Measuring Market Power and Degree of Collusion of UK Liquid Milk Suppliers*

Preliminary and Incomeplete: Please Do Not Quote or Cite

Wanwiphang Manachotphong†

January 2008

Abstract

In September 2007, the Office of Fair Trading (OFT) accused the major UK supermarkets (Asda, Morrisons, Safeway, Sainsbury and Tesco) and major dairy processors (Arla, Dairy Crest, Wiseman, Lactalis McLelland and The Cheese Company) for fixing price of liquid milk. In response to the case, the majority of the accused party admitted that they were involved in anti-competitive behavior and agreed to pay the combined penalty of over 116 million pound. However, each of them would receive a significant reduction in the financial penalty if they continue to provide full co-operation to the OFT. This paper employs the conduct parameter approach to investigate collusive behavior among UK major milk retailers–Asda, Morrisons, Safeway, Sainsbury and Tesco. It also shows the benefits from using the Principal Component Analysis (PCA) data-reduction method in demand and conduct estimations. We found that despite their confession pertaining to anti-competitive behavior, the observed price levels are closer to the competitive outcome than the perfectly collusive outcome. Firms' ability to set price above marginal cost mainly comes from the inelasticity of consumers' demand rather than from the collusive behavior of firms.

* I am grateful to Dr. Howard Smith and Dr. John Thanassoulis, their guidance and support. I am also grateful to Professor Steve Bond and Dr. Chris Wilson for very useful discussion and suggestions. For financial support, I am grateful to the UK's Milk Development Council who made obtaining this data set possible. Any remaining errors are my own.

† Department of Economics, University of Oxford. Email: wanwiphang.manachotphong @economics.ox.ac.uk.

1 Introduction and Summary

In September 2007, the Office of Fair Trading (OFT) accused the major UK supermarkets (Asda, Morrisons, Safeway, Sainsbury and Tesco) and major dairy processors (Arla, Dairy Crest, Wiseman, Lactalis McLelland and The Cheese Company) for fixing price of liquid milk. In response to the case, the majority of the accused party admitted that they were involved in anti-competitive behavior and agreed to pay the combined penalty of over 116 million pound. However, each of them would receive a significant reduction in the financial penalty if they continue to provide full co-operation to the OFT. (OFT, 2007)

This paper uses the conduct parameter approach to investigate collusive behavior among UK major milk retailers–Asda, Morrisons, Safeway, Sainsbury and Tesco. The conduct parameter approach suggests that firm's ability to fix price above marginal cost comes from three sources–inelastic demand, high market concentration and firms collusiveness. The relation among those factors can be factorized from the following profit-maximizing monopoly condition:

$$P(Q) = MC_i - \theta Q \frac{\partial P(Q)}{\partial q_i} \tag{1}$$

where P denotes price, Q denotes total supply in the market, MC_i denotes marginal cost of firm i, q_i denotes quantity supplied by individual firm i and θ denotes the conduct parameter. This equation are the foundation of the conduct parameter approach. The value of θ ranges from 0 to 1; where θ equals 0 for perfectly competitive environment; while θ equals 1 for perfectly collusive environment. Section 4 will show that equation (1) can also be expressed as:

$$\frac{p - MC_i}{p} = -\frac{\theta}{\varepsilon}$$

where ε denotes price elasticity of demand. Thus, if price and marginal cost of milk are known and demand elasticity can be estimated, we can back out for θ.

Past research used the conduct parameter approach in a wide range of applications. Bresnahan (1989) formalized the conduct parameter method and the New Empirical Industrial Organization (NEIO) technique where marginal cost and conduct can be simultaneously estimated. Cort(1999) criticized that the NEIO technique–where marginal cost and conduct are simultaneously estimated–only provides reliable estimates if firms play a static conjectural variation[1] game. Genesove and Mullin (1998), hereby GM (1998), and Clay and Troesken (2003) examined accuracy of the NEIO technique. They compared the values of MC_i and θ estimated by the NEIO technique with the true values. Both claimed that the differences are minimal. In contrast, Kim and Knittel (2006) and Wolfram (1999) performed similar experiments but found that the NEIO incorrectly estimate MC_i and θ. Our paper, however, is not subject to the NEIO critique because we have information on marginal cost.

We obtained the direct estimate of supermarkets' marginal cost from the Milk Development Council (MDC) and obtained the data on milk price and demand from the TNS WorldPanel survey. Our data set covers the three-year period of September 2002 to August 2005. This amounts to 156 weekly periods. We use the log-linear and Almost Ideal Demand System (AIDS) to estimate demand elasticity. The log-linear is the most popular reduced-form demand function while the AIDS is one of the most widely used structural demand function. We find that, after taking into account the endogeneity and omitted variable problems, both demand functions give similar results.

To obtain consistent estimates of demand elasticity, both log-linear and AIDS functions account for prices of 56 other grocery categories which are substitutes and complements of liquid milk. We use the principal component analysis (PCA) to reduce the number of 56 price dimensions into 11. Given as small sample size as 156, the estimation would not have been feasible otherwise. Complete discussion on the implementation of the PCA in demand estimation can be found in Hoderlein and Lewbel (2007).

This paper adds to small literature showing benefits from using the PCA in demand and conduct parameter estimation. We do so by comparing our results with that obtained from GM (1998)'s demand specification. In GM (1998), only the own-price of product is included in the demand function. They correct for endogeneity of price using the instrumental variables method. In our case, however, price of milk adjusted only twice during the sampling period and was correlated with prices of other groceries. We need to include prices of other groceries to identify the own-price coefficient (through its relation with prices of other goods), and to correct for the omitted variable bias.

Turning to the estimation results. We find that demand for milk is inelastic (about -0.4), while the conduct value is surprisingly low. The value of θ ranged between that belongs to the perfect Bertrand[2] and the Cournot

[1] A static oligopolistic competition model in which firms make strategic decisions upon their expectations of other firms' reactions. Examples of the conjectural variations model are such as Cournot, Bertrand, Stackelberg leadership and perfect collusion (monopoly).

[2] Nash equilibrium using price as the strategic variable.

competition[3]. There is not enough evidence to support successful collusion by the supermarkets. We, therefore, conclude that firms' ability to fix price mainly came from the inelasticity of demand. Had the supermarkets been perfectly collusive, price of milk would have been set at a much higher level.

The rest of the paper proceeds as follows. Section 2 describes the UK liquid milk market while section 3 introduces the data. Section 4 discusses the theoretical framework. Section 5 discusses how we implement demand estimation. Section 6 elaborates on the PCA data-reduction method which we use to solve the dimensionality, identification and omitted variable problems. Section 7, then, reports estimation results. Finally, section 8 concludes.

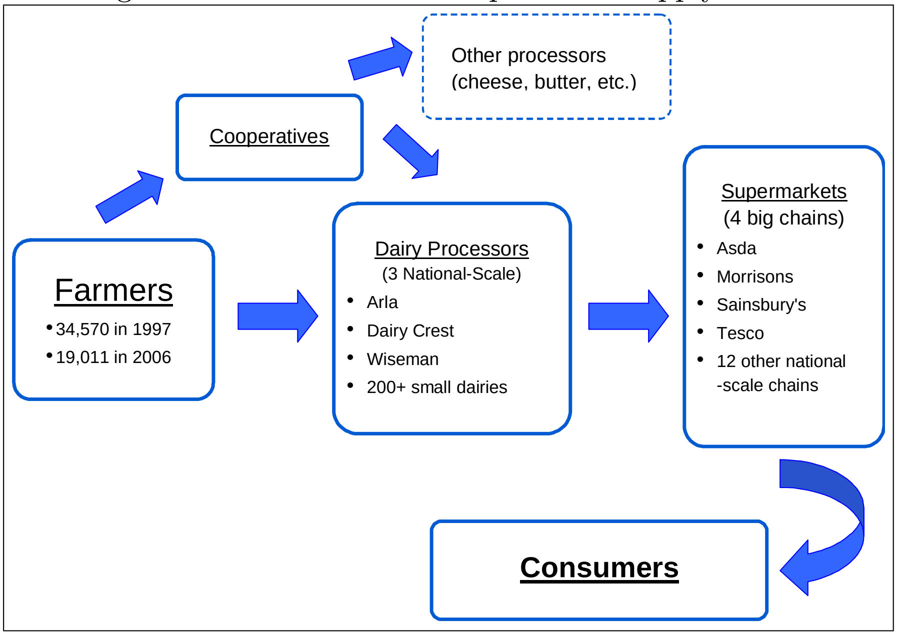
2 Market for liquid milk in the UK

The UK's liquid milk supply chain constitutes of 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, purify and bottle them for sale in the retail market. Consumers either buy milk directly from a store or have it delivered at their doorsteps by a milkman. Most milkmen are managed by dairy processors. Supermarkets make direct contract with dairy processors on price, volume and transportation arrangement of bottled liquid milk. Figure (1) depicts the structure of the liquid milk supply chain.

In 2006, there are about 19,011 dairy farms throughout the UK. This number decreased from 34,570 in 1997 (Milk Development Council, 2007a, MDC). There are over 200 liquid milk processors in the UK. Only three of which–Arla, Dairy Crest and Wiseman–operates at the natinal scale. These three biggest processors account for 90 percent of the total liquid milk sold to UK grocery retailers (Competition Commission, 2007, CC). In 2007, there are 17 supermarket chains operating nationwide. Since Morrisons acquired Safeway in 2004, only the four largest firms–Asda, Morrisons, Sainsbury's and Tesco–currently hold the majority of market share. These four supermarket firms are called "The Big4". The market share of the Big4 has been growing during the past few year. In 2005, they account for approximately 70 percent of liquid milk sold to households.

[3] Nash equilibrium using quantity as the strategic variable

Figure 1: Structure of the Liquid Milk Supply Chain



On the other hand, doorsteps milk consumption has been falling. Table (1) shows the market share trends of the Big4, other supermarkets and the doorstep milkman.

Despite seasonal fluctuations of milk demand and supply, liquid milk price has been relatively stable. During the time period analyzed in this paper (September 2002 to August 2005), price of private-label[5] liquid milk sold in the Big4 stores adjusted twice. Both times–the 30^{th} week of year 2003 and the 13^{th} week of year 2005–the adjustment was due to dairy farmers' direct action where farmers demonstrated to presure the supermarkets to raise their retail milk price. Figure (2) plots price and consumption of liquid milk in the UK.

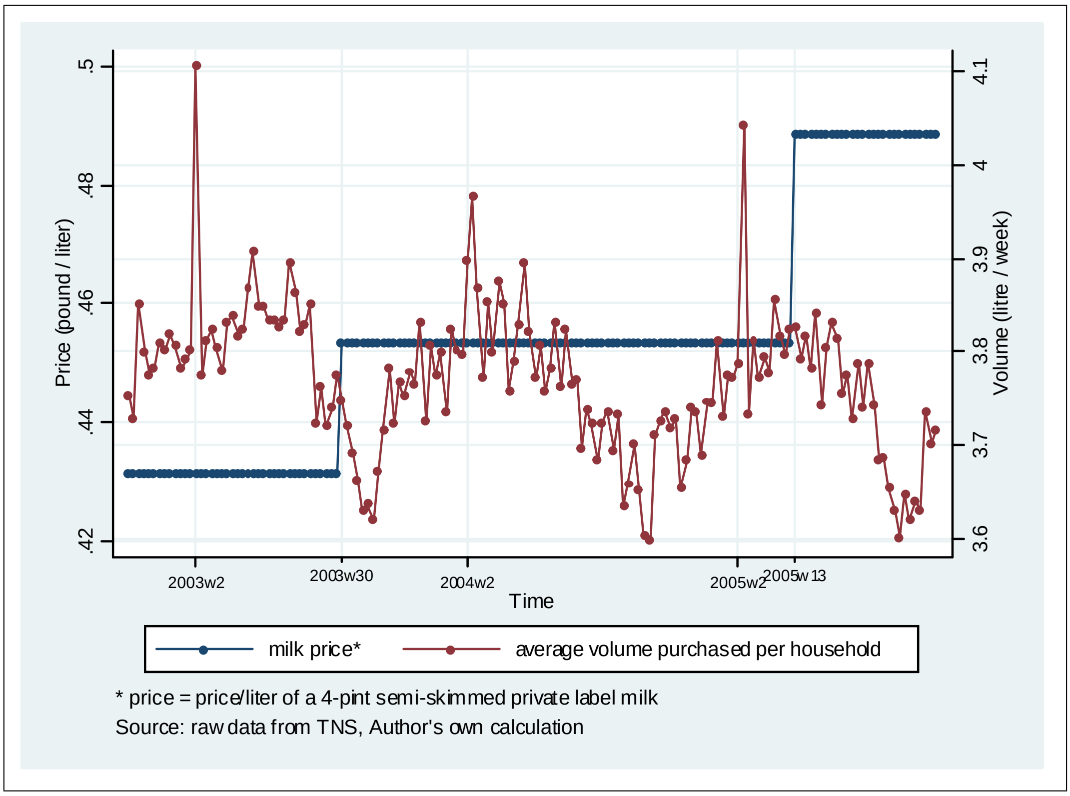
Turning to household consumption, liquid milk spending constitutes about 3.7 percent of a household's total grocery spending. This makes milk the third largest category–after vegetable (10 per cent) and fruits (5.5 per cent)–in terms of spending. Each day, about 11.3 million liter of liquid milk is consumed by households (MDC, 2007b). Liquid milk comes in different varieties according to fat content, purifying technology, brand, flavour, etc. Table (2)

[5]Store's own brand.

Table 1: Market Share of Supermarkets and the Milkman

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[4]	7.41	6.66	–	–
Sainsbury	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 $= \Sigma_{i=1}^{N} s_i^2$	0.13	0.13	0.15	0.16
Total Big4's share	63.02	64.94	67.40	68.86

Figure 2: Price and Household Consumption of Liquid Milk



summarizes proportion of different milk varieties sold in UK supermarkets in year 2005.

3 The Data

3.1 Consumer's 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 represent that of the UK population. When agreed to join the program, 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, we observe price, quantity, product category and the store identification code. As for milk

Table 2: 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 (%), National Brand(%)

* Based on sales in calendar year 2005.

Source: raw data from TNS, author's tabulation

items in particular, we were able to acquire additional information on brand, fat content, pack size, purifying technology and whether or not it was labeled as organic.

Our dataset 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.

We adopt a household selection method similar to that of Hausman and Leibtag (2004) in order to avoid information distortion from uncommitted households. The terms uncommitted households refers to those who participate for a very short duration of time or who were not serious about recording their purchases. We dropped records from any given household if they participated for less than 12 months. In addition, we dropped those who appeared to be active for less than 83.33 per cent of their participating duration. For example, if household A were participating in the homescan program for 12 months, they pass the first criterior. However, if their purchase records appeared for less than 10 months in the 12-month participating period, we drop them because the participation rate would be less than 83.33 per cent. The total of 14,480 households passed our 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 degree of collusion among liquid milk suppliers is done at the aggregate level. We evaluate the market power of the Big4 as one aggregate entity. Then, assess whether there is any evidence of collusion. To do this, we need to estimate the elasticity of aggregate demand for milk.

For the aggregate demand estimation, we sum up the total volume of milk purchased by households then divide by the number of household in each weekly period. This gives us the average per-household volume for each weekly period. As for price, we choose to use price per liter of "four-pint semi-skimmed private label milk" as a representative price of liquid milk. This is because it is the most popular type of milk consumed in the market. In 2005, private label milk accounts for about 90 percent of liquid milk sold in the Big4 supermarkets. One-fourth of that comes from four-pint semi-skimmed milk.

Price and quantity of other product categories are also updated on a weekly basis. We calculate the weighted-average (by spending) price index of each product category as follows:

$$p_{k,t} = \sum_{i=1}^{n} \left(\rho_{k_i,t} \times \frac{spending_{k_i,t}}{\Sigma_{i,t} spending_{k_i,t}} \right); k_i \in k \tag{2}$$

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

3.2 Direct Estimation of Marginal Cost

The marginal cost of liquid milk paid by the Big4 supermarkets was estimated by the MDC. From Wiseman plc's (Wiseman) financial account, MDC observed the company's turnover and volume of liquid milk sold. Since Wiseman's business is mainly devoted to liquid milk and some by-product cream, it is possible to obtain reasonable estimates of prices they received from selling liquid milk. According to Thanassoulis and Smith (2007) and MDC, Wiseman processes and delivers 70 percent of their liquid milk to the Big4 supermarkets.

The MDC estimated Wiseman's price by dividing its total revenue from liquid milk by total volume of liquid milk sold. This amount is inclusive of transportation from Wiseman to each of the Big4 outlets. Thus, we believe that it is a reasonable estimate of the Big4's marginal cost of liquid milk. However, since Wiseman's data is observed at the aggregate annual level, our direct estimation of marginal cost neither reflects seasonal effects nor accounts for the payment structure, e.g. whether the supermarkets paid two-part tariffs or linear price. Lack of accurate information on these two fronts could be problematic if we aim to study the short-run dynamic of competition among supermarkets. However, since our focus is on the average collusiveness of firms during the sample period (Sep 2002 - Aug 2005), using annual-average information would not have any significant effect on our analysis.

We take more caution, however, on our assumption of firms' symmetries. One symmetry is among the three main milk processors and another is among the Big4 supermarkets. In particular, we assume that Wiseman and the two other main milk processors–Arla and Dairy Crest–received the same price from the Big4 on average. This in turn, implies that none of the Big4 firms have higher bargaining power (against the processors) than others. So, they pay the same price for their liquid milk on average. Table(3) shows MCD's direct estimate of the annual-average price per liter paid to Wiseman. We use this as the marginal cost faced by the Big4, during our sample period.

Table 3: Annual-average price per liter paid to Wiseman plc

Time	Marginal Cost (pound / Liter)
Apr 2002 - Mar 2003	0.3342
Apr 2003 - Mar 2004	0.3512
Apr 2005 - Mar 2005	0.3555
Apr 2005 - Mar 2006	0.3665

Source: MDC, Author's tabulation.

4 Theoretical Framework

At this point, we need to find a theoretical framework which provides a foundation to evaluate the collusiveness of firms. This paper uses the conduct parameter method (CPM) to investigate the firms collusiveness. The CPM views "market power" and "collusiveness" as two separate issues. Market power refers to firms' *ability* to charge higher than marginal cost while collusiveness refers to firms' *anti-competitive behavior*. We will show later that the market power comes from three main sources which are inelastic demand, high market concentration and collusive behavior of firms. Therefore, collusiveness is just one means to gain market power. When high price-cost margin is observed, we can only infer that there should be high market power. Whether or not firms are involved in anti-competitive behavior has to be tested.

4.1 Direct measure of the conduct parameter (θ)

The conduct parameter (θ) concept was developed to represent degree of collusiveness among firms in a given oligopolistic industry. The conduct parameter was developed upon the idea that if the demand function is known and firms' marginal cost is observed, their collusiveness can be evaluated given a decision rule. The most commonly used decision rule is the conjectural variations model.

The conjectural variations model is a static oligopolistic competition model in which firms make strategic decisions upon their expectations of other firms' reactions. Examples of the conjectural variations model are such as Cournot, Bertrand, Stackelberg leadership and perfect collusion (monopoly). Given an inverse demand function $P = P(Q)$, firm $i's$ marginal cost MC_i, and assuming that firm i chooses an optimal quantity q_i such that its profit is maximized, the first-order condition gives:

$$P = MC_i - \theta Q \frac{\partial P(Q)}{\partial q_i} \tag{3}$$

This implies that when firms are perfectly collusive (monopoly case), $\theta = 1$ and when firms are perfectly competitive (perfect-competitive Bertrand), $\theta = 0$. Thus, $\theta \in [0, 1]$. In the next section, we derive θ and discuss how an outcome between the two extreme cases can be interpreted. Before we move on, however, it is worth discussing past research on this body of literature as well as the strengthes and weaknesses of the conduct parameter method (CPM).

The relation given in (3) has been adopted in a wide range of applications. In cases where researchers observe price, quantity and marginal cost, it is used to find the industry's conduct θ. GM (1998) evaluates collusiveness of firms in the US sugar industry during the period of 1890 - 1912. Kim and Knittel (2006) and Wolfram (1999) measures market power and conduct of the electricity market in Calfornia and the UK respectively. Our work falls into this category of application.

In many cases where the marginal cost is not observed, however, equation (3) can be used in two ways. First, if the cost and demand shifters are observed, it can be used to simultanously estimate the marginal cost and the conduct θ. This type of work is called the New Empirical Industrial Organization (NEIO) (Bresnahan, 1989). Second, if reserchers are willing to assume a conduct θ or specify the incentive-compatibility range of conduct θ, they can find the implied marginal cost (Berry, Levinsohn and Pakes, 1995) (Rosen, 2007).

4.2 Derivation of the Conduct Parameter (θ)

Assume that firms supply one common product and each of them cannot influence the price. If total market demand for the product is denoted by Q and price of product is denoted by p, we can write an inverse linear demand function as:

$$p(Q) = \alpha_1 + \alpha_2 Q$$

where $p(Q)$ is price of the product as a function of total quantity demanded Q.

The profit of firm i can then be expressed as a function of price-cost margin times quantity sold:

$$\pi_i = (p(Q) - MC_i) \times q_i$$

where π_i is the profit of firm i, $p(Q)$ is price of the product, MC_i is the marginal cost of firm i and q_i is quantity supply by firm i. Assuming that firms choose to produce at a quantity q_i which maximizes its profit π_i, we can obtain the profit-maximizing level of q_i from the following first order condition:

$$\begin{aligned}
\frac{\partial \pi_i}{\partial q_i} &= p + (q_i \times \frac{\partial p_i}{\partial Q} \times \frac{\partial Q}{\partial q_i} - MC_i) = 0 \\
\frac{p - MC_i}{p} &= -q_i \times \frac{\partial p_i}{\partial Q} \times \frac{\partial Q}{\partial q_i} \times \frac{1}{p} \times \frac{Q}{Q} \\
\frac{p - MC_i}{p} &= \underbrace{-\frac{\partial p_i}{\partial Q} \times \frac{Q}{p}}_{-1/\varepsilon} \times \frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q}
\end{aligned} \tag{4}$$

where ε is the price elasticity of demand for product i. The expression on the left hand side is called the Lerner's Inde

$$\frac{p - MC_i}{p} = \text{Firm } i\text{'s Lerner's index } (L_i)$$

L_i represents firm i's market power, the ability to set price above the marginal cost. As we can see from (4), market power inversely depends on price elasticity of demand (ε). Lower demand elasticity allows firms to set higher price over marginal cost, thus gives firms more market power. Given the individual firm's Lerner's index, we can find the industry's Lerner's index (L)

as a weighted average of all the firms'.

$$\begin{aligned}
L &= \sum_i s_i \times \frac{p - MC_i}{p} \\
&= \sum_i s_i \times -\frac{1}{\varepsilon} \times \frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q} \\
&= -\frac{1}{\varepsilon} \times \underbrace{\sum_i s_i \times \frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q}}_{\theta} \\
\text{Thus, } L &= -\frac{\theta}{\varepsilon} \qquad (5) \\
\text{or, } \theta &= -\varepsilon \times \left(\sum_i s_i \times \frac{p - MC_i}{p} \right) \qquad (6)
\end{aligned}$$

We define $\theta \in [0, 1]$ as the conduct parameter. The value of θ informs us about the level of collusiveness of firms in the equilibrium. The term $\frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q}$ tells us how much the quantity produced by a single firm (q_i) matters to the total market quantity. Intuitively, if firms are perfectly collusive upon quantity supply, then firms would increase and decrease the quantity supply together. An x percent increase in quantity by each firm would lead to an x percent increase in the total quantity supply Q. The term $\frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q}$ should equal to 1 in the perfectly collusive case. In contrast, if firms are perfectly competitive, then an x percent increase in quantity by firm i would not matter to the total quantity supply Q. The term $\frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q}$ should equal to 0 in the most competitive case. Now, consider the θ part in the Lerner's index.

$$\theta = \sum_i s_i \times \frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q} \qquad (7)$$

- **Bertrand or perfect competition**: in a static perfect competition equilibrium. One firm is very small compared to the entire market size. Therefore $\frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q} = 0$ which implies that $\theta = 0$. From (7), we can write:

$$\begin{aligned}
\theta &= \sum_i s_i \times 0 \\
\theta &= 0
\end{aligned}$$

- **Cournot in equilibrium**: in a static Cournot equlibrium, each individual firm decide on the profit-maximizing quantity taking into account the best decision upon quantity by other firms. In equilibrium, the quantity supplied by firm i (q_i) would equal to its market share (s_i). Therefore, a 1 percent increase in firm i's quantity would increase the market quantity by s_i percent (firm i's market share). Therefore, $\frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q} = s_i$ which implies that $\theta = \Sigma_i s_i^2$. From (7), we can write:

$$\begin{aligned} \theta &= \sum_i s_i^2 \\ \theta &= \text{the Herfindahl Index } (H) \end{aligned}$$

In a special case where firms are symmetric, $\theta = s_i$ or $\theta = \frac{1}{N}$ where N is the total number of firms.

- **Monopoly or perfect collusion**: in a static perfect collusion equilibrium, firms agree to increase or decrease quantity supply together. When we observe an x percent increase in quantity supplied by firm i, we can expect to observe an x percent increase in quantity supplyied by all other firms. Therefore, an x percent increase in q_i would lead to the same percentage increase in Q. Thus, $\frac{\partial Q}{\partial q_i} \times \frac{q_i}{Q} = 1$ which implies that $\theta = 1$. From (7), we can write:

$$\begin{aligned} \theta &= \sum_i s_i \times 1 \\ \theta &= 1 \end{aligned}$$

The value of $\theta \in [0, 1]$ where 0 implies the most competitive conduct and 1 implies the most collusive conduct.

To make explicit the relation between market concentration and market power, we can express the Lerner's index (5) interms of the Herfindahl's index (H);

$$L = -\frac{\phi H}{\varepsilon}$$

where $\phi \in [0, \frac{1}{H}]$ takes the same role as θ but is adjusted to allow the presence of H. The alternative conduct parameter takes value of $\phi = 0, H$ and $\frac{1}{H}$ in the perfect Bertrand, Cournot and perfectly collusive equilibrium respectively.

5 Demand Estimation

Previously, we discuss the assessment of market power which refers to firms' ability to set price above marginal cost. This is formally called the Lerner's index $L = \frac{p-MC}{P} = \frac{\theta}{-\varepsilon} = \frac{\phi H}{-\varepsilon}$.The formula suggests that high mark-up can be caused by three factors–low elasticity of demand (ε), high market concentration (H) and firms' collusive behaviour (θ or ϕ).

In this paper, we observe milk price, marginal cost and firms' market shares. Therefore, we are already equipped with L and H . The last element that we need to estimate before we can assess the competitive conduct θ is the elasticity of demand ε.

Before performing demand estimation, we first choose the most appropriate demand functional form given the nature of our analysis and availability of information. Then, we determine the estimation strategies adopted to obtain most consistent estimates.

5.1 Demand Functions

Past developments on demand estimations have provided us with a wide range demand functional forms. Depending on their preferences and data availability, researchers can choose from the most flexible (reduced form) to the most structured (structural estimation) demand function. There is a trade-off between estimation flexibility and theoretical consistentcy. On one hand, a flexible demand function allows us to explore the relationship among variables freely without having to impose any restriction. On the other hand, the result may not conform with consumers' rationality.

By rationality, we mean that a representative consumer chooses a consumption bundle which maximizes her utility subject to a budget constraint. The Marshallian demand function is one of those derived from a rational consumer's behavior. The function satisfies the four following principles–adding-up, homogeneity, slutsky symmetry and negative semi-definiteness. Adding-up implies that a consumer would never spend more than their wealth. Homogeneity suggests that the same percentage change in wealth and overall price level would not affect a consumer's budget allocation. Finally, Slutsky symmetry and negative semi-definiteness of the slutsky matrix hold when budget allocation is resulted from utility-maximization subject to consumers' budget constraint. Many structural demand functions, i.e. the Almost Ideal Demand System, can ensure that these properties are satisfied. Reduced-form functions, on the other hand, does not guarantee these properties.

However, limitations of the reduced form estimation is acceptable if those who use them acknowledge the limitation and employ the method in a case where limitations do not appear to impose serious problems. As stated in Deaton and Muellbauer (1980) "These restrictions [reduced form models' limitations] do not mean the model cannot be applied in practice, only that its applications must be restricted to those cases where its limitations are not thought to be serious (Deaton and Muellbauer, 1980, p.66)". What should be the most appropriate demand function depends on the nature of each analysis and the availibility of data.

In this study, we will compare results from three different demand functions–the GM (1998)'s log-linear demand specification (used as the baseline), the log-linear demand function and the Almost Ideal Demand Sytem (AIDS). The log-linear is the most popular reduced-form demand function while the AIDS is one of the most widely used structural demand function. All three of them are reasonably flexible and can be applied with continuous demand for liquid milk.

- **GM(1998)'s log-linear baseline:** in GM(1998), various demand functions were used to test the sensitivity of result to choice of demand function. They found that their results were not significantly sensitive to choice of demand function. The following general form of demand curve was used to generate a variety of reduced-form demand functions:

  $$q_i(p_i) = \beta(\alpha - p_i)^{\gamma}$$

  where q_i is quantity of product i demanded, p_i is price of product i. The parameter $\gamma = 1$ for the linear demand function, $\gamma = 2$ for the quadratic demand function, and $\gamma < 0$ and $\alpha = 0$ for the log-linear demand function. Since GM (1998) found that results were not sensitive to demand functional forms, we choose to use their log-linear function as our baseline:

  $$\log q_i = \log(-\beta) + \gamma \log(p_i) \tag{8}$$

  Notice that, to economize on data, GM(1998)'s demand specification does not include other price variables except for own-price of product (purified sugar in their case). They correct for potential endogeneity of price using the instrumental variables method.

  In our case, we suspect that using only price of liquid milk may lead to two following problems–omitted variable bias and identification of price coefficient. Our demand specifications, which we will discuss below, include prices of milk's substitutes and complements in order to alleviate these two problems.

- **Log-linear demand function**: this demand function is chosen for two reasons. First, it allows us to make direct extension to GM(1998). Second, the log-linear demand function is very flexible and has been widely used among the reduced form literature. By definition, the coefficient associated with log($price$) would be price elasticity of demand while the coefficient associated with log(expenditure) would be income elasticity of demand[6]. However, since there is no theoretical restriction imposed, the four slutsky conditions of rational demand systems cannot be guaranteed. The log-linear demand function can be written as:

$$\log q_i = \delta_1 + \delta_2 \log x + \sum_k \delta_{i,k} \log p_k \tag{9}$$

where q_i is the quantity of product i demended, x is expenditure, p_k is price of product $k \in (1, ., i, .., N)$, and $\delta_1, \delta_2, \delta_{i,k}$ are parameters to be estimated. By definition, parameter $\delta_{i,k}$ $(k = i)$ is the own-price elasticity of demand for milk and $\delta_{i,k}$ $(k \neq i)$ is the cross-price elasticity of price i on demand for product k. Therefore, homogeneity condition would be satisfied if $\Sigma_k \delta_{i,k} + \delta_1 = 0$ $(\forall k \in 1, ., i, .., N)$ (Deaton and Muellbauer, 1980).

- **AIDS**: the AID function is expressed as a budget share function. It was developed from the double logarithmic demand model where the log-linear demand funtion was modified in order to satisfy the adding-up condition. The AIDS also satisfies homogeneity, slutsky symmetry and semi-definiteness conditions. We use the same notation as in Deaton and Muellbauer (1980) to express the AID function as follows:

$$w_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log(x/P) \tag{10}$$

where w_i is the household's budget share of good i, p_j is price of good j, x is total spending, α_i, γ_{ij}, β_i are parameters and P is price index defined by

$$\log P = \alpha_0 + \sum_k \alpha_k \log p_k + \frac{1}{2} \sum_k \sum_l \gamma_{kl} \log p_k \log p_l \tag{11}$$

where α_0, α_k and γ_{kl} are parameters. In our case, where data is aggregated over households, the AIDS can be expressed as follows (Deaton and Muellbauer, 1980):

$$\bar{w}_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log(\bar{x}/kP) \tag{12}$$

where $\bar{w}_i$ is the share of aggregate spending on good i in the total spending by all households, $\bar{x}$ is the average total spending by all households and k can be viewd as the homogeneity index of households. If $k = 1$ then, we assume high equality of the budgets among households.

[6] if income is perfectly correlated with expenditure.

Thus, the term $\bar{x}/k$ can be viewed as the representative budget level (Deaton and Muellbauer, 1980). For simplicity, we set $k = 1$.

Using parameter estimates from the AIDS budget share equation (**??**), the price elasticities of demand and income elasticity can be calculated as follows:

Own-price elasticity of good i $(\varepsilon_{ii}) = \frac{\gamma_{ii} - \beta_i(\bar{w}_i - \beta_i \ln(\bar{x}/kP))}{\bar{w}_i} - 1$
Cross-price elasticity between good i and j $(\varepsilon_{ij}) = \frac{\gamma_{ij} - \beta_i(\bar{w}_j - \beta_j \ln(\bar{x}/kP))}{\bar{w}_i}$
Income elasticity of good i $(e_i) = \frac{\beta_i}{\bar{w}_i} + 1$

5.2 Estimation Issues

To obtain consistent estimates of price elasticity, we take into consideration the following issues; 1) dimensionality problem, 2) identification problem, and 3) endogeneity of price and expenditure. This section discusses how we tackle each one of them in details.

5.2.1 Dimensionality problem

The dimensionality problem usually refers to two issues, multi-colinearity of variables and too many 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 are such as 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 which individual product(s) to consume. Thus, rather than having to estimate a large demand system, researcher can only focus on products under the same decision tree's branch (or same category). This implies that products under different branches are weakly-separable. Thus, the multi-stange 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, they are separately determined as functions of product characteristics. Products that are more similar in characteristics would have higher cross-price elasticity and vice-versa.

The GCCT approach aggregates products together and find a price index of each product group. This way, researchers could reduce the number of price regressors from total number of individual products to 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 adding-up, homogeneity, slutsky symmetry and negative semi-definiteness (Lewbel, 1996). One restriction requires independence between each relative price and all the price indexes (see more in Lewbel, 1996). The test of independence is very difficult, if not impossible, to implement.

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 data reduction still maintains the four properties of a rational demand system. Hoderlein and Lewbel (2007) provides a comprehensive discussion on this. They also show how cross-price elasticities can be calculated eventhough their actual prices are not included in the PCA-modified AIDS budget share equation.

This paper uses the PCA to remedy the dimensionality problem because it is flexible while still maintains that four Slutsky conditions. Section (6.3) shows in details how we implement the PCA in our demand estimation.

5.2.2 Identification Problem

As mentioned in the previous sections, price of liquid milk adjusted twice during our observation period. Such low variation makes it difficult to identify the price coefficient. We remedy this by including prices of other products in the demand function. Identification of the coefficient associated with milk price can, then, be gained through its relation with other prices. Thus, other than solving the dimensionality problem, the PCA also solves the identification problem.

Another remedy is to use the "relative price of milk" rather than the absolute price of milk. This can be done by selecting a benchmark product that has more variation in price. Then, rescale price of milk with price of that product. However, the selection of the benchmark product or group of product could be ad-hoc.

5.2.3 Endogeneity Problem

In a case where many variables are of interest, researchers have to fix the endogeneity of each of them in order to obtain consistent estimates of associated parameters. In our case, the only two variables of interest are price of milk ($p_{milk,t}$) and average total expenditure ($\bar{x}_t$). Thus, it does not matter whether other price variables or any of the principal components are endogenous. As long as they serve as good controls for $p_{milk,t}$ and $\bar{x}_t$ then we should be fine.

Endogeneity of milk price ($p_{milk,t}$) Endogeneity of price is one common problem in demand estimation. Intuitively, while price could determine quantity demanded by the consumers, it could also *be determined* by the total quantity supplied by the producers. If this two-way relation exists and is not accounted for, price would not be exogenous in the demand equation. As a result, the coefficient associated with price would be inconsistently estimated.

In our case, however, milk price did not seem to vary with the weekly milk demand (see figure 2). All the Big4 supermarkets adjusted milk price[7] only twice during the three-year observation period. Both adjustments were due to farmers' direct action, not in respond to demand. Therefore, it is reasonable to assume that price of milk is exogenous in the demand and budget share equation (9) and (10).

Endogeneity of expenditure ($\bar{x}_t$) On the other hand, we believe that average total expenditure, which we use to proxy income, is not exogenous. While total spending (income) can determine quantity demanded, it is reasonable to assume that vice-versa is also true. For example, a household with higher (lower) budget plan on grocery is likely to spend more (less) on grocery including milk; on the other hand, a household who plans to buy more (less) milk is likely to spend more (less) on grocery. Failing to account for this endogeneity would result in inconsistent estimate of income elasticity.

To remedy the endogeneity of the average total household expenditure ($\bar{x}$), we use the estimated ($\hat{x}$) instead of actual ($\bar{x}$). Since ($\hat{x}$) is linearly correlated with ($\bar{x}$) but is not correlated with the budget share shocks by construction, the coefficient (β_{milk}) would not be subject to endogeity bias anymore. Assume that the following equation explains average household's expenditure on grocery:

$$\bar{x}_t = a + b_{milk} p_{milk,t} + \sum_{j=1}^{n} b_j \log pc_{j,t} + \epsilon_t \tag{13}$$

where $\bar{x}_t$ is the actual average household expenditure on grocery, $p_{milk,t}$ is price liquid milk at time t, $pc_{j,t}$ is the value of principal component j at time

[7] proxied by price per litre of a 4-pint semi-skimmed private label milk.

t, a, b_{milk} and b_j are parameters to be estimated. Then, we can express the estimated expenditure as:

$$\hat{x}_t = \hat{a} + \hat{b}_{milk} p_{milk,t} + \sum_{j=1}^{n} \hat{b}_j \log pc_{j,t}$$

where $\hat{x}_t$ is the estimated average household expenditure on grocery, $\hat{a}, \hat{b}_{milk}$ and $\hat{b}_j$ are estimated parameters from equation (13), $p_{milk,t}$ and $pc_{j,t}$ are the same as defined in equation (13). The next section discusses how the PCA is implemented in demand estimation.

6 PCA as a remedy to dimensionality problem

The PCA approach is new compared to all other remedies to the dimensionality problem. To the best of our knowledge, the first formal work using it is Hoderlein and Lewbel (2007). They prove that the PCA-modified AIDS still satisfies the four slutsky properties of a rational demand system. The paper shows how to estimate own- and cross- price elasticities between gasoline and various goods in the US.

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 computer network traffic system to face recognition. The main idea of the PCA is that it finds common characteristics of the multi-dimensional dataset and reconstruct each characteristic into a new principal component. Therefore, if variables in the orginal dataset are highly correlated, we would be able to use small numbers of components to account for much of the variation.

Using 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. Our construction of PCA was done in the following steps:

1. **Set-up**: supposed our original dataset contains K variables of N observations and we would like to reduce the dimension to M variables. Thus, $M < K$.

   In our case, the dataset contains prices of 57 product categories, one of which is milk. We are only interested in finding the own-price elasticity of milk. Thus, 56 other price variables are regarded as controls (so $K = 56$). We only have 156 observations ($N = 156$) and would lose one degree of freedom for each regressor added. This, we would like to use PCA to account for as much variation using smaller number of variables M (where $K > M$).

2. **Standardization of the original data**: since the PCA's scaling depends on the unit of the original data, we 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, we obtain the maximum number of eigenvectors which is K. In a case where the covariance matrix does not have full rank, we 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 dataset. The second eigenvector accounts for the second most variation of the dataset, etc. In the worst case ,where variables are completely independent from each other, each eigenvector would account for $1/K$ proportion of the variation (each accounts for equal amount 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 original data, the less number of components needed to account for a given level of variation.

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

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

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

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

This paper uses the PCA to transform a dataset of 56 price variables (56 different grocery categories excluding liquid milk) into lower dimensions of principal component. Table (4) shows the proportion of variation accounted 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, we will show that after a certain number of components are included, adding more component would not be robust to the result.

To show how different components represent different aspects of variations in the data, figure (3) plots prices of eggs, toilet rolls, bread and breakfast cereal against time; while figure (4) plots of the first four principal components against time.

6.1 Estimated Demand Specifications

Having discussed estimation strategies adopted to obtain consistent demand estimation, we now turn to the actual empirical specifications. By replacing non-milk price variables with M principal components, we can expresse the GM(1998) log-linear demand specification (8), our log-linear demand specification (9) and the AIDS budget share specification (12) as follows:

- GM(1998) log-linear baseline specification

$$\log q_{it} = \delta_1 + \delta_{milk} \log p_{milk,t} + u_t$$

  where subscript t denotes different weekly periods, q_{it} is the average consumption of milk (in liter) per household, $p_{milk,t}$ is price per liter of 4-pint semi-skimmed private label milk which is identical in all the Big4 supermarkets. The parameters δ_1 and δ_{milk} are the same as $\log(-\beta)$ and γ in equation (8) respectively. Even when conditional on $p_{milk,t}$, we cannot ensure that the error component u_t is randomly distributed. This is because there may be other variables, especially income and prices of other products that are correlated with q_{it} but not included as explanatory variables. Thus, the error component u_t would be the summation of those omitted variables and is unlikely to be normally or randomly distributed.

Table 4: PCA of Prices of 56 Grocery Product Categories

M^{th} component	Eigenvalue	% of variation	Cumulative
1	12.07	21.56	12.56
2	10.94	19.53	41.09
3	4.42	7.89	48.98
4	3.04	5.44	54.42
5	2.15	3.84	58.26
6	1.73	3.09	61.35
7	1.60	2.86	64.21
8	1.47	2.62	66.84
9	1.27	2.26	69.10
10	1.17	2.09	71.19
11	1.08	1.93	73.12
12	1.04	1.86	74.98
13	0.94	1.68	76.66
14	0.83	1.48	78.14
...	...	...	...
56	0.03	0.05	100.00

Figure 3: Examples of Tornqvist Price Indices

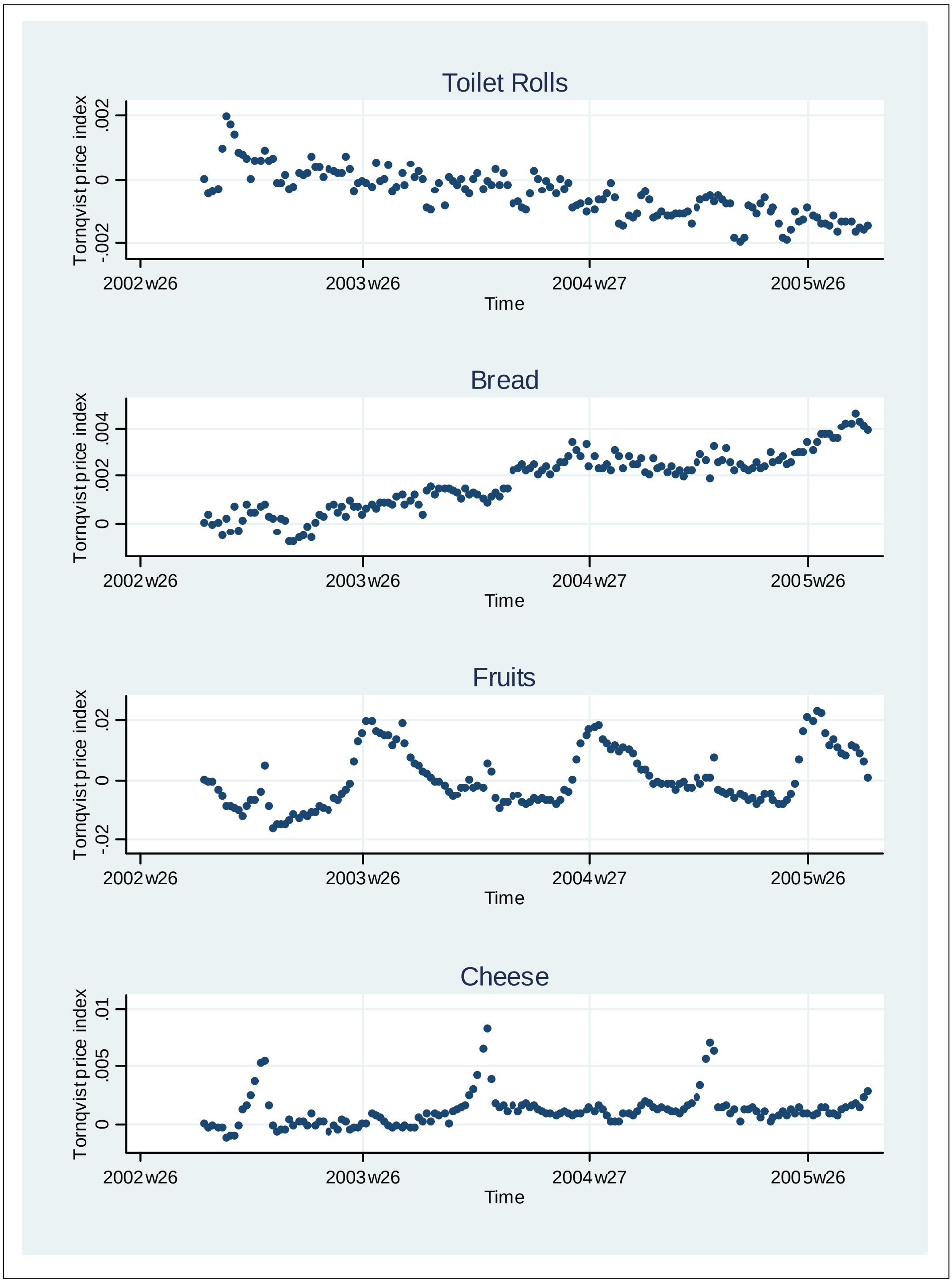
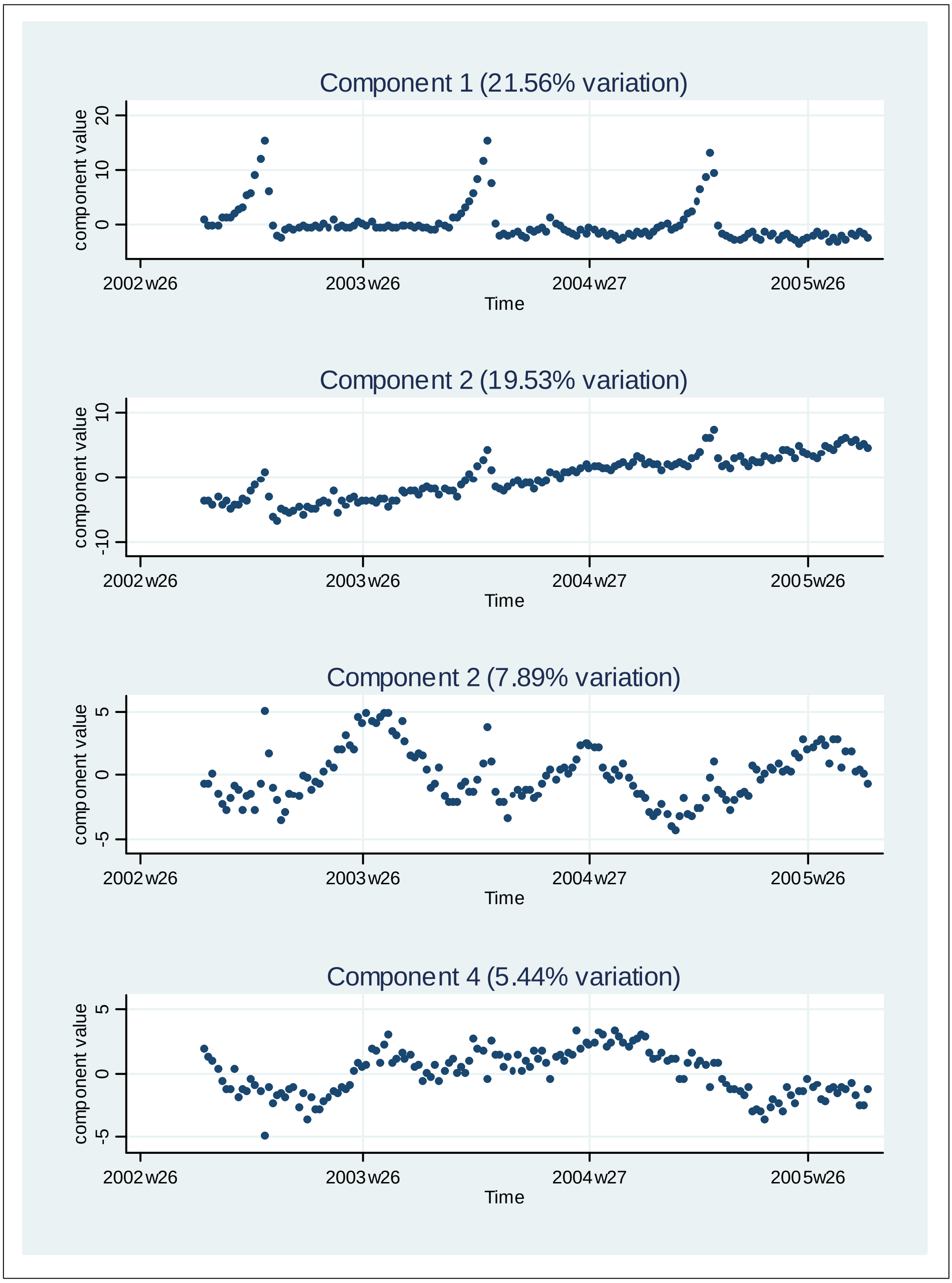


Figure 4: Plots of the first four principal components



- Log-linear demand specification

$$\log \bar{q}_{it} = \delta_1 + \delta_2 \log \hat{x}_t + \delta_{milk} \log p_{milk,t} + \sum_{m=1}^{M} \delta_m \log pc_{m,t} + \epsilon_t$$

where subscript t denotes different weekly periods, q_{it} is the average consumption of milk (in liter) per household, $p_{milk,t}$ is price per liter of a 4-pint semi-skimmed private label milk which is identical in all the Big4 supermarkets, $\hat{x}_t$ is the estimated average household's total spending on grocery, pc_m is the m^{th} principal component, M denotes total number of principal components included in the specification and ϵ_t is the error component where $E[\epsilon_t|\bar{x}_t, p_{milk,t}, pc_{m,t}] \sim N(0, \sigma^2)$. The parameters to be estimated are $\delta_1, \delta_2, \delta_{milk}$ and δ_m. In the log-log demand model, δ_2 is the income elasticity while δ_{milk} is the own-price elasticity of milk.

- AIDS budget share specification

$$\bar{w}_{milk,t} = \alpha + \gamma_{milk} \log p_{milk,t} + \sum_{m=1}^{M} \gamma_m \log pc_{m,t} + \beta_{milk} \log(\hat{x}_t / kP_t) + \varepsilon_t \tag{14}$$

where subscript t denotes different weekly periods, $\bar{w}_{milk,t}$ is the average aggregate budget share of milk in households' grocery consumption, $p_{milk,t}$ is price per liter of a 4-pint semi-skimmed private label milk which is identical in all the Big4 supermarkets, $\bar{x}_t$ is the estimated average household's total spending on grocery, pc_m is the m^{th} principal component, M denotes total number of principal components included in the specification and ϵ_t is the error component where $E[\epsilon_t|\bar{x}_t, p_{milk,t}, pc_{m,t}] \sim N(0, \sigma^2)$. The calculation of the PCA-modified translog AIDS price index $P\,_t$ is different from that of the original AIDS. The formal derivation can be found in Hoderlein and Lewbel (2007).

For practicality and calculation tractability, we approximate P_t (the AIDS translog price index defined in equation (11)) with the corrected Stone's price index. The corrected Stone's price index was introduced to correct bias caused by using the orginal Stone's price index in the AIDS estimation. Asche and Wessells (1997) discusses effects from using approximations of the AIDS price index. They suggest that the original Stone's price index would lead to inconsistent parameter estimates. The original Stone's price index can be written as:

$$\log P_t^S = \sum_{k=1}^{56} w_{k,t} \log p_{k,t} \tag{15}$$

where t subscript denotes time period, P_t^S is the Stone's price index for period t, $w_{k,t}$ is the weight given to product category k and $p_{k,t}$ is price of product category k defined in (2). In our case, P_t^S is constructed from prices of 56 different product categories in total. The weight $w_{k,t}$ is defined by:

$$w_{k,t} = \frac{spending_{k,t}}{\Sigma_{k=1}^{56} spending_{k,t}}$$

Alternatively, Moschini (1995) shows that the inconsistency can be resolved by using the "corrected Stone's price index". The new index rescales all prices in relation to the base period. We use the following price index for our estimation:

$$\log P_t^{CS} = \sum_{k=1}^{n} w_{k,t} \log \frac{p_{k,t}}{p_{k,1}}$$

where P_t^{CS} is the corrected Stone's price index and $p_{k,1}$ is price of product k in the base period. All other variables are as defined in (15). More discussion on the corrected Stone's price index can be found in (Moschini, 1995) and (Asche and Wessells, 1997).

By using the corrected Stone's price index, the calculation of price elasticities and income elasticity are also simplified to the followings:

Own-price elasticity of milk (ε_{milk}) $= (\frac{\gamma_{milk}}{\bar{w}_{milk}} - \beta_{milk} - 1)$
Cross-price elasticity between milk and good j ($\varepsilon_{milk,j}$)
$= (\frac{\gamma_{milk,j}}{\bar{w}_{milk}} - \beta_{milk}\frac{\bar{w}_j}{\bar{w}_{milk}})$
Income elasticity of milk (e_{milk}) $= \frac{\beta_{milk}}{\bar{w}_{milk}} + 1$

7 Estimation Results

This section reports parameter estimates from three demand functions–GM(1998) baseline, log-linear and AIDS. It then calculates the implied own-price elasticities of demand for milk and income elasticities of milk given those estimates. Finally, it discusses the implication of our results.

Table (5) contains parameter estimates from the baseline GM (1998) specification and our log-linear demand specification. Different specifications include different numbers of principal components. A new year dummy is included to control for seasonal spending shocks and price adjustments during the Christmas and new year holidays. Most coefficients are significant at 5 percent level. The adjusted-R^2 of the GM (1998) specification is 0.121

while the adjusted-R^2 of our log-linear specifications with at least 8 principal components are at least 0.6.

By construction of the log-linear demand function, the coefficient associated with $\log p_{milk}$ estimates the own-price elasticity of demand for milk, e.g. $\delta_{milk} = \frac{d \log \bar{q}_{milk}}{d \log p_{milk}} = \frac{d\bar{q}_{milk}}{dp_{milk}} \times \frac{p_{milk}}{\bar{q}_{milk}}$. The elasticity, in absolute term, increases as we add more principal components in the spefication. However, it becomes unrobust to higher numbers of principal components after 11 components has been included. This implies that, in our case, we need at least the first 11 components to control for price variations of products that are substitute and complements of liquid milk. By including less components, we would be subject to omitted variable bias. If the missing price components are correlated with milk price, milk price would pick up the variation of those components through the correlation. The estimated coefficient then, would not represent only the effects from milk price, but also effects from those missing variables. Our log-liner specifications give the elasticity estimates of -0.2610 to -0.3997. The baseline GM(1998) specification, which does not include any PCA, gives a much lower elasticity of -0.1781.

Similarly, by construction of the log-linear demand function, the coefficient associated with $\log \hat{x}/P$ estimates income elasticity of milk, e.g. $\delta_2 = \frac{d \log \bar{q}_{milk}}{d \log \hat{x}/P} = \frac{d\bar{q}_{milk}}{d\hat{x}/P} \times \frac{\hat{x}/P}{\bar{q}_{milk}}$. The estimates of δ_2 are greater than 1 in all log-linear demand specifications. If we assume that the estimation is accurate, liquid milk would be luxury good. This is, however, quite unlikely for UK household. We will discuss in the next paragraphes that the AIDS model gives more reasonable estimates of income elasticity.

Table (6) contains parameter estimates from the AIDS demand function. Different specifications include different numbers of principal components. Similar to the log-linear demand case, a new year dummy is included to control for seasonal budget share shocks and price adjustments during the Christmas and new year holidays. Most coefficients are significant at 5 percent level. The adjusted-R^2 of all specifications with at least 8 principal components are at least 0.68.

Unlike in log-linear case, the AIDS coefficients are not direct estimates of price and income elasticities. Table (7) reports implied price and income elasticities by the AIDS estimation. Estimates of price elasticity are close to those estimated by the log-linear model. They vary from -0.189 to -0.422. The elasticity, becomes unrobust to higher number of principal components after 11 components has been included. This finding is similar to the log-linear demand case. The result suggests that we need at least 11 principal components to control for omitted variable bias.

Table 5: Regression Results for Log-Log model of Liquid Milk Demand

Dependent variable = $\log(\bar{q}_{milk})$							
Variable (parameter)	specification						
	(GM '98)	(1)	(2)	(3)	(4)	(5)	(6)
$\log p_{milk}$ (δ_{milk})	-0.1781 (0.0378)	-0.2610 (0.1066)	-0.2644 (0.1055)	-0.3644 (0.1033)	-0.3817 (0.1037)	-0.3854 (0.1039)	-0.3997 (0.1059)
$\log \hat{x}/P$ (δ_2)	-	1.6532 (0.8559)	1.5925 (0.8472)	1.9675 (0.8082)	2.0784 (0.8092)	2.1605 (0.8162)	2.2311 (0.8232)
Principal Components							
pc_1	-	-0.0027 (0.0020)	-0.0025 (0.0020)	-0.0036 (0.0019)	-0.0038 (0.0019)	-0.0040 (0.0019)	-0.0042 (0.0019)
pc_2	-	-0.0040 (0.0016)	-0.0038 (0.0016)	-0.0038 (0.0015)	-0.0039 (0.0015)	-0.0041 (0.0015)	-0.0041 (0.0015)
pc_3	-	-0.0095 (0.0041)	-0.0092 (0.0040)	-0.0108 (0.0038)	-0.0113 (0.0038)	-0.0117 (0.0039)	-0.0120 (0.0039)
pc_4	-	-0.0186 (0.0075)	-0.0181 (0.0074)	-0.0216 (0.0071)	-0.0226 (0.0071)	-0.0233 (0.0071)	-0.0239 (0.0072)
pc_5	-	-0.0044 (0.0043)	-0.0041 (0.0042)	-0.0056 (0.0040)	-0.0061 (0.0040)	-0.0066 (0.0041)	-0.0069 (0.0041)
pc_6	-	-0.0108 (0.0059)	-0.0104 (0.0059)	-0.0128 (0.0056)	-0.0136 (0.0056)	-0.0142 (0.0056)	-0.0146 (0.0057)
pc_7	-	0.0056 (0.0010)	0.0055 (0.0010)	0.0056 (0.0010)	0.0057 (0.0010)	0.0058 (0.0010)	0.0058 (0.0010)
pc_8	-	0.0154 (0.0086)	0.0147 (0.0086)	0.0183 (0.0082)	0.0194 (0.0082)	0.0202 (0.0082)	0.0209 (0.0083)
pc_9	-	-	0.0019 (0.0009)	0.0019 (0.0009)	0.0019 (0.0009)	0.0019 (0.0009)	0.0019 (0.0009)
pc_{10}	-	-	-	0.0001 (0.0009)	0.0001 (0.0009)	0.0000 (0.0009)	0.0001 (0.0009)
pc_{11}	-	-	-	-0.0041 (0.0010)	-0.0042 (0.0010)	-0.0042 (0.0010)	-0.0042 (0.0010)
pc_{12}	-	-	-	-	-0.0014 (0.0010)	-0.0014 (0.0010)	-0.0014 (0.0010)
pc_{13}	-	-	-	-	-	-0.0008 (0.0010)	-0.0008 (0.0010)
pc_{14}	-	-	-	-	-	-	-0.0008 (0.0011)
Other variables							
New Year Dummy	-	-0.2979 (0.1764)	-0.2863 (0.1746)	-0.3668 (0.1667)	-0.3907 (0.1669)	-0.4080 (0.1684)	-0.4224 (0.1698)
constant	1.1861 (0.0299)	-5.1708 (3.3251)	-4.9429 (3.2914)	-6.4487 (3.1411)	-6.8841 (3.1451)	-7.1994 (3.1718)	-7.4796 (3.1999)
Observations	156	156	156	156	156	156	156
Adjusted-R^2	0.121	0.602	0.6107	0.6500	0.6525	0.6517	0.6505

Table 6: Regression Results for AID model of Liquid Milk Demand

Dependent variable = $\bar{w}_{milk}$

Variable (parameter)	Specification					
	(1)	(2)	(3)	(4)	(5)	(6)
$\log p_{milk}$ (γ_{milk})	0.0321 (0.0068)	0.0330 (0.0069)	0.0227 (0.0061)	0.0226 (0.0062)	0.0225 (0.0062)	0.0242 (0.0062)
$\log \hat{x}/P$ (β_{milk})	0.0167 (0.0061)	0.0154 (0.0061)	0.0017 (0.0056)	0.0017 (0.0056)	0.0020 (0.0057)	0.0014 (0.0056)
Principal Components						
pc_1 x 10^{-3}	0.0341 (0.0547)	0.0303 (0.0547)	−0.0879 (0.0470)	−0.0878 (0.0472)	−0.0875 (0.0472)	−0.0815 (0.0476)
pc_2 x 10^{-3}	−0.0423 (0.0713)	−0.0560 (0.0712)	−0.0019 (0.0634)	−0.0011 (0.0638)	0.0006 (0.0641)	−0.0168 (0.0641)
pc_3 x 10^{-3}	−0.2648 (0.0583)	−0.2646 (0.0581)	−0.2014 (0.0503)	−0.2013 (0.0505)	−0.2015 (0.0506)	−0.2058 (0.0504)
pc_4 x 10^{-3}	−0.6802 (0.0816)	−0.6650 (0.0810)	−0.5786 (0.0722)	−0.5791 (0.0725)	−0.5825 (0.0734)	−0.5697 (0.0722)
pc_5 x 10^{-3}	−0.1599 (0.0894)	−0.1578 (0.0895)	−0.0327 (0.0780)	−0.0326 (0.0783)	−0.0338 (0.0785)	−0.0381 (0.0783)
pc_6 x 10^{-3}	−0.3123 (0.0847)	−0.3121 (0.0842)	−0.2583 (0.0726)	−0.2583 (0.0729)	−0.2583 (0.0731)	−0.2626 (0.0725)
pc_7 x 10^{-3}	0.2523 (0.0885)	0.2523 (0.0881)	0.2030 (0.0764)	0.2029 (0.0767)	0.2038 (0.0770)	0.2045 (0.0763)
pc_8 x 10^{-3}	0.5679 (0.1220)	0.5577 (0.1223)	0.3667 (0.1089)	0.3668 (0.1093)	0.3707 (0.1101)	0.3666 (0.1095)
pc_9 x 10^{-3}	-	−0.2650 (0.0988)	−0.2032 (0.0859)	−0.2032 (0.0862)	−0.2047 (0.0866)	−0.2031 (0.0857)
pc_{10} x 10^{-3}	-	-	0.6699 (0.1075)	0.6701 (0.1079)	0.6748 (0.1091)	0.6604 (0.1075)
pc_{11} x 10^{-3}	-	-	0.1959 (0.0975)	0.1957 (0.0979)	0.1972 (0.0983)	0.1990 (0.0975)
pc_{12} x 10^{-3}	-	-	-	−0.0149 (0.0907)	−0.0147 (0.0909)	−0.0140 (0.0900)
pc_{13} x 10^{-3}	-	-	-	-	0.0369 (0.0985)	0.0306 (0.0974)
pc_{14} x 10^{-3}	-	-	-	-	-	0.2107 (0.1047)
Other Variables						
New Year Dummy	−0.0092 (0.0014)	−0.0089 (0.0014)	−0.0052 (0.0012)	−0.0053 (0.0012)	−0.0053 (0.0012)	−0.0052 (0.0012)
constant	0.0015 (0.0218)	0.0069 (0.0217)	0.0505 (0.0200)	0.0503 (0.0201)	0.0490 (0.0205)	0.0529 (0.0202)
Observations	156	156	156	156	156	156
Adjusted-R^2	0.683	0.6873	0.7682	0.7666	0.7652	0.7701

Table 7: Elasticities From Different Demand Models

Model / Elasticity	Specification						
	GM('98)	(1)	(2)	(3)	(4)	(5)	(6)
No. of Principal Components:	0	8	9	11	12	13	14
Log-linear model:							
price elasticity (δ_{milk})	-0.178	-0.261	-0.264	-0.364	-0.382	-0.385	-0.400
income elasticity (δ_2)	-	1.653	1.593	1.968	2.078	2.161	2.231
AIDS model*:							
price elasticity ($\frac{\gamma_{milk}}{w_{milk}} - \beta_{milk} - 1$)	-	-0.189	-0.164	-0.418	-0.419	-0.422	-0.379
income elasticity ($\frac{1}{w_{milk}}\beta_{milk} + 1$)	-	1.429	1.398	1.043	1.044	1.053	1.035

Notes: based on calculation using sample mean $w_{milk} = 0.039$.

Table 8: Estimates of the Conduct Parameter

Timing	Mean Price	Mean MC	Lerner's Index	Conduct Parameter (θ)		
	(pound/liter)		(L)	GM('98)	Log-Linear	AIDS
Before 1^{st} direct action:						
	0.4313	0.3396	0.2128	0.0379	0.0812	0.0891
Between 1^{st} and 2^{nd} direct action:						
	0.4533	0.3536	0.2200	0.0392	0.0840	0.0922
After 2^{nd} direct action:						
	0.4886	0.3650	0.2530	0.045	0.0966	0.1060

Notes: Lerner's index $L = \frac{\bar{p} - \overline{MC}}{\bar{p}}$; $\bar{p}$ is mean price and $\overline{MC}$ is mean marginal cost. The conduct parameter $\theta = -\varepsilon \times L$; $\varepsilon = -0.1781, -0.382$ and -0.419 in the GM(1998)'s log-linear baseline, our log-linear specification and AIDS demand specifications respectively.

Turning to income elasticity, AIDS's estimates of income elasticity ranges from, 1.429 to 1.035 but stays around 1.043 to 1.035 after 11 components has been included. Thus, unlike the log-linear demand case, AIDS suggests that liquid milk is a normal good.

Table (8) reports our estimates of the conduct parameter under different circumstances. The conduct estimates from our log-linear and AIDS specifications are similar; while those obtained from the GM (1998) baseline specification gives much lower values. Price of liquid milk was adjusted twice during our 3-year sample period, both were due to farmers' direct actions. Thus, a separate conduct was estimated for each of the following events– before 1^{st} direct action, between 1^{st} and 2^{nd} direct action and after 2^{nd} direct action. As for the consumers' price elasticity of demand, we pick the ones from 12-component specifications (specification 4) and assume that it stays constant throughout the sample period.

We discussed in section(4.2) that possible values of the conduct parameter lie between 0 and1 where $\theta = 0$ represents the most competitive conduct (perfect Bertrand) and $\theta = 1$ represents the most collusive conduct (monopoly). If firms play a static Cournot game in equilibrium, then $\theta = \Sigma_{i=1}^{N} s_i^2$ (the Herfindahl Index). In our case, $\Sigma_{i=1}^{N} s_i^2$ equals 0.13, 0.13, 0.15 and 0.16 in 2002, 2003, 2004 and 2005 respectively. These values are higher than all of our estimated conduct parameters. Thus, overall, the conduct still lied between perfect Bertrand[8] and Cournot competition[9]. As the conduct parameters are significantly lower than 1, we do not have sufficient evidence to support perfect collusion.

8 Discussion and Conclusion

This paper studies degree of collusiveness among major liquid milk retailers in the UK–Asda, Morrisons, Sainsbury's and Tesco. We adopt the conduct parameter approach which suggests that firms' ability to charge higher than cost (market power) comes from three sources–inelastic demand, high market concentration and firms collusiveness. Using our data on price, marginal cost and market share of each firm, we estimate demand elasticity and firms' competitive conduct.

The demand elasticity was estimated using log-linear and AIDS demand functions. All our demand specifications use the PCA to account for prices of substitutes for and complements of milk. The PCA method was proved very beneficial for estimation of demand systems with many goods. We were

[8] Nash equilibrium using price as the strategic variable.

[9] Nash equilibrium using quantity as the strategic variable.

able to reduce the number of price variables from 56 to 11. With as small number of observations as 156 in our case, the reduction is significant.

To illustrate the benefit of the PCA, we compare our results with that obtained from GM(1998)'s log-linear demand specification. The GM(1998)'s demand specification uses only milk price as the explanatory variable. Although milk price is exogenous in both specifications, the GM(1998) baseline specification consistently underestimates demand elasticity and conduct parameter. This is because of two reasons. First, there is not enough variation in milk price to identify the coefficient; and second, milk price is correlated with prices of other products which are not included in the GM (1998) baseline specification. In our demand specifications, the principal components of prices help identify the coefficient associated with milk price and help controlling for omitted variable bias.

To find the minimum price variation and control variables needed to obtain consistent estimates, we compare results from specifications with 8 to 14 principal components. After including 11 components, parameter estimates become unrobust to additional variation. Altogether, 11 components account for 73.12 percent of variation in all the 56 category prices.

Own-price elasticity of milk from the log-linear and AIDS demand models are -0.382 and -0.419 respectively. This implies that demand for liquid milk faced by the Big4 is inelastic–one percent increase in price would reduce demand by about 0.4 percent.

Turning to our estimates of the conduct parameter (θ). If we assume that consumers' elasticity remained constant during our observation period, the value of θ would increase every time there was a farmers' direct action. This is because farmers' direct actions facilitated the supermarkets to collectively raise price. Such collective action may have not been successful otherwise.

Despite our findings that the value of θ have been incrasing through time, the overall results still suggest that firms are far from perfectly collusive. The value of θ ranged from 0.0812 and 0.0966 in the log-linear model and from 0.0891 to 0.1061 in the AIDS model. Comparing these values with the perfect competition's conduct (or perfect Bertrand $\theta = 0$), the Cournot's conduct ($\theta = \Sigma_{i=1}^{N} s_i^2 = 0.13$ to 0.16) and perfect collusion (or monopoly $\theta = 1$), the estimated θ fall between perfect Bertrand and Cournot.

Due to such inelastic demand ($\varepsilon \approx -0.4$), it is possible for a perfectly collusive group of firms to set much higher price than the observed level. The monopoly price level, however, cannot be estimated with precision here because we do not know the elasticity of demand at the monopoly price and quantity level.

We conclude that although the Big4 were involved in price-fixing behavior, the observed price levels were still much lower than the perfectly collusive scenario. The inelasticity of demand enabled the Big4 to raise price above marginal cost without losing much sales. Price-fixing behavior results in large welfare transfer from consumers to firms. The total dead weight loss due to reduction in demand, however, should be relatively small.

References

[1] 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.

[2] Berry, S., J. Levinsohn, et al. (1995). "Automobile Prices in Market Equilibrium." Econometrica 60(4): 889-917.

[3] Bresnahan, T. F. (1989). Empirical Studies of Industries with Market Power. Handbook of Industrial Organization. R. Schmalensee and R. D. Willig, Elsevier Science Publisher B.V. 2.

[4] Clay, K. and W. Troesken (2003). "Further tests of static oligopoly models: whiskey 1882-1898." The Journal of Industrial Economics LI(2): 151-166.

[5] Commission Competition (CC). (2007). "Milk Supply Chain Profitability." Groceries Market Inquiry Retrieved 09 May, 2007, from http://www.competition-commission.org.uk/inquiries/ref2006/grocery/prov_findings/ prov_find_app_8_3.pdf.

[6] Corts, K. S. (1999). "Conduct parameters and the measurement of market power." Journal of Econometrics 88: 227Ð250.

[7] Deaton, A. and J. Muellbauer (1980). Economics and Consumer Behavior. Cambridge, Cambridge University Press.

[8] Genesove, D. and W. P. Mullin (1998). "Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914." RAND Journal of Economics 29(2): 355-377.

[9] Gorman, W. M. (1959). "Separable utility and aggregation." Econometrica 27: 469-81.

[10] Hausman, J. A. and E. Leibtag (2005). Consumer Benefits from Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart.

[11] Hoderlein, S. G. N. and A. Lewbel (2007). Price Dimension Reduction in Demand Systems With Many Goods, Mannheim University and Boston College.

[12] 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.

[13] Lewbel, A. (1996). "Aggregation Without Separability: A Generalized Composite Commodity Theorem." The American Economic Review 86(3): 524-543.

[14] Milk Development Council (MDC). (2007b). "Liquid Milk Expenditure and Volume." Retrieved 23 December, 2007, from http://www.mdcdatum.org.uk/RetailerDataPrices/tnsliquidmilk.html.

[15] Milk Development Council (MDC). (2007a). "Producer Numbers." Retrieved 04 December, 2007, from http://www.mdcdatum.org.uk/FarmDataPrices/producernumbers.html.

[16] Moschini, G. (1995). "Units of Measurement and the Stone Price Index in Demand System Estimation." American Journal of Agricultural Economics 77(1): 63-68.

[17] Office of Fair Trading (OFT). (2007). "OFT welcomes early resolution agreements and agrees over £116m penalties." OFT Press Releases 2007 Retrieved 05 January, 2007, from http://www.oft.gov.uk/news/press/2007/170-07.

[18] Pinkse, J. and E. M. Slade (2004). "Mergers, brand competition, and the price of a pint." European Economic Review 48: 617-643.

[19] Rosen, A. (2007). Identifcation and Estimation of Firms' Marginal Cost Functions With Incomplete Knowledge of Strategic Behavior. Cemmap Working Paper. London, Institute of Fiscal Studies, Department of Economics, UCL.

[20] Smith, H. and J. Thanassoulis (2007). Estimating Bargaining Power In The Supply Chain. Oxford, Departmet of Economics, Oxford University.

[21] Wolfram, C. D. (1999). "Measuring Duopoly Power in the British Electricity Spot Market." The American Economic Review 89(4): 805-826.