estimation of Competitive Conduct When Firms Are Efficiently Colluding: Addressing the Corts Critique. mimeo. Texas, Texas A&M University.

Puller, S. L. (2007). "Pricing and Firm Conduct in California's Deregulated Electricity Market." The Review of Economics and Statistics 89(1): 75-87.

Ramsey, J. B. (1969). "Tests for specification errors in classical linear least squares regression analysis." Journal of the Royal Statistical Society B(31): 350-371.

Rhee, B. (1996). "Consumer Heterogeneity and Strategic Quality Decisions." Management Science 42(2): 157-172.

Rivers, D. and Q. H. Vuong (1988). "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." Journal of Econometrics 39(347-366).

Rosen, A. (2007). "Identification and Estimation of Firms' Marginal Cost Functions With Incomplete Knowledge of Strategic Behavior." Cemmap Centre for Microdata Methods and Practice(CWP03/07).

Rotemberg, J. J. and G. Saloner (1986). "A Supergame-Theoretic Model of Price Wars during Booms." The American Economic Review 76(3): 390-407.

Rysman, M. and S. Greenstein (2005). "Testing for Agglomeration and Dispersion." Economics Letter 86(3): 405-411.

Salvo, A. (2004). Inferring Conduct under the Threat of Entry: The Case of the Brazilian Cement Industry. London, London School of Economics and STICERD.

Sargan, J. D. (1958). "The Estimation of Economic Relationships Using Instrumental Variables." Econometrica 26(3): 393-415.

Seim, K. (2006). "An Empirical Model of Firm Entry with Endogenous Product-Type Choices." RAND Journal of Economics 37(3): 619-642.

Shaked, A. and J. Sutton (1982). "Relaxing Price Competition Through Product Differentiation." Review of Economic Studies 49(1): 3-13.

Sinnott, R. W. (1984). "Virtues of the Haversine." Sky and Telescope 68: 159.

Smith, H. and J. Thanassoulis (2008). Bargaining in the Supply Chain: The Case of Liquid Milk. Working Paper. Oxford, University of Oxford.

Sorensen, A. T. (2006). "Social Learning and Health Plan Choice." RAND Journal of Economics 37(4): 1-29.

Source Strategy Incorporated (1998, 1999, 2000, 2001, 2002, 2003, 2004). exas Hotel Performance Factbook. San Antonio, Source Strategies Incorporated.

Stahl, K. (1982). "Differentiated Products, Consumer Search, and Locational Oligopoly." Journal of Industrial Economics 31(1/2): 97-113.

Swann, G. M. P. (1985). "Product Competition in Microprocessors." Journal of Industrial Economics 34(1): 33-53.

Swann, G. M. P. (1986). Quality Innovation: An Economic Analysis of Rapid Improvements in Microelectronic Components. London, Francis Pinter Publishers.

Taubman, P. (1976). "The Determinants of Earnings: Genetics, Family, and Other Environments: A Study of White Make Twins." American Economic Review 66(5): 858-870.

Texas Comptroller's Office (various years). Hotel Occupancy Tax Receipts.

Texas Tourism Office of the Governor (2004). Final Report. Annual Report.

Tirole, J. (2003). The Theory of Industrial Organization. Cambridge, MA, MIT Press.

Toivanen, O. and M. Waterson (2005). "Market Structure and Entry: Where's the Beef?" RAND Journal of Economics 36(3): 680-699.

Train, K. (forthcoming). Discrete Choice Methods with Simulation (2nd edition). Cambridge, Cambridge University Press.

Train, K. E. (2003). Discrete Choice Methods with Simulation. Cambridge, Cambridge University Press.

U.S. Census Bureau (2002). ZIP Code Statistics from the Economic Census, U.S. Census Bureau.

UK Data Archive (UKDA) (2000). National Food Survey, 2000. Study Description/Documentation, UKDA.

Villas-Boas, S. B. and R. Hellerstein (2006). "Identification of Supply Models of Retailer and Manufacturer Oligopoly Pricing." Economics Letters 90: 132-140.

Villas-Boas, S. B. (2007). "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data." Review of Economic Studies 74: 625-652.

Vitorino, M. A. (2007). Empirical Entry Games with Complementarities: An Application to the Shopping Center Industry. Working Paper, GSB University of Chicago.

White, H. (1980). "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity." Econometrica 48: 817-838.

Winter, M. (2008). "Corporatism and Agriculture in the U.K.: The Case of the Milk Marketing Board." Sociologia Ruralis 24(2): 106-119.

Wolfram, C. D. (1999). "Measuring Duopoly Power in the British Electricity Spot Market." The American Economic Review 89(4): 805-826.

Wooldridge, J. M. (2002). Econometric Analysis of Cross Section and Panel Data. Cambridge, MA, MIT Press.