

Food Marketing Policy Center

Evaluating Coordinated Effects: An application to the US beer industry

by Frederik Beelitz

Food Marketing Policy Center
Research Report No. 114
July 2009

Research Report Series

<http://www.fmpc.uconn.edu>



University of Connecticut
Department of Agricultural and Resource Economics

Evaluating Coordinated Effects: An application to the US beer industry

Frederik Beelitz

St. Antony's College, University of Oxford

Thesis submitted in partial fulfillment of the requirements for
the degree of Master of Philosophy

May 13, 2009

Abstract

The paper brings Friedman's (1971) collusive game to data and investigates whether the merger between the fifth and fourth largest brewer (G. Heileman and Stroh) of the US beer industry in the mid 1990's had a significant impact on the incentives to collude in the industry. It does so by firstly estimating a random coefficient Logit demand system for the US beer market. In a second step the demand estimates are used to conduct a merger simulation (Davis, 2006) quantifying coordinated effects of the merger. The results show that the change in the likelihood of collusion for the non merging parties was negligible, but significantly increased for the merged party.

Contents

1	Introduction	3
2	The US beer industry	6
3	Collusion model	9
3.1	Single period payoffs	9
3.2	Incentive compatibility constraint	10
3.3	Calculating marginal costs	12
3.4	Estimating critical discount factors	13
4	The econometric demand model	16
4.1	Previous demand estimations	16
4.2	Demand estimation used in this paper	19
4.2.1	Logit model	22
4.2.2	Random coefficient Logit model - The full model	24
5	Endogeneity and instruments	27
6	Estimation	29
7	The data	34
8	Results	38
8.1	Demand results	38
8.1.1	Logit model results	38
8.1.2	Full model results	42
8.1.3	Full model elasticities	45
8.2	Collusion game results	50
8.2.1	Presentation of merger simulation results	50
8.2.2	Interpreting the results	55
9	Conclusion	58

Acknowledgement *I am very grateful to Dr. Ronald Cotterill of the Food Marketing Policy Center at the University of Connecticut for kindly making the extensive and unique data available, to my supervisor Howard Smith for all his support and to Christian Rojas for sharing further data with me.*

Word Count *The computerized word count is 15,986 words. Taking page 54 as representative page, I count 12 words on average for 32 lines, multiplying by 73 pages results in a word count of 28,032 words.*

1 Introduction

When assessing the competitive effects of a merger, antitrust authorities will generally inspect whether the merger leads to unilateral effects or coordinated effects. The former arises when the merged entity would enjoy an excessively dominant position in the market. Coordinated effects exist if the merger changes the nature of competition making firms significantly more likely to collude and raise prices. This paper empirically evaluates the change in the likelihood of collusion following the merger of the fourth (Stroh) and fifth (G Heileman) largest US brewer in 1996. The US beer industry with its high concentration ratio, high degree of non price competition and frequent public statements made by firms hinting at coordination (Tremblay, 2005) seems to be an industry particularly prone to coordinated actions. Thus far however no merger in the US beer industry was blocked on the grounds of coordinated effects. The aim of this analysis is to assess whether coordinated effect concerns should have been raised by the Federal Trade Commission (FTC) regarding the merger between Stroh and G. Heileman. I follow a two step procedure and firstly estimate a random coefficient demand system based on Berry, Levinsohn and Pakes (1995, hereafter BLP) and Nevo (2000). The estimates of the demand system combined with pricing rules under different modes of competition are then used to determine the likelihood of collusion pre and post merger. The research therefore builds on both, the empirical demand estimation literature and the empirical coordinated effects literature.

Competition authorities base their decisions on coordinated effects mainly on qualitative measures and follow a rich theoretical framework. This is due to the fact that the theory of coordinated effects is well studied (Abreu (1986), Compte, Jenny, Rey (2002), Kuhn (2004), Sabbatini (2004)), while empirical applications are less frequent. Most of the existing empirical studies have focused on evaluating whether observed prices are more consistent with collusive or competitive price levels (Bresnahan (1982,1987), Nevo (1998,2001), Pinkse and Slade (2004)). Only very little empirical work has been done so far that analyses whether collusion is sustainable as an equilibrium (Kovacic et al, 2007). This paper contributes to this part of the literature by applying a methodology that brings an actual model of tacit collusion to data. It is the first (in conjunction with Davis and Huse, forthcoming) to empirically simulate Friedman's (1971) collusive framework. The framework consists of an infinitely repeated

multiproduct Bertrand game where firms follow grim-trigger strategies. The advantage of using a basic framework as opposed to more complex collusion set-ups (e.g. optimal punishment mechanisms (Abreu,1988)), is that the results will capture the essence of repeated interactions while keeping the analysis fairly tractable.

The methodology relies strongly on a precise and realistic demand system. The demand estimates and the subsequent cross price derivatives are key in the simulation of both, prices and demand under different modes of competition. Traditional demand estimations such as the Almost Ideal Demand System (Deaton and Muellbauer, 1980) or multi-stage budgeting approach (Hausman et al, 1994) fail to incorporate consumer heterogeneity. The random coefficient demand estimation (BLP, Nevo (2000)) is accepted as the most flexible, reliable and realistic model to estimate demand and is mainly criticized for its computational difficulty. It is based on the discrete choice literature (Lancaster (1971), Mcfadden (1974)) but allows for heterogeneous coefficients on price and product characteristics. I mainly follow Nevo (2000) for my demand estimation by modelling the heterogeneity of coefficients on the basis of true demographic distributions. I explicitly model for unobservable product characteristics and correct for endogeneity by constructing instrumental variables from the panel structure of the data. The data used in this analysis is quarterly supermarket data covering 58 cities and 64 brands over five years (1988 to 1992). This provides an extensive data set which is representative for US beer consumption, since 50% of beers sold in the US are retailed via supermarkets and grocery stores. This paper is the first to estimate a random coefficient demand system of beer in the US. It therefore contributes to the empirical demand literature by highlighting differences to other types of demand estimations of the US beer market (Hausman et al (1994) and Rojas (2008)). The results of my demand estimation suggest that the own price elasticities of the random coefficient model are on average slightly higher (-6.3) than previous studies (-5 in Hausman et al (1994) and -3.5 in Rojas (2008)). The cross price elasticities are of similar dimension. The results of the subsequent merger simulation provide two key insights that stand in stark contrast with the results associated with Friedman's theoretical model. Firstly, I find that smaller firms were more prone to collusion than larger firms. Economic theory would suggest that smaller firms gain more from defecting and therefore are less tempted to collude. Secondly, the collusive effects of the merger on non merging firms were negligible but significantly increased for the merging parties. A fall in concentration is traditionally assumed

to have an identical effect on all players.

The structure of the paper is as follows. Section 2 presents the US beer market. Section 3 describes the theoretical model of the collusion game. Section 4 outlines the econometric model underlying the demand estimation. In Section 5, I explain how the problem of endogeneity of prices is addressed. Section 6 outlines the theoretical estimation procedure as conducted in Matlab. In section 7, I give a brief overview of the data. Section 8 presents the results of the demand model and the merger simulation before section 9 concludes.

2 The US beer industry

Brewing as it is known today in the US emerged in the mid nineteenth century when German immigrants began making German-style Lager instead of English style Ales. American consumers quickly took to the lighter Lagers produced by Anheuser-Busch, Coors, Miller, Pabst and Schlitz and by 1918 1,568 brewers were selling their products in the US. The prohibition laws from 1919 illegalized the production or sale of beer and put all brewers out of the beer making business. After the prohibition laws were abolished in 1933, breweries quickly returned to brewing, bringing the number of competitors to 700 in 1937. The concentration of the industry has increased drastically ever since. From 1947 to 2001, the number of brewers of Lager fell from 421 to 24 and the market share of the four largest producers rose from 17 to 94 percent (Tremblay, 2005). Over that period, overall consumption increased from 48 million to 202 million barrels of beer in 1994.

Table 1 shows the development of the market shares of the four largest firms over the period 1970 to 1995. In the mid 1990's the industry was highly concentrated with Anheuser Busch, Coors and Miller producing about 80 % of domestic beer consumption. Interestingly, a higher concentration led to a higher number of products supplied and in 2001, 2800 brands were on the market with each brewer selling an average of 35 brands. However, the majority of these brands covered niche demands. Particularly the introduction of Light beers in the late 1970's to mid 1980's accelerated the trend of brand proliferation and customer segmentation across the industry. Anheuser Busch clearly was (and still is) the market leader and its two leading brands Budweiser and Bud Light captured approximately one third of beers nationwide in 1995.

Table 1
Domestic market shares (%)¹

Brewer	1970	1975	1980	1985	1990	1995
Anheuser-Busch	17.8	23.4	28.4	38.2	45.8	47.6
Coors	5.8	7.9	7.8	8.3	10.2	11
Miller	4.1	8.5	21.1	20.8	22.9	23.2
Stroh	2.7	3.4	3.5	13	8.6	5.8

The market is dominated by one type of beer: American style Lager. It

¹Source: Tremblay (2005)

constitutes over 90% of domestic beer with the remaining 10% split between Ales, Stouts, imports and specialty brews. As a result of the homogeneous supply mix, the difference in taste of competing brands is not as accentuated as it might be in markets with a higher variety of products. In fact, most US beer consumers find it difficult to choose their favorite beers when beer labels are covered (Scherer, 1996) . Consequently, advertising and promotional activities have become one of the key differentiating factors. In 2000 the top three brewers together spent about \$ 744 million on advertising or 8.6% of total sales. This makes the brewing industry one of the top 'advertising spenders' across all 'consumables' industries (Tremblay, 2005). However, it is only the largest firms in the industry that are able to finance these significant expenditures. In 1995, the three largest brewers outspent their competitors by a ratio of 5 to 1.

Beer in the US is sold via four different channels: on premise (accounting for 26% of sales in 1995), liquor stores (17%), supermarkets (19%) and convenience stores (20%). Supermarkets and convenience stores accounted for 40 % of sales in 1995. The figure has been steadily increasing with convenience stores now being the number one retail channel for beer ².

The strong industry concentration today is a result of aggressive merger and acquisition strategies throughout the latter half of the century with about 200 acquisitions recorded over the last 60 years. This was the most influential factor in altering the competitive landscape of the US brewing industry and raised concerns of non competitive behavior. The reasons for the high number of acquisitions include access to a wider distribution network, achieving economies of scale or simply reducing competition. The high number of inefficiently producing firms resulted in a rather lenient antitrust policy toward US brewers throughout the 1960's and 1970's. Antitrust policy became more stringent for large producers and more lax for smaller brewers in the 1980's in an effort to bridge the gap between larger and smaller producers. G.Heileman benefitted from this and followed a particularly aggressive merger strategy acquiring 17 companies over the period 1961 to 1987 making it the fifth largest brewery by 1994. After experiencing financial difficulty however, Heileman was acquired by Stroh in 1996. The merger was not challenged by the FTC and strengthened Stroh's position as fourth largest producer. It further isolated the top four brewers from the smaller ones. Their respective market shares in 1996 were as follows: Anheuser Busch (48.9%), Miller (23.1%), Coors (11.2%) and Stroh

²Source: SAB Miller/Rave industry analysis 2003

(8.3%).

No merger in the US beer industry was ever challenged for coordinated effects. This is surprising given that the US beer industry seems particularly prone to collusive behavior for numerous reasons. Firstly, a high degree of concentration and the domination of a few firms might facilitate an agreement. Secondly, the low level of product differentiation and similar cost structures would have made coordination easier (Tremblay, 2005). Thirdly, the retail prices of beers were readily available and this would have facilitated the detection of a potential defector from the collusive agreement.

The developments in the US beer industry throughout the 1980's and its competitive landscape in the mid 1990's indicate that it was particularly vulnerable to collusive behavior. It is therefore well worth investigating the collusive effects of the merger between Stroh and G. Heileman.

3 Collusion model

In this section, I define the framework used to determine the critical discount factors pre and post merger. Comparing both these discount factors provides an indication for the change in the likelihood of a collusive agreement when market structures are altered by a merger.

3.1 Single period payoffs

The one period stage game used here is identical to the one used in most unilateral effects merger simulation literature (Werden and Froeb (1994), Hausman et al (1994), BLP and Nevo(2000)). In particular I apply a differentiated multi-product Bertrand pricing game to determine the single period payoffs of the infinitely repeated game.

Suppose there are J different brands in the market of interest and each firm f of all F active firms produces a subset Λ_f of all J products. Assuming no fixed costs profits for firm f are

$$\pi_f = \sum_{j \in \Lambda} (p_j - c_j) M s_j(p), \quad (3.1)$$

where c_j is the marginal cost of production for product j and is assumed not to change post merger, p_j are prices of product j , $s_j(p)$ is the market share of product j which is a function of prices of all products in the market and M is the overall potential market size. Note that the overall market size used here is substantially different to traditional antitrust market definition analyses. Here M is defined as the total market size including the share of the outside good (see section 7 for details on the estimation of M). A great advantage of using this type of 'market sizing' is that one can keep the overall market size fixed, while allowing for a flexible quantity of products sold. This makes the final results of the analysis less sensitive to the definition of the market.

Any profit maximizing firm f will choose a price-vector of its products p_f such that it maximizes equation 3.1 subject to prices being positive and goods being sold in positive quantities. Assuming that there exists a pure strategy Nash equilibrium and the constraints do not bind, the profit maximizing prices of any product j sold by firm f must satisfy the following first order condition

$$s_j(p) + \sum_{k \in \Lambda_f} (p_k - c_k) \frac{\partial s_k(p)}{\partial p_j} = 0, \quad \text{for all } j \in J \quad (3.2)$$

This results in a system of J equations, which is solved by the Nash equilibrium set of prices denoted by the vector $p^{NE} = (p_1^{NE}, p_2^{NE}, \dots, p_J^{NE})$. Substituting p^{NE} and the resulting demand levels ($s(p^{NE})$) into equation 3.1 yields the one period competitive payoff $\pi_f^{NE} = \pi_f(p^{NE})$.

Similarly there exists a collusive solution, $p^{cl} = (p_1^{cl}, p_2^{cl}, \dots, p_J^{cl})$, to the set of J equations

$$s_j(p) + \sum_{k=1}^J (p_k - c_k) \frac{\partial s_k(p)}{\partial p_j} = 0, \quad (3.3)$$

where all firms jointly maximize their profits and act as a cartel³. Denote the one period per firm collusive profits as $\pi_f^{cl} = \pi_f(p^{cl})$.

Any firm might find it profitable to deviate from the collusive payoff and might want to undercut its rival's high collusive prices to gain a larger market share. The firm in question will therefore want to set its prices $p_f^{def} = (p_j^{def}, \dots, p_\Lambda^{def})$ for all $j \in \Lambda_f$ such that it maximizes the following profit equation

$$\sum_{j \in \Lambda} (p_j - c_j) M s_j(p_f^{def}, p_{-f}^{cl}). \quad (3.4)$$

Denote defection payoffs as $\pi_f^{def} = \pi_f(p_f^{def}, p_{-f}^{cl})$, where p_{-f}^{cl} is the vector of collusive prices of all products not sold by the deviating firm f .

3.2 Incentive compatibility constraint

Following Friedman (1971), I will consider the feasibility of sustaining a collusive equilibrium in an infinite repetition of the above stage game when firms adopt a 'grim trigger' strategy. In a 'grim trigger' strategy setting a defector receives his defection payoffs in one period and subsequently earns Nash equilibrium profits in all following periods: a deviation triggers an end to cooperation. A firm is assumed to maximize the net present value (NPV) of its profits. It will make optimal decisions at each point in time (node of the game tree) and

³Note that the summation in the collusive outcome is over all products and not only over the products sold by an individual firm.

will therefore behave in such a way as to satisfy the condition for a subgame perfect equilibrium. Friedman (1971) shows that a large number of subgame perfect equilibria are possible if all players follow a 'grim trigger' strategy and are sufficiently patient.

If the above stage game is infinitely repeated the anticipated NPV from colluding is simply the infinite sum of the collusive payoffs discounted at a firm specific rate δ_f , which can be written as $V_f^{cll} = \frac{\pi_f^{cll}}{1-\delta_f}$. The NPV from defection for firm f will be $V_f^{def} = \pi_f^{def} + \frac{\delta_f \pi_f^{NE}}{1-\delta_f}$, i.e f receives defection profits in the first period, followed by Nash equilibrium profits in all subsequent periods. A firm will only find it profitable to collude when its anticipated NPV from colluding is higher than that of defection, resulting in the well known incentive compatibility constraint

$$\frac{\pi_f^{cll}}{1-\delta_f} \geq \pi_f^{def} + \frac{\delta_f \pi_f^{NE}}{1-\delta_f}, \quad (3.5)$$

$$\text{or } \delta_f \geq \frac{\pi_f^{def} - \pi_f^{cll}}{\pi_f^{def} - \pi_f^{NE}} \quad (3.6)$$

Inequality 3.6 gives a condition under which a collusive agreement is stable. It defines a minimum critical discount value above which collusion is not only possible but also profitable for firm f in an infinitely repeated game. The firm specific critical discount factor (δ_f) therefore acts as a measure of likelihood of collusion. If the true discount factor ($\widehat{\delta}_f$) of a firm f lies above the critical discount factor δ_f , the firm will find it profitable to engage in a collusive agreement. Therefore, the higher δ_f the smaller is the probability that $\widehat{\delta}_f$ will lie above δ_f . In other words, a higher δ_f implies a smaller likelihood of collusion.

The aim of the paper is to estimate these firm specific discount values pre and post merger. The difference between the set of critical discount values will give an indication as to how much more (or less) likely collusion is under a new market structure. The crucial determinant needed to calculate the different profit levels in equation 3.6 is a precise demand estimation ($s_j(p)$). The demand estimation used in this paper is a random coefficient Logit model and the details and results of this estimation will follow in subsequent sections. The rest of section 3 however outlines how each of the one period payoffs in inequality 3.6 are computed to determine critical discount values.

3.3 Calculating marginal costs

After estimating a reliable demand system, the first step in the merger simulation consists of calculating marginal cost to determine the price cost margin of each product. The available data set does not include marginal costs and this information is (understandably) not publicly available. I therefore use a widely used application in empirical IO and retrieve the marginal costs from the first order conditions of the profit maximization problem (equations 3.2).

In line with most of the merger simulation literature (Nevo (2001), Davis (2006)) and to facilitate notation I introduce a $J \times J$ ownership matrix $\mathbf{\Delta}$. The elements $\Delta_{j,k}$ (jth row and kth column) take the value 1 if product j and k are sold by the same firm and zero otherwise. I also introduce the superscripts *pre* and *post* on selected variables, indicating that the elements of the matrix or vector are pre or post merger variables respectively. Denote \mathbf{B} as the $J \times J$ matrix of own and cross price derivatives where element $b_{j,k} = \frac{\partial s_j(p)}{\partial p_k}$ indicates the change in demand for product j resulting from a unit price change of product k . The column vectors $s(p)$ and c are the vectors of market shares (as a function of all prices) and marginal costs respectively. The J system of equation 3.2 can therefore be rewritten in matrix notation as

$$s(p^{pre,NE}) + (p^{pre,NE} - c)\mathbf{\Delta}^{pre} \circ \mathbf{B}' = 0 \quad (3.7)$$

Let \circ be the element by element or 'Hadamard' product multiplier. Rewriting equation 3.7 to solve for marginal costs yields

$$c = (\mathbf{\Delta}^{pre} \circ \mathbf{B}')^{-1} s(p^{pre,NE}) - p^{pre,NE} \quad (3.8)$$

To estimate marginal costs in equation 3.8 the competitive price vector ($p^{pre,NE}$) and the competitive demand levels ($s(p^{pre,NE})$) need to be known. These two elements however can only be established via equation 3.2 where in turn c needs to be known. To circumvent this problem, I assume that the observed prices and demand levels over the last period of my data set are at their competitive outcomes and the firms compete in prices. Hence $p^{pre,NE}$ and $s(p^{pre,NE})$ can be directly observed from the data. Note that I do not differentiate between *pre* and *post* merger marginal cost levels, implicitly assuming that marginal costs are not affected by the merger.

3.4 Estimating critical discount factors

To estimate the firm specific critical discount factors δ_f^{pre} and δ_f^{post} as defined by inequality 3.6 the following payoffs need to be calculated $\pi_f^{pre,NE}$, $\pi_f^{post,NE}$, $\pi_f^{pre,cll}$, $\pi_f^{post,cll}$, $\pi_f^{pre,def}$ and $\pi_f^{post,def}$.

a. Calculating $\pi_f^{pre,NE}$, $\pi_f^{post,NE}$

After having solved for marginal costs, it is now straightforward to calculate the pre merger Nash equilibrium profits ($\pi_f^{pre,NE}$), since one only needs to calculate the following F equations to obtain per firm profit levels

$$\pi_f = \sum_{j \in \Lambda} (p_j^{pre,NE} - c_j) M s_j(p^{pre,NE}), \quad (3.9)$$

As in section 3.3 the demand and price levels are observed in the data.

To determine post merger Nash equilibrium profits, one firstly needs to calculate post merger Nash equilibrium prices ($p^{post,NE}$). These solve the following system of first order conditions

$$s(p^{post,NE}) + (p^{post,NE} - c) \Delta^{post} \circ \mathbf{B}' = 0 \quad (3.10)$$

The only difference to equation 3.7 is the change in the ownership structure. The new set of equilibrium prices will now determine a new set of market shares for each product and the results can replace $p^{pre,NE}$ in equation 3.9 to determine the post merger Nash equilibrium profit levels for every firm. One would generally expect the post merger Nash equilibrium profits to rise for both merging and non merging parties. The reason is the reduction in concentration in the market and the subsequent slight drop in competition (unilateral effects) and hence $\pi_f^{post,NE} > \pi_f^{pre,NE}$.

b. Calculating $\pi_f^{pre,cll}$, $\pi_f^{post,cll}$

In a next step one can calculate the firm specific collusive prices and payoffs. By setting all elements in the ownership matrix equal to one (Δ^{cll}), the model allows for joint profit maximization among the players in the industry. That is, every competitor in the market will set the price of its products taking into consideration its effect on *all* other products. Therefore the optimal collusive

price is determined by maximizing the profit function $\pi_f^{cl} = \sum_{j \in J} (p_j - c_j) M s_j(p)$. The resulting J optimality conditions (in matrix notation)

$$s(p^{cl}) + (p^{cl} - c)\Delta^{cl} \circ \mathbf{B}' = 0 \quad (3.11)$$

determine the collusive price vector (p^{cl}). There is no differentiation between pre and post merger collusive prices. This is due to the fact that collusive prices will be identical pre and post merger and hence $p^{pre,cl} = p^{post,cl} = p^{cl}$. It follows logically that $\pi_f^{pre,cl} = \pi_f^{post,cl} = \pi_f^{cl}$.

Following Davis et al. (2008), there is no reason to believe that collusive prices would be different under a new ownership structure. Intuitively, this can be explained by the fact that by acting as a cartel, the ownership of the products are irrelevant to the firms since they jointly maximize their profits for all products in the market. Mathematically, it can be explained by the simple argument that post and pre merger all elements in the ownership matrix are set equal to one and therefore it is irrelevant which form the ownership matrix takes 'pre collusion'.

A set of strong assumptions underlies the calculation of collusive prices and profits. Firstly, I assume that a collusive agreement is made by all F players in the industry. This is a rather stark assumption and could easily be altered in extensions to the model. Secondly, I assume that firms collude at full collusive prices. This might exaggerate collusive profits and might in reality raise the suspicion of antitrust agencies. Thirdly, I assume that every firm sets collusive prices to all of its products rather than only a subset of its product mix. Equivalently, I assume that a defecting firm will defect from the collusive agreement with all of its products also.

c. Calculating $\pi_f^{pre,def}$, $\pi_f^{post,def}$

The defection price for firm f is the optimal price it can set for all its products given that all other firms set their prices at p^{cl} . Hence firm f will want to set the prices of its products (denoted by the vector p_f^{def}) such that they maximize the profit equation 3.4. The resulting first order condition takes the following form

$$s([p_f^{pre,def}, p_{-f}^{cl}]) + ([p_f^{pre,def}, p_{-f}^{cl}] - c)\Delta^{pre} \circ \mathbf{B}' = 0 \quad (3.12)$$

The optimal defection prices ($p_f^{pre,def}$) satisfy equation 3.12 and as above, the

profit levels are calculated by estimating the new demand levels at defection prices for f and collusive price for all other firms and plugging the results back in the profit function of the form 3.9.

Given that the post merger collusive prices remain the same as in the pre merger case, non merging parties will have identical defection prices and subsequent defection payoffs when renegeing on the collusive agreements. Hence $\pi^{post,def} = \pi^{pre,def}$ for non merging parties.

The merged entity however will revise its defection prices post merger due to its new augmented product mix. This evidently results in different levels of defection payoffs and therefore $\pi^{post,def} \neq \pi^{pre,def}$. It is more difficult to predict whether defection payoffs will increase (making collusion more difficult to sustain) or decrease (making collusion easier to sustain) post merger for the merging parties. On one side, the new larger entity might find it less profitable to defect from the collusive agreement, because the gain in additional market share is now smaller than it was previously for the smaller individual firms. On the other side, the new product mix might have been changed in such a way that the variety of products now covers a larger spectrum of products and product differentiation. This means that a less drastic fall in prices is required to achieve a similar if not greater market share.

4 The econometric demand model

4.1 Previous demand estimations

The most basic approach to demand estimation in heterogeneous product markets is based on a representative agent model, whose preferences are defined over the respective products. By assuming a constant price elasticity, one can simply regress quantities purchased of product j on average prices of all products to obtain a basic system of J (assuming there are J products) demand equations of the form

$$\ln(q_j) = \alpha_j + \sum_{k=1}^J \eta_{jk} \ln(p_k) + \varepsilon_j$$

Deaton and Muellbauer (1980) developed their almost ideal demand system (AIDS) largely on this linear form, in which the shares of the various brands are linearly related to the logarithm of real total expenditure and the logarithms of relative prices. Several problems arise from this methodology.

Firstly, consumer heterogeneity is not taken into consideration. This makes an estimation of demand in different geographic markets or in the same market at different points in time imprecise and difficult. Consumer preferences vary by markets or they might change over time. Consequently, one would expect a precise demand estimation to allow for varying price coefficients. This is not the case in representative agent models and suggests that a model accounting for consumer heterogeneity is needed.

Secondly, the basic models suffer from the often cited 'too many parameter' problem regarding the estimation of elasticities. If a market includes J products, J^2 parameters need to be estimated (i.e. one own price coefficient and $J - 1$ cross price coefficients for every one of the J products). The available data set might simply not be large enough to cope with that number of parameters.

Thirdly, these demand estimations are not flexible enough for a variety of IO applications such as the hypothetical introduction of a new good into the choice set of existing products. The latter two shortcomings are a direct result of defining preferences on products per se rather than product characteristics.

Three types of models have been developed to address the problem of estimating too many parameters. Firstly, Gorman (1971) formulates a multi-stage budgeting approach of demand estimation which aims at reducing the number

of elasticity parameters to be estimated. The idea is to split the market in different segments and estimate an AIDS type demand for each segment. Because cross price elasticities only need to be estimated within each segment, a smaller number of elasticities need to be computed. Hausman, Leonard and Zona (1994) are the first to apply this methodology to the US beer industry. They segment the products into categories such as Light, Premium and Superpremium and can subsequently find cross price elasticities more easily.

A pitfall of the multi-stage budgeting type of estimation however is the fact that different grouping and segmentation procedures are used in its applications. This can lead to demand estimations of the same product having very different results as pointed out by Ackerberg et al. (2006).

Secondly, Lancaster (1971) sets up a model which projects the consumer problem onto characteristic space rather than product space. Each product is thereby defined by a set of characteristics and preferences are defined over these product characteristics. The consumer's purchase decision can now be determined by the bundle of characteristics associated with a particular product that maximizes his utility (Mcfadden,1974) rather than the product per se. Traditionally, the Logit model is applied to calculate the individual probabilities of purchasing a product (i.e. market shares). This structural discrete choice model reduces the number of parameters to be estimated, since only the cross effects between product characteristics need to be estimated regardless of the number of products.

The usefulness of applications of the framework set out by Lancaster (1971) and Mcfadden (1974) was questioned due to two major limitations. The early models of discrete choice models did firstly not allow for unobservable product characteristics and secondly the functional forms used in the models restricted own and cross price elasticities substantially.

The first limitation was formulated by Berry (1994). It is highly likely that product characteristics included in the model do not capture all indirect utility a consumer attains from the purchase of a particular product. Berry therefore suggests to include an unobserved product characteristic term in the utility function, which is unobserved by the econometrician but observed by consumers and producers. Examples of 'unobservables' might be taste of the product, reliability of the product or brand recognition. The introduction of an unobservable product characteristics term however leads to endogeneity and further computational problems which will be described in more detail in section 5. BLP address this shortcoming by introducing a computational methodology that estimates

parameters consistently despite endogeneity. Details of this methodology will be described in section 4.2.

The second limitation was pointed out by Mcfadden (1974) himself, who acknowledged that the Logit model and its inherent IIA assumption (independence of irrelevant alternatives) used to compute individual purchasing probabilities led to questionable elasticities⁴. To circumvent this problem, one can allow the marginal utilities of product characteristics to be individual specific. This lets marginal utility vary by consumers and not only leads to more sensible elasticities but also introduces the afore mentioned consumer heterogeneity into the preference parameters. In their highly influential paper, BLP develop a 'Random coefficient Logit model', which incorporates consumer heterogeneity and allows for a consistent estimate of price and product characteristic parameters. Building on BLP, Nevo (2000) extends the model by introducing empirical demographic distributions to determine individual specific characteristic preferences. The author is thereby able to explicitly model consumer's heterogeneity of preferences over product characteristics. The model used in this paper is largely based on Nevo (2000) and will be described in further detail in the following section.

In the IO literature, the random coefficient model is seen as a reliable and precise demand estimation and is mostly criticized for (1) its computational complexity and (2) the assumption that consumers only purchase one single product out of the differentiated product supply mix. The latter criticism is valid since most consumers purchase or stock several different brands simultaneously. One can justify this restrictive assumption by claiming that consumers will only *consume* one product at a time (Nevo, 2000) and defining the choice period appropriately. The former criticism is only relevant when the results of the demand estimation are urgently required and is not a criticism of the estimation per se.

Thirdly and more recently, Pinkse, Slade and Brett (2002) develop a distance metric technique for estimating demand. The model specifies the cross price terms as a function of a brand's location in product characteristic space relative to other brands. It therefore allows for a projection of the consumer problem onto product characteristics while keeping the computational simplicity of AIDS type models. This type of model has been applied to the US (Rojas, 2006) and UK (Pinkse and Slade, 2004) beer industry and their results will serve

⁴see section 4.2 for further details of Logit model

as benchmarks for my subsequent results. A limitation of the distance metric procedure compared to the random coefficient model is the fact that it does not allow for heterogeneous price or product coefficients. It therefore paints a slightly less accurate and realistic picture of consumer demand and switching behavior.

4.2 Demand estimation used in this paper

The theoretical econometric model determining the demand system in this paper is a discrete choice model with random coefficients, first introduced by Berry (1994) and BLP. In each $t = 1, \dots, T$ market $i = 1, \dots, I$ consumers can purchase one of $J = 1, \dots, J_t$ products or no product at all. A market is defined as a city-quarter combination and within each market product characteristics, aggregate quantities sold and average prices are observed for each product. A market is defined as a city-quarter combination. Note that in my model not all products are offered in all markets, in fact every market sees a different choice set of products and hence the need for a subscript t under J_t . Following Nevo (2000) consumer i attains an indirect utility of purchasing product j which is of the following linear additively separable form⁵

$$u_{ijt} = \beta_i x_{jt} - \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt} \quad (4.1)$$

p_{jt} is the price of product j in market t , x_{jt} is a $K \times 1$ vector of product characteristics that co-determine the utility level achieved from purchasing the product in question and remain constant over time, ξ_{jt} is the set of unobserved (by the econometrician) product characteristics and ε_{ijt} is a mean zero stochastic error term. α_i is the marginal disutility of price and β_i is a column vector of taste coefficients for each of the K product characteristics. The observed characteristics used in this model are a measure of availability of the product in the relevant market, alcohol content, size of the average container, whether the beer is imported and advertising and promotional activities. Evidently, these product characteristics will not fully cover the utility level a consumer attains from purchasing one of the brands. There are a variety of unobserved (to the econometrician) product characteristics that influence the buying decision of

⁵Nevo(2000) also introduces an income term in the indirect utility function, which I ignore because the term is cancelled out in the subsequent demand estimation

the consumer. In this model they might include brand value and taste of beer. ξ_{jt} captures these unobservables. It is most likely that ξ_{jt} will be correlated with prices, since producers will take these unobservable product characteristics into consideration for their pricing decision. To correct for this endogeneity problem, a set of appropriate instruments needs to be defined. That is, a set of variables that are correlated with prices but largely independent of the unobservable term. It is possible to facilitate the search for an instrumental variable by adding brand dummies into equation 4.1. This splits the unobservable term into a brand specific fixed effect (ξ_j) and a market specific deviation of unobservable components from brand specific means ($\Delta\xi_{jt}$) and hence $\xi_{jt} = \xi_j + \Delta\xi_{jt}$. It is now easier to find an instrumental variable, because the instrument only needs to capture the correlation between prices and market specific deviations rather than having to cover the brand specific correlation also. For the rest of section 4, ξ_{jt} will be used as notation for the unobserved product quality rather than $\xi_j + \Delta\xi_{jt}$ for simplicity. The inclusion of brand dummies does pose further computational problems, which will be addressed at the end of section 6.

The random coefficient model includes consumer heterogeneity in the coefficients alpha and beta. This heterogeneity enters via demographic variables, which will be split into observed (D_i) and unobserved variables (v_i). Consequently, α_i and β_i can be defined by the following

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i, \quad v_i \sim P_v^*(v), \quad D_i \sim \widehat{P}_D^*(D) \quad (4.2)$$

where D_i is a $d \times 1$ vector of observed demographic variables and v_i is a set of demographics which influence the purchasing decision but are not observed. The application of the model in this paper sees income, age, marital status and sex used as observed demographic variables. Even though we have no detailed information on the distribution of these demographics, it is possible to estimate an empirical distribution by using additional data sources such as the Current Population Survey (CPS).

The use of actual data to estimate the distribution of observable demographics reduces the reliance on arbitrary distributional assumptions and therefore increases the fit of the model. The unobserved demographic variables v_i might include informations that are not obvious for the beer purchasing decision but might nevertheless influence it, such as a consumer's tolerance of the beer.

Regarding the distribution of these unobservables, one usually has no information as no publicly available survey is likely to cover that much detail of consumer preferences. Therefore a distributional assumption of v_i is required. In this model it will be assumed to be of a multivariate Normal form. The $(K + 1) \times d$ matrix of coefficients Π measures the effect demographic variables have on the marginal utility for each of the observed product characteristics. The $(K + 1) \times (K + 1)$ matrix of parameters Σ captures the effect of unobservable demographics on taste characteristics.

One can define $\theta = (\theta_1, \theta_2)$ as a vector containing all the parameters of the model. Combining equations 4.1 and 4.2 splits the indirect utility into a mean utility level (δ_{jt}) and its individual specific deviation ($\mu_{ijt} + \varepsilon_{ijt}$) and results in the following equation

$$u_{ijt} = \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \mu_{ijt}(x_{jt}, p_{jt}, v_i, D_i; \theta_2) + \varepsilon_{ijt} \quad (4.3)$$

$$\mu_{ijt} = [-p_{jt}, x_{jt}](\Sigma v_i + \Pi D_i)$$

$$\delta_{jt} = \beta x_{jt} - \alpha p_{jt} + \xi_{jt}$$

θ_1 consists of the constant coefficients α and β and θ_2 is a matrix of parameters capturing the interaction between demographics and prices and product characteristics (Π and Σ). Given this utility specification, consumers are assumed to purchase the good that maximizes their indirect utility.

Consumers can only chose a single product that maximizes their utility not a bundle of products. The choice set of products a consumer can choose from however does not only include the existing products but also an 'outside good'. This outside good is equivalent to the consumer deciding to purchase no product at all. The model therefore allows consumers to switch out of the market by not choosing to purchase any product at all, ensuring that the overall volume consumed does not stay constant when prices change. To include the outside good in the choice set of the consumer, it is important to define a utility level for it. I follow the standard procedure (BLP (1995) and Nevo (2000)) and simply normalize the mean utility to zero (i.e. $\delta_0 = 0$).

Since every consumer is defined by their individual demographics and product specific shocks, the set of individual characteristics that will induce the purchase of product j can be defined by the following set

$$A_{jt}(x_{.t}, p_{.t}, \delta_{.t}, ; \theta_2) = \{(D_i, v_i, \varepsilon_{i0t}, \dots, \varepsilon_{iJt}) | u_{ijt} \geq u_{ilt} \forall l = 0, 1, \dots, J\}$$

The set A_{jt} defines the individuals whose characteristics are such that the utility of purchasing product j in market t is higher than over all other products in market t . The market share of product j is therefore simply the probability that observed and unobserved demographics (v_i and D_i) and the product specific shocks (ε_{ijt}) fall into the region A_{jt} .

Assuming ties occur with zero probability, the probability of consumers purchasing the product is given by integrating over the distributions of these variables.

$$\begin{aligned} s_{jt} &= \int_{A_{jt}} dP^*(D, v, \varepsilon) \\ &= \int_{A_{jt}} dP^*(\varepsilon | D, v) dP^*(v | D) d\widehat{P}_D^*(D) \\ &= \int_{A_{jt}} dP_\varepsilon^*(\varepsilon) dP_v^*(v) d\widehat{P}_D^*(D) \end{aligned} \tag{4.4}$$

The last step is only valid, when assuming that v_i and D_i are independently distributed. This integral therefore yields the probability of individuals having demographic variables such that product j is the product that maximizes their utility and therefore calculates the market shares. The distributional assumptions of the demographic variables and the error term are of crucial importance to define market shares and elasticities. Two different adaptations of the Integral in equation 4.4 and their implications will be described in the following.

4.2.1 Logit model

In the Basic Multinomial Logit model all consumer heterogeneity is assumed to be included in the error term rather than in the coefficients of product characteristics and price. This implies that μ_{ijt} is equal to zero in equation 4.3. The error term (ε_{ijt}) is assumed to be iid with a Type I extreme value distribution. Because the demographic variables do not enter equation 4.3 when μ_{ijt} is equal to

zero, the integral in equation 4.4 simply becomes $\int_{A_{jt}} dP_{\varepsilon}(\varepsilon)$. Given the assumed distribution and the resulting cumulative distribution function $P_{\varepsilon}(\varepsilon) = e^{-e^{\varepsilon}}$ this integral has a full analytic solution of the form⁶

$$s_{jt} = \frac{\exp(x_{jt}\beta - \alpha p_{jt} + \xi_{jt})}{1 + \sum_{k=1}^J \exp(x_{kt}\beta - \alpha p_{kt} + \xi_{kt})} \quad (4.5)$$

The elasticities are obtained by differentiating equation 4.5 with respect to prices, multiplying by the respective price and dividing by the market share. This results in the following equations

$$\eta_{jkt} = \frac{\partial s_{jt} p_{kt}}{\partial p_{kt} s_{jt}} = \begin{cases} -\alpha p_{jt}(1 - s_{jt}) & \text{if } j = k \\ \alpha p_{kt} s_{kt} & \text{otherwise} \end{cases} \quad (4.6)$$

Note that the own price elasticities are determined by a constant term (α), market shares (s_{jt}) and the price of the product (p_{jt}). Since the market shares are generally relatively low values, the own price elasticity will be driven by the price. The functional form assumed for prices in the utility function is therefore crucial in determining elasticities. As Nevo (2000) points out, if the log of prices would enter the utility function, the own price elasticities would be approximately constant. The heavy dependency of the own price elasticity on an (arbitrary) choice of functional form makes the informative value of the elasticities questionable. Furthermore, Rasmussen (2007) notes that a low price implies a lower own price elasticity and therefore a higher price mark-up. There is however no economic rationale that would support this assumption. On the contrary, one usually observes higher price mark ups for products with higher marginal costs (luxury products compared to budget products).

The cross price elasticities resulting from the Logit model are also unreasonable. As can be seen from 4.6, the cross price elasticity of a product, say j , depends on the constant marginal utility of income α , the price of product k and the market share of product k and is independent of product j 's market share or price. The latter means that a price increase in product k will have the same percentage effect on all other k_{-1} product market shares. A simple example can illustrate this point very clearly. As the price of Budweiser increases consumers will switch equally to Heineken and Kingsbury (a non alcoholic beer). This

⁶See Train (2003) for a derivation

however is not realistic as the product characteristics of Heineken evidently are more similar to Budweiser's than those of Kingsbury.

This problem arises, because the error term (and consumer heterogeneity) has an iid structure. Consumers will rank the products based on their individual specific error term (ε_{ijt}). This means that if a consumer chooses Budweiser over Heineken and Kingsbury (in that preference order), $\varepsilon_{i,Budweiser} > \varepsilon_{i,Heineken} > \varepsilon_{i,Kingsbury}$. However, the Logit model does not correlate $\varepsilon_{i,Heineken}$ and $\varepsilon_{i,Budweiser}$ which means that the $\varepsilon_{i,Heineken}$ for the average Budweiser-consumers is the same as for the whole population on average. And since all ε are identically distributed, the Logit model assumes that the proportion of individuals in the population ranking Kingsbury or Heineken as second best is identical. Hence, consumers will not switch to more similar products, but will switch to all products equally.

The great advantage of the Logit model is its simplicity and ease of computation, its disadvantages however are too substantial to use it as a reliable demand model. I will use the Logit model only in a first step in this paper to obtain initial parameter estimates, which I can refine with the random coefficient model.

4.2.2 Random coefficient Logit model - The full model

To overcome the problem leading to unrealistic elasticities, the model needs to allow for correlation between the error term across products. So that consumers switching to another product when prices increase will most likely switch towards one with a shock that is correlated with their initially preferred product. The correlation between the utilities of different products will be determined by the individual specific utility term (μ_{ijt}). The correlation will therefore be a function of demographic information and product characteristics. That is, products with similar characteristics will be correlated and consumers with similar demographics will have similar product ranking and therefore substitution patterns. As a result, the coefficients of price and product characteristics capture consumer heterogeneity and each individual will have different marginal utilities for product characteristics and price. Coming back to the integral 4.4, it is now difficult to find an analytic solution to this integral. If one is to keep the iid Type I extreme Value distribution for the error term ε_{ijt} , then the individual market share of product j in market t is defined by (Rasmussen, 2007)

$$s_{ijt} = \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^J \exp(\delta_{kt} + \mu_{ikt})} \quad (4.7)$$

Equation 4.7 defines the probability of individual i in market t purchasing product j . It is now possible to rewrite the integral in equation 4.4 and therefore the overall market share as

$$s_{jt} = \int_v \int_D \frac{\exp(\delta_{jt} + \mu_{ijt})}{1 + \sum_{k=1}^J \exp(\delta_{kt} + \mu_{ikt})} dP_v^*(v) d\widehat{P}_D^*(D) \quad (4.8)$$

Hence the market share of product j in market t is determined by the average of the individual market shares weighed by the occurrence of that type of individual in the population. The resulting elasticities (Rasmussen, 2007) take the following form

$$\eta_{jkt} = \frac{\partial s_{jt} p_{kt}}{\partial p_{kt} s_{jt}} = \begin{cases} -\frac{p_{jt}}{s_{jt}} \int_v \int_D \alpha_i s_{ijt} (1 - s_{ijt}) dP_v^*(v) d\widehat{P}_D^*(D) & \text{if } j = k \\ \frac{p_{kt}}{s_{jt}} \int_v \int_D \alpha_i s_{ijt} s_{ikt} dP_v^*(v) d\widehat{P}_D^*(D) & \text{otherwise} \end{cases} \quad (4.9)$$

The two integrals in equation 4.9 average the price sensitivities and the individual probabilities of purchasing the product across the I individuals. Three important features of the full model need to be highlighted.

Firstly, the elasticities are more realistic than those of the Logit model. Recall that in the Logit model own price elasticities were determined by the functional form of prices in the utility function. In the random coefficient model however own price elasticities will be determined by individual specific price sensitivities (α_i). This means that higher price sensitivities will lead to higher average own price sensitivity (in absolute values). This is an intuitive result, since an increase in the price of a product whose consumers are very price sensitive will lead to a higher number of consumers switching to a neighboring product. Regarding the cross price elasticities, the random coefficient model allows for different substitution patterns rather than identical ones for each product (as was the outcome in the Logit model). Equation 4.9 suggests that the cross price elasticities now depend on market shares of both products in question and not only on one. And because the market shares are now driven by demographics and their interaction with prices and product characteristics

via μ_{ijt} , the cross price elasticities will also be driven by the demographics. More similar consumers will therefore switch to more similar products.

Secondly, the integral in equation 4.8 and the elasticities now have no closed analytic form (unlike the Logit model) and need to be estimated by simulation methods, which will be described in section 6. This is the key computational difference of the random coefficient Logit model compared to the basic Logit model.

Thirdly, in order to calculate the integral, information on demographic distribution need to be known or assumed. As already noted, random draws of the CPS are used to approximate for the distribution of the observed demographic variables (D) and the unobserved demographic variables (v) are assumed to follow a multivariate normal distribution.

5 Endogeneity and instruments

The framework set out in Section 4 suffers from endogeneity between the unobservable product characteristics and prices. A valid instrumental variable (IV) for prices is therefore required to correct for this problem. Recall that by introducing brand specific dummies into the indirect utility function the unobservable product characteristics consist only of the market specific deviation ($\Delta\xi_{jt}$). These market specific deviations (e.g. local demand shock) will be known to producers and they will set their market specific mark-ups accordingly. One can think of price mark-ups following the function $f_j(\Delta\xi_{jt})$ and hence the price of the product being determined by $p_{jt} = c_j + f_j(\Delta\xi_{jt})$, namely marginal costs plus a price mark up. A valid instrumental variable should therefore capture as best possible the product specific marginal cost (informative), while being independent of the market specific mark-up (valid).

I use a set of instrumental variables first suggested by Hausman et al (1994). The price of product j in city m and quarter s is instrumented by the average price of product j in all other cities surrounding m for each quarter.

Two crucial assumptions underlie this set of instrumental variables. Firstly, my identifying assumption is that market specific unobservables ($\Delta\xi_{jt}$) are independent across markets. This ensures that the price mark up in one city is independent of the price mark up in another city and implies the IV's are valid. This is a strong assumption and might be violated. If for example it becomes publicly known that an unhealthy ingredient is being used in the production process of one brand, it is very likely that demand for that product will fall over all markets.

Secondly, in order for prices in other cities to reflect the marginal costs of a product in the instrumented city, common marginal cost for each brand are assumed across markets. This ensures that the IV's are informative. One might however expect these costs to vary strongly across markets, which would result in poor correlations across prices of the same product. In order to best correct for potential differences in these costs, I average prices of each brand in 6 specified regions within the US. Hence I do not use the national average of prices in other cities but the regional average. This should partially correct for differences in production costs, since grocery stores in cities in the designated regions are likely to purchase their beer from the same plant. Similarly, transportation and shelving costs are likely to be more similar in each region than across the US,

because of differences in fuel prices and/or staff costs. Nevo (2000) and Rojas (2008) even instrument for these costs individually by using supermarket-wage level data and population density as proxies for shelf space and labour cost. I was not able to do so due to a lack of data, particularly regarding supermarket wage level data.

BLP suggest a further set of instrumental variables which consist of (1) each product characteristic of the product being instrumented, (2) the sum of each product characteristic of all the products marketed by the same producer and (3) the sum of each characteristic for all rival firm products. The underlying assumption of these instruments is that products whose rivals have more similar characteristics will have lower price mark ups. Results of regression estimates with both sets of instruments are presented in section 8.

6 Estimation

This section describes in detail the estimation routine determining the demand parameters based on the underlying data and the theoretic econometric model. The aim of the estimation is to find values for the parameters θ that will minimize the distance between the observed market share and the market share suggested by the econometric model. The code for the estimation was written in Matlab 7 and based on the code used in Nevo (2000). I will outline the computation of both the Logit and random Logit models in an effort to highlight their computational differences. For a consistent estimation of the random coefficients, a four step procedure has been adopted.

1. Calculating market shares with initial values of $\delta_{.t}$ and θ_2

Firstly, the estimated market shares $s(\delta_{.t}; \theta_2)$ for a given value of δ_{jt} and θ_2 are calculated. The integral in equation 4.8 has no closed form analytical solution and has to be solved numerically. The most common method used to approximate an integral of this form is by simulation (Train, 2003). The simulation procedure consists of computing an average probability of purchasing product j in market t across ns individuals, based on their individual probability of purchasing the product (s_{jti}).

$$s_{jt} = \frac{1}{ns} \sum_{i=1}^{ns} s_{ijt} \quad (6.1)$$

The individual probability of purchasing a particular product is given by 4.7. In this model an individual is defined by a set of observed demographic variables (D_{id}) obtained from random draws of their respective distributions, a set of unobserved demographic demand shocks (v_{it}) drawn from a multivariate normal distribution and an individual specific error term (ε_{it}). I have drawn 20 individuals per city and not per market. Therefore in one city it is always the same individuals making the purchase decision regardless of quarter and year. This is necessary for the second step to work and for the calculated market shares to sum up to one. By assuming that ε follows a Type I extreme value distribution, equation 6.1 has a closed form solution⁷

⁷To facilitate the equation in 6.2 I have included the p variable in the x vector

$$s_{jt}(\delta_{jt}, \theta_2) = \frac{1}{ns} \sum_{i=1}^{ns} \frac{\exp[\delta_{jt} + \sum_{k=1}^K x_{jt}^k (\sigma_k v_i^k + \pi_{k1} D_{i1} + \dots + \pi_{kd} D_{id})]}{1 + \sum_{m=1}^J \exp[\delta_{mt} + \sum_{k=1}^K x_{mt}^k (\sigma_k v_i^k + \pi_{k1} D_{i1} + \dots + \pi_{kd} D_{id})]} \quad (6.2)$$

Calculating this market share however requires a set of starting values for the mean utility level (δ_{jt}) and parameters $\theta_2 = (\sigma, \pi)$. These initial values can be chosen arbitrarily, since convergence will ensure that the optimal parameters for θ_2 will be determined. Hence, the simulation procedure calculates the market share for a given mean utility level and given θ_2 .

Equation 6.2 needs only to be calculated for the random coefficient full model and not for the Logit model. The reason being that in the basic Logit model all consumer heterogeneity enters the indirect utility function through the error term and market share is simply determined by equation 4.5. Note that the applied simulation methodology does not minimize the simulation errors and more efficient estimates could be determined with other computationally more complex simulation techniques. Using these techniques however is out of the scope of this paper.

2. *Finding the mean utility values (δ_t) that minimize the distance between observed and calculated market shares*

The second step in the estimation routine consists of finding a δ_{jt} such that the calculated market share obtained in the first step equals the observed market shares (S_t) while keeping the initial value of the coefficient matrix θ_2 . Hence for each market the following system of equations has to be solved.

$$s(\delta_t, \theta_2) = S_t \quad (6.3)$$

Berry (1994) proves that equation 6.3 has a unique solution implying that there exists a unique mean utility level (δ^*) satisfying the equation $\delta_{.t}^* = s_{.t}^{-1}(S_t)$. For the Logit model to satisfy equation 6.3 it suffices to compute the equation $\delta_{.t} = \ln S_t - \ln S_{0t}$ to identify the mean utility level which satisfies equation 6.3⁸. For the full model the system of equations 6.3 needs to be computed numerically, since the parameters θ_2 enter equation 6.2 in a non linear way. To overcome this problem one can solve the equation by recursive iteration. BLP suggest a contraction mapping which consists of iterating down the series 6.4

⁸See Appendix for details of derivation.

until a δ^H has been established whose difference from δ^{H-1} is not greater than some tolerance level.

$$\delta_{.t}^{h+1} = \delta_{.t}^h + \ln S_{.t} - \ln s(\delta_{.t}^h, \theta_2) \quad (6.4)$$

More specifically, the process requires an initial value for δ^h (I use the result from the Logit model) to which the difference $\ln S_{.t} - \ln s(\delta_{.t}^h, \theta_2)$ is added. The resulting δ^{h+1} is then used to determine a new $s(\delta_{.t}^{h+1}, \theta_2)$ and a new $\delta_{.t}^{h+2}$ is calculated with the same procedure. This recursive iteration process is repeated until the difference between two mean utility levels is below some tolerance level. In order to speed up the iteration process the tolerance level is adjusted according to the number of iterations. So, for the first 100 iterations the tolerance level is set at 10E-8 and for each 50 iterations after that the tolerance level increases by an order of ten. Even though this reduces computation time, the time cost of the iteration remains substantial.

3. Linear representation of the structural error term

In a third step, the estimation routine requires a linear representation of the structural error in terms of estimated parameters from the previous steps. The structural error term here is simply the unobserved product characteristic term as defined in section 4. This linear representation is needed to subsequently define an estimator that minimizes its squares. Rewriting equation 4.3 and defining $\Delta\xi_{jt}$ as ω_{jt} yields

$$\omega_{jt} = \delta_{jt}(S_{.t}; \theta_2) - (x_{jt}\beta + \alpha p_{jt}), \quad (6.5)$$

where the first term is the result obtained from step 2 and x_{jt} and p_{jt} are observed product characteristics and price respectively. Recall that because brand specific dummies are included in the vector of product characteristics x_{jt} , ω_{jt} is defined as the city specific deviation from the unobserved brand characteristics ($\Delta\xi_{jt}$) and not ξ_{jt} . Note that the parameters $\theta_2 = (\sigma., \pi.), \alpha$ and β are still the initial starting values. In order to find the parameter values ($\theta = (\theta_1, \theta_2)$) that minimize this error term, one could minimize the objective function $\omega(\theta)' \omega(\theta)$, i.e. minimize the squared errors. This OLS type regression would only be possible if the error term defined in equation 6.5 is uncorrelated with all independent variables. As outlined in section 5 however, it is reasonable to assume that brewers will take into consideration their (to the econometrician) unobserved product

characteristics when setting their product prices. As a result, prices have to be instrumented for with the instrumental variable Z such that $E[Z\omega(\theta^*)] = 0$. This condition means that at the true set of parameter values (θ^*), the population moment is equal to zero and implies that the chosen instrumental variable is valid. As described in section 5, the Hausman (1995) type instrumental variable has been used in this paper with the necessary identifying assumptions that satisfy the validity requirements. In order to find a set of parameters that satisfy the population moment condition as closely as possible, it is necessary to construct a GMM estimate, which minimize the weighted quadratic distance of the sample analog of the moments condition $Z\omega(\theta)$, i.e.

$$\theta = \arg \min_{\theta} \omega(\theta)' Z \Phi^{-1} Z' \omega(\theta) \quad (6.6)$$

The weight matrix Φ is a consistent estimate of $E(Z'\omega\omega'Z)$. Following Nevo (2000) it is possible to firstly compute an estimate of θ , say $\hat{\theta}$ using $\Phi = Z'Z$ ⁹. Secondly, $\hat{\theta}$ is used to compute a new weight matrix: $E[Z'\omega(\hat{\theta})\omega'(\hat{\theta})Z]$, which in turn is used to calculate a new estimate of θ .

4. Finding the set of parameters that minimize the GMM objective function

The fourth and final step of the estimation routine consists of the search for the value of θ that minimizes the GMM objective function, i.e. satisfies equation 6.6. Differentiating the GMM objective function with respect to θ yields the standard Two stage least squares (2SLS) GMM estimate $\hat{\theta} = (X'Z\Phi^{-1}Z'X)^{-1}X'Z\Phi^{-1}Z'\delta$. In the Logit model all parameters enter this GMM estimate in a linear fashion, i.e. $\hat{\theta}_1 = (X'Z\Phi^{-1}Z'X)^{-1}X'Z\Phi^{-1}Z'(\beta x_{jt} - \alpha p_{jt} + \xi_{jt})$ ¹⁰. Hence consistent estimates of θ_1 can be obtained by 2SLS regression of $\ln S_{jt} - \ln S_{0t} = \delta_{.t} = \beta x_{jt} - \alpha p_{jt} + \xi_{jt}$ with an appropriate set of instruments. This regression equation follows from step 2 and the condition for the estimated market shares to equal the observed market shares.

For the random coefficient model however, the GMM estimator is less straightforward to calculate, because θ_2 enters the error term non-linearly and θ_1 enters linearly. Therefore a non linear search over θ has to be performed. To reduce computation time, it is possible to restrict the non linear search to θ_2 by determining the matrix of coefficients θ_1 by the following 2SLS regressor

⁹This is possible only if homoscedastic errors are assumed which in turn implies an optimal weight matrix proportional to $Z'Z$

¹⁰Recall that in the Logit model the parameters θ_2 are not existent because $\mu_{ijt} = 0$.

$$\hat{\theta}_1 = (X'Z\Phi^{-1}Z'X)^{-1}X'Z\Phi^{-1}Z'\delta(\hat{\theta}_2) \quad (6.7)$$

I use the Nelder-Mead (1965) non derivative "simplex" search method for the non linear optimization procedure¹¹. Once a set of estimates for the θ coefficients are obtained, the whole process (starting at step 1) is repeated (with the estimated θ) until the minimum value in the GMM objective function is found.

A computational difficulty when using brand dummy variables is that the taste coefficients (β) cannot be retrieved, because a large part of the necessary information will be included in the brand dummy coefficients. By applying Chamberlain's (1982) minimum distance procedure, one can easily overcome this difficulty. Essentially it consists of regressing the brand dummy coefficients (d) on the product characteristics (X) plus the unobserved brand quality (ξ), i.e. $d = X\beta + \xi$. Assuming $E(\xi|X) = 0$, the consistent GLS estimator for the taste coefficients is

$$\hat{\beta} = (X'V_d^{-1}X)^{-1}X'V_d^{-1}\hat{d}$$

where \hat{d} is the vector of coefficients estimated from the procedure described in the previous section and V_d is the covariance matrix.

The computation time required for the estimation procedure was of around 180 minutes with step 2, the contraction mapping process, being the most time consuming. I have not included the Matlab code in the appendix but it can be made available upon request.

¹¹A detailed description of this methodology and alternatives can be found in Nevo(2000)

7 The data

The source of the data set is the Information Resources Inc. (IRI) Infoscan database and was kindly made available by the Food Marketing Policy Center¹². The data is extensive supermarket scanner data and covers beer purchases from all supermarkets in 58 US metropolitan areas (cities). Beers with regional market shares below 3% are not included in the data, which results in the data covering on average 92 % of supermarket beer purchases in every city. The data set consists of quarterly observations over the period 1988 to 1992 (inclusive) and covers 64 different beers produced by 13 brewers. All products and their respective producers are listed in Table 2. The Table shows the dominant position some brands hold in the US beer market with notably Budweiser, Coors Extra Gold and Miller Lite being among the brands with the highest average market shares. A market is defined as a city quarter observation and since not all cities are observed in each quarter, the total number of markets is 928. In fact only 49 of the 58 metropolitan areas are covered over all twenty quarters with the other cities covered over different time periods¹³. Another feature of the data set is that the number of brands sold differs in every market. On average 37 brands are sold in each market ranging from 24 (quarter 1 1989 Wichita, KS) to 48 (quarter 4 1991 Buffalo/Rochester, NY). This results in 33,891 observations.

Table 3 summarizes the descriptive statistics of the variables that are covered in the data set. The IRI data includes quarterly average product prices per brand, volume and units sold of each product in each market and a measure of product coverage in the respective market. Product coverage is measured as a ratio of the sum of all commodity value (ACV) sold by stores carrying the product divided by the ACV of all stores in the market. Consequently, a ratio of unity implies that the product is sold in all stores within a city. The average coverage ratio is relatively high (76%) suggesting that most products when offered in one city are offered across most grocery stores. Additional information regarding product characteristics had to be gathered from a variety of sources. These are listed in the bottom half of Table 3 and include alcohol content (Alcohol), size of the average container sold (Size), national advertising levels (Advertising), a dummy that indicates whether the beer is imported (Import) and the percentage of volume sold with promotional activity (Promotional).

¹²I am very grateful to Dr. Ronald Cotterill for making this extensive dataset available

¹³See Appendix for a full listing of included cities and their respective observed quarters.

Table 2
List of brewers (acronyms), brands and average market shares¹⁴ in %

Brewer	Brand	share	Brewer	Brand	share	Brewer	Brand	share
Anheuser Busch (AB)	Budweiser	14.4		Lone Star	1.1		Hamm Light	0.5
	Bud Dry	2.3		Lone Star Light	0.5		Olympia	1.5
	Bud Light	6.4		Old Style	1.6		Blue Ribbon	1.7
	Busch	5.9		Old Style Light	1.1		Red White and Blue	0.4
	Busch Light	2.5		Rainier	2.4	Phillip Morris (PM)	Genuine Draft	3.3
	Michelob	1.3		Schmidts	0.9		Meister Brau	1.5
	Michelob Dry	0.8		Sterling	1.1		Meister Brau Light	0.5
	Michelob Golden Draft	1.0		Weidemann	1.0		MGD Light	2.3
	Michelob Light	1.3		White Stag	0.8		Miller High Life	3.1
	Natural light	3.2	Genesee (GE)	Genesee	3.7		Miller Life	9.0
Adolph Coors (AC)	O'Douls	0.5		Kochs	0.8		Milwaukee's Best	6.1
	Coors	2.3	Grupo Modelo (GM)	Corona	0.5	Stroh (S)	Goebel	0.8
	Coors Extra Gold	8.4	Goya (GO)	Goya	0.4		Old Milwaukee	3.8
	Coors Light	6.5	Heineken (H)	Heineken	0.6		Old Milwaukee Light	1.5
	Keystone	0.7	Labatt (LB)	Labatt	0.4		Piels	3.2
	Keystone Light	1.4		Labatt Blue	0.9		Schaefer	2.1
	Black Label	1.0		Rolling Rock	0.4		Schlitz	1.1
	Blatz	1.2	Molson (MO)	Molson	0.4		Stroh	1.1
	Heidelberg	1.0		Molson Golden	0.5	Matts FX (MA)	Matts	0.8
	Henry Weinhard Ale	1.2		Old Vienna	1.5		Utica Club	1.0
G. Heileman (GH)	Henry Weinhard P.R.	0.8	Pabst (P)	Falstaff	0.4			
	Kingsbury	0.3		Hamm	1.1			

Note that the average market shares tabulated are only over those markets in which the product is being sold and are slightly different from national market shares. The table therefore gives an overview of the average importance of each brand in the markets they are sold in.

Advertising data was obtained from the Leading National Advertising annual publication and covers average national brand advertising across ten media types. It varies considerably across brands and producers. Premium brands like Budweiser and Miller reach quarterly advertising expenditures of up to 40 million \$ and smaller regional brands like Sterling or Red White and Blue have no advertising expenditures at all. Information on alcohol content was collected from various specialized sources (especially Case et al, 2000).

Table 3
Summary of statistics

Variable	Units	Mean	Std. deviation	Min	Max
<i>IRI data</i>					
Price	\$ per Liter	1.48	0.47	0.1	3.54
Volume sold	Liters (000)	192.52	520.7	0.11	21698.2
Coverage	%	74.01	28.61	0.26	100.0
Units sold	(000)	57.88	149.9	28.1	6111
<i>Other variables</i>					
Alcohol	%	4.48	0.94	0.4	5.25
Import	0/1	0.13			
Size	Liters per unit	3.11	0.96	0.61	10.63
Advertising	\$ (000,000)	3.54	6.25	0	40.28
Promotional	%	30.46	20.97	0	100

Table 3 shows that the average alcohol content is 4.48% with the non alcoholic beer Odoul's having an alcohol content of 0.4% and the strongest beer Henry Weinhard Ale having 5.25% alcohol content. I have opted not to categorize the different brands into segments such as 'Premium', 'Superpremium' or 'Light'. Firstly, I was not able to find a consistent segmentation of the beers in my sample. Secondly, all category definitions were mainly pegged to the two product characteristics advertising and alcohol content. And since I already include those separately in my estimation, the added value, if any, of including the different categories would be marginal. The product variable Size has simply been calculated by dividing the volume sold by the number of units, thereby giving an average value of the container size of each beer.

Next to product characteristics, demographic variables are also required for the demand estimation. They will determine the heterogeneity of the price and product characteristics coefficient in my demand system. The demographics have all been obtained from the Current Population Survey 1990. For each city I have drawn twenty individuals and their respective sex (Male = 1), annual income, marital status (married = 1) and age. This resulted in 4640 demographic observations. Only individuals aged 21 or over were sampled, since the legal drinking age in most US states is 21. Table 4 summarizes these variables across all sampled individuals.

Table 4

List of demographic variables

Variable	Unit	Mean	Std. deviation	Min	Max
Sex	0/1	0.52			
Income	\$ (000)	22.74	16.46	0.62	99.96
Marital Status	0/1	0.54			
Age	years	37.31	12.47	21	76

A final data specification necessary to apply the demand estimation, is the determination of an overall market size. This market sizing is relatively difficult, since no given methodology can be applied. The overall market size has to be approximated. In previous studies, this has been done by defining a variable to which the market is proportional to and multiplying it by an assumed proportionality factor. Nevo (2000) takes the population in a city as the proportional variable for his estimation of ready to eat cereal demand. Bresnahan et al (1997) use the number of office based employees as proportional variable to estimate the demand for computers. In their demand estimation for cars BLP take the total number of households as the variable to which the overall market is proportionate to.

Following a similar method, I assume that the market size is proportional to the drinking age population in each city. The proportionality factor I choose is equal to one Liter per week. In other words, I assume that the overall potential market size for beer consumption consists of every individual above the legal drinking age to drink 1 Liter of beer per week. Even though this might seem like a rather arbitrary value, the results of the demand estimation are robust to changes in the market size.

8 Results

8.1 Demand results

This section summarizes the results obtained from the Logit demand estimation and the random coefficient demand estimation. As already noted, the Logit model leads to unrealistic cross price elasticities and can therefore not be used as a reliable demand system for the merger simulation. Logit results are nevertheless presented here, because they give very useful insights into the subsequent design of the final model. In particular, the results show (a) the importance of instrumenting for price, (b) which instruments to use and (c) highlights the importance of product characteristics in explaining the variation of the market shares across brands.

8.1.1 Logit model results

Table 4 shows the results of the parameters α and β of six different regressions of the equation $\ln S_{jt} - \ln S_{0t} = \beta x_{jt} - \alpha p_{jt} + \xi_{jt}$ with their standard errors in parenthesis. All regressions are based on 33,891 observations and all include time dummy variables, i.e. yearly and quarterly dummies to correct for seasonal and 'macroeconomic' fluctuations. Regressions (1)-(3) are OLS regressions. Columns (4)-(6) show the results of three 2SLS regressions, where the price level has been instrumented with regional average price levels ('Hausman') or the summation of product characteristics ('BLP'), as described in section 5.

A further source of differentiation for equations (1)-(6) is the inclusion of different dummy variables. Next to time dummy variables (present in all estimations and not made explicit in the table), I have selectively corrected for either brand fixed effects and/or city fixed effects; both are denoted by 'brand' and 'city' in the next to last row. For example, equation (3) is an OLS regression which has been estimated with all observed product characteristics and time, city and brand specific dummies.

When comparing equations (1) and (2), one can see the importance of including product characteristics to explain the variation of the error term. In particular the better fit of model (2) ($R^2=0.828$) and the statistical significance

of the coefficients indicate that the observed product characteristics are essential in explaining the differences in market shares for each product. The R^2 lies well above that of other random coefficient estimation such as BLP (1995) or Nevo (2001) suggesting that the chosen product characteristics are very precise at explaining the difference in market shares.

Furthermore, the price coefficient more than doubles. This results in the percentage of inelastic demand estimations for every brand in every market to drop to 18.2 % as can be seen in the last row of the table. The Logit model does not restrict own price elasticities and they vary for every brand-market combination. Hence, the elasticities follow a distribution and the last row of the table indicates the percentage of own price elasticities that are below unity¹⁵. An elasticity below unity in absolute terms means that the percentage change in quantity demanded is smaller than that in prices. This in turn implies an increase in total revenue as prices rise, which does not seem a very sensible outcome for beer. This row therefore gives an indication as to how realistic the estimated demand system is.

When correcting for brand fixed effects, the price coefficient remains constant but the β coefficients change. This is to be expected given that the brand dummy coefficients in equation (3) will 'mop up' a large part of the information captured by the product characteristics coefficients in equation (2).

(3) has a high R^2 (86.69%), meaning that only 13% of the variation of the market shares is explained by the structural error term. Note that this error term is not quite identical to the one defined as $\Delta\xi_{jt}$ in section 4.2. The error term in model (3) is the set of unobservable product characteristics that are not explained by city, brand or time variation and is therefore a reduced form of the city specific variation $\Delta\xi_{jt}$.

The instrumental variable regressions in equation (4)-(6) address the endogeneity between the price level and the unobserved product characteristics. Note that I have not presented any results with *only* BLP type IV's. The outcomes of such estimations were unrealistic with price coefficients always taking positive values. I therefore report on results obtained with Hausman instruments or in the case of equation (5) Hausman *and* BLP type instruments.

¹⁵Own price elasticities have been calculated according to equation 4.6

Table 4
Results of Logit model

x	OLS			2SLS(IV)		
	(1)	(2)	(3)	(4)	(5)	(6)
Price	-0.4796 (0.0189)	-0.9657 (0.0159)	-0.9506 (0.0294)	-0.3214 (0.0850)	-1.1689 (0.0766)	-1.2802 (0.0777)
Alcohol	-	0.0623 (0.0044)	-0.1352 (0.0255)	0.0805 (0.0252)	-0.1064 (0.0237)	-0.1137 (0.0237)
Coverage	-	4.343 (0.0171)	4.0904 (0.0171)	4.2877 (0.1975)	4.0863 (0.0171)	4.0842 (0.0172)
Import	-	0.1599 (0.0189)	0.5748 (0.1240)	-0.1840 (0.1480)	0.6475 (0.1264)	0.6846 (0.1266)
Size	-	0.1823 (0.0058)	0.2249 (0.0059)	0.2138 (0.0019)	0.2098 (0.0077)	0.2021 (0.0077)
Advertising	-	5.75e-05 (7.55e-07)	8.75e-06 (1.24e-06)	8.59e-06 (1.56e-06)	8.78e-06 (1.24e-06)	8.79e-06 (1.24e-06)
Promotional	-	0.7776 (0.0245)	0.8882 (0.0228)	1.0815 (0.0303)	0.8471 (0.0265)	0.8261 (0.0266)
Instruments				Hausman	Hausman, BLP	Hausman
R^2	0.1978	0.8287	0.8696			
1st stage R^2				0.9170	0.9474	0.9472
1st stage F test				3984.77	3870.31	4005.07
Overid. test				714.27(30.14)	649.12(37.65)	295.15(30.14)
Dummies				brand	city, brand	city, brand
% inelastic demands	90.87	18.23	19.91	99.75	5.34	1.24

The high 1st stage R^2 and 1st stage F-test values indicate that the instruments are valid and strong¹⁶. I also tested for endogeneity (Hausman test) of prices and found that the residual of the first stage regressions (for regressions (4)-(6)) is highly statistically significant when included as independent variable in an OLS type regression.

The importance of correcting for the endogeneity of prices becomes apparent when comparing equations (3) with (5) and (6). Even though most of the product characteristic coefficients do not vary much, the price coefficient becomes considerably more negative reaching -1.280 in equation (6). This would indicate an upward bias of the OLS estimator. The structural error term is therefore positively correlated with prices. This is an intuitive conclusion and confirms our assumption made when introducing the unobserved product characteristics in section 4.2, namely that products with a higher unobserved quality are priced higher. By correcting for this endogeneity, the bias is reduced and the coefficient becomes more negative. The Hausman type instruments seem to address this bias better than the combination of Hausman and BLP instruments, resulting in only 1.24% of demands to be inelastic in equation (6).

A potential imprecision of both sets of instruments however is that the overidentification test is rejected in all models, suggesting that the identifying assumptions made in section 5 might not be valid. The overidentification test (or Sargan test) results are shown in the third to last row with the respective critical value at the 5% significance level in parenthesis. However, it is unclear whether the rejection of the test is due to the large number of observations or the fact that the IV's are not valid. It is well known that with a large enough sample a Chi-squared test will reject essentially any model (Nevo, 2000). The test value of the overidentification test is minimized by including city fixed effects into the 2SLS regression (comparing (4) with (5) and (6)). This suggests that including demographic information in the model improves the validity of the instruments. Hence, in order to minimize the probability of rejecting the overidentification test, my final model should as best possible correct for demographics by incorporating demographic information.

¹⁶The contribution of the instruments to the R^2 can further be tested when comparing the R^2 of the first stage regression with the R^2 of the first stage regression excluding the instruments. The difference in R-square's for the sets of instruments varies between 20% and 25 %, confirming the strength of the instruments.

Based on the discussion of these results, my aim is to develop a full model that approximates the structure of regression (6), because (a) the instruments in equation (6) seem to correct for the endogeneity bias better than any other (combination of) instruments, (b) the probability of overidentification of the IV's is minimized when correcting for demographic city specific effects and using Hausman instruments and (c) the observed product characteristics seem to explain the variation in market shares very well.

8.1.2 Full model results

Table 5 presents the results of the full model. The estimates are based on the utility function 4.3 and were estimated with regional average prices as instrumental variables. The results are based on 33,891 observations and the regression includes time and brand dummy variables. The first column shows the means of the distribution of marginal utility coefficients α and β and their respective asymptotically robust standard errors in brackets. The product characteristic coefficients (β 's) and the price coefficients (α 's) follow a stochastic distribution which is determined by the empirical distribution function of the demographic variables with which they are interacted. The second column shows the impact of the unobserved demographic variable (assumed to be normally distributed) on the marginal utility of each product characteristic. The last four columns show the elements of the matrices Σ and Π and indicate how strong the influence of each of the demographic variables is on the β 's and α 's. Note that not all the product characteristics have been interacted with demographic variables. 'Advertising' and 'Promotional' are assumed to be largely independent of demographics. Their effect on consumer's utility is assumed to be homogeneous across the population.

The interpretation for the estimates in Table 5 can be explained using the example of the product characteristic 'Size'. The marginal utility of 'Size' follows a distribution determined by the dummy variable 'Sex' (1=male, 0=female) and a normal distribution. The mean of the marginal utility is 0.0184, suggesting that the mean consumer prefers a larger beer container size. The impact of the normal shock and the variable 'Sex' on the marginal utility is determined by the parameters 0.286 and 0.903 respectively. Ignoring the normal shock for now, the coefficient on the dummy variable 'Sex' means that the marginal utility of the container size is 0.903 higher for male consumers than for female consumers.

In other words, male consumers prefer larger unit sizes when consuming beer than do female consumers. This certainly seems a rather intuitive result.

The mean β 's were obtained from the Minimum Distance procedure described in section 6. They are all of the expected sign and intuitively sensible. The average consumer likes imports, beers with high advertising levels, beers with a high coverage rate and beers that are sold in large unit sizes. Recall that the MD procedure was necessary because of the inclusion of brand dummies and the subsequent difficulties in identifying the contribution to the utility of each product characteristics (Step 4 in section 6). A Chi-squared test (Chamberlain, 1982) can test the significance of the observed characteristics capturing this brand dummy information. The result of this test (MD χ^2) is reported in table 5 and confirms the significance of the product characteristics.

All slope coefficients are statistically significant with the exception of the product characteristic 'Alcohol'. Alcohol content therefore does not seem to be a product characteristic that influences utility of the consumer significantly. This result seems to fall in line with previous demand estimations for beer (Pinkse and Slade (2004) and Rojas (2008)), which also find a positive marginal utility for alcohol but only just statistically significant at 10% significance levels. Interestingly, when interacted with the demographic variable 'marital status' (1=married,0=single), one can see that the effect of marginal utility is negative for married individuals. Thus married individuals actually prefer less alcohol when consuming beer compared to their unmarried counterparts.

The 'Std. deviation' figures are all statistically insignificant. This implies that the variation of the respective slope coefficient can not be explained by a normally distributed shock. Coefficients on the demographic interaction variables however are almost all statistically significant, which clearly indicates that the chosen demographics can explain the true heterogeneity of the coefficients. These results show that my aim of incorporating demographics in a more precise way in my model has been achieved and the chosen demographics approximate the (city specific) heterogeneity very well.

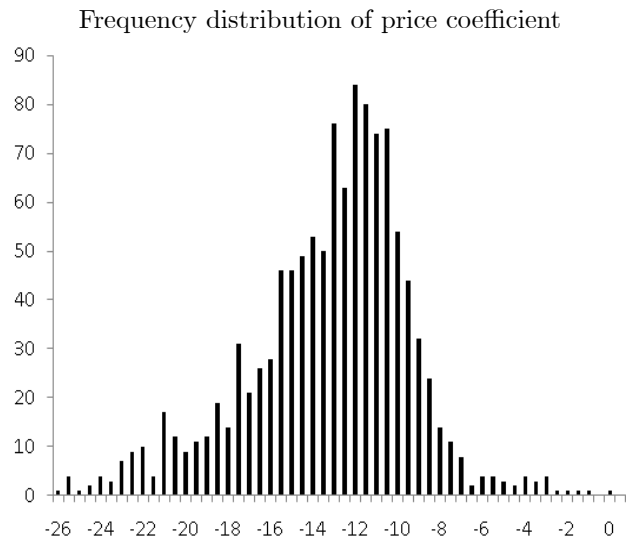
The mean price coefficient is negative (-14.1467) and statistically significant. The difference to the estimated price coefficient in the Logit model is quite considerable, suggesting that the included heterogeneity highlights the price sensitivity of beer consumers. The frequency distribution of the price coefficient (Figure 1) shows a strong skewness to the left and does not seem to follow a Normal distribution.

Table 5
Results of Full model

Product characteristics x	Mean (α and β 's)	Std. deviation (σ)	Interaction with Demographic variables			
			Sex	Age	marital status	
Price	-14.1467 (2.7491)	0.2709 (0.5063)	-	-0.2760 (0.1213)	-	0.3964 (0.0696)
Alcohol	0.0273 (0.0393)	0.0253 (0.1239)	-	-	-1.1182 (0.2624)	-
Coverage	0.1608 (0.0408)	0.0427 (0.5453)	-	0.2069 (0.2264)	-	0.6300 (0.0940)
Import	1.0167 (0.3979)	0.4890 (2.4820)	-	-	-	-
Size	0.0814 (0.0075)	0.2860 (0.1620)	-	0.9030 (0.3170)	-	-
Advertising	0.0001 (1.26e ⁻⁶)	-	-	-	-	-
Promotional	2.3512 (0.8712)	-	-	-	-	-
MD R ²						0.22394
MD χ^2						16502.6
% of α 's >0						0.1

This strengthens the point that the normally distributed 'shocks' (v) cannot explain the heterogeneity of price sensitivity across the population. For price coefficients, the driver of the variation are the distribution of income and age respectively. The positive coefficient on Income (0.3964) shows that consumers with higher income levels will generally be less price sensitive, which is a very logical conclusion. The negative coefficient on 'Age' (-0.276) implies that older individuals will be more price sensitive compared to younger consumers. This seems slightly less intuitive, but might be due to the fact that older consumers tend to switch their consumption to other alcoholic drinks more readily than do younger consumers. Note that the price coefficient can become positive. But this is the case for only 0.1 % of all sampled individuals. This is a highly insignificant number and underlines the sensibility of the model.

Figure 1



8.1.3 Full model elasticities

The own price elasticities that result from the full model have been calculated according to equation 4.9 for every product in every market. Because the integral in equation 4.9 does not have a closed analytic solution, I have applied a

simulation procedure over the sampled individuals¹⁷. Next to estimating elasticities, this calculation also allows for the estimation of the vector of derivatives **B**. The full model treats each city-quarter combination as a separate market and in order to obtain the overall price elasticities, I have taken the average own price elasticity for each product across all markets. Table 6 presents a sample (see appendix for all 64 own price elasticities) of own price elasticities and also includes the estimated price cost margin resulting from the demand system. The marginal cost is estimated by equation 3.8. and the price cost margin is simply calculated by $\frac{p-c}{p}$.

Table 6

Average own price elasticities (η_j) and price cost margins (*PCM*)

Brewer	Product	η_j	<i>PCM</i>
Anheuser-Busch	Budweiser	-3.960	41.04%
	Busch Dry	-6.688	23.57%
	Busch	-5.802	35.25%
Adolph Coors	Coors	-6.552	16.37%
	Coors Extra Gold	-7.443	13.38%
Labatt	Labatt	-9.041	11.33%
Molson	Molson	-8.519	17.44%
Phillip Morris	Miller Lite	-4.916	23.65%
Stroh	Old Milwaukee	-6.052	16.50%
	Schlitz	-6.514	14.35%

All own price elasticities are negative and greater than unity in absolute terms. The median own price elasticity is -6.345 and the estimated elasticities are generally slightly higher than those reported by Pinkse and Slade (-4.6), Hausman, Leonard and Zona (-4.98) and Rojas (-3.53). The median price cost margin is 20.02% and is considerably higher than what the industry literature usually suggests as industry average price cost margin for that period (Tremblay (2005) suggests around 10%). The reason for this discrepancy might be the fact that my data covers retail prices only and not wholesale prices. As a result, the price mark up for brewers might be biased upward. All price cost margins are presented in the Appendix and yield very sensible results with premium

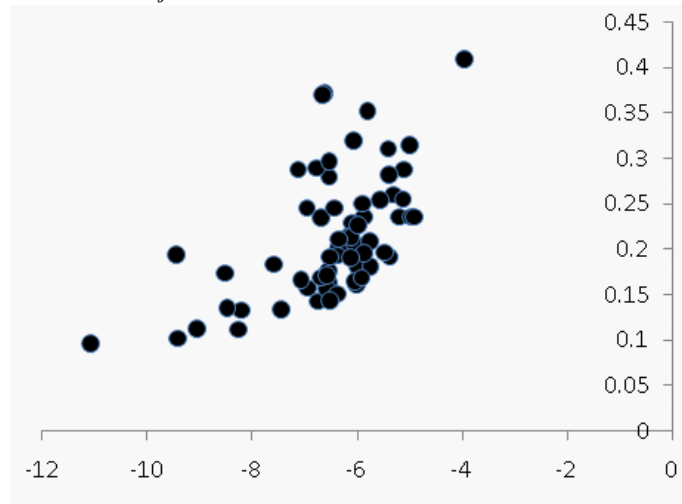
¹⁷The simulation follows the same procedure as equation 6.2

brands such as Budweiser, Busch and Miller Lite having above average price cost margins.

One can see from Table 3 that the correlation between price cost margin and own price elasticities is negative. This negative relationship is persistent across all 64 brands as can be seen from Figure 2. Economic theory predicts exactly that relationship. High margin products (luxury goods) will attract consumers that are less price sensitive. A price increase for these products will therefore result in fewer consumers switching away from the product than would be the case for low margin products.

Figure 2

Relation of η_j (x-axis) and PCM (y-axis) for all 64 brands



The cross price elasticities are all presented in the Appendix. I have selected a few that highlight the general substitution patterns resulting from my demand estimation in table 7. Each element η_{jk} in row j and column k of table 7 show the change (as percentage) in market share of product j following a 1% increase in prices of product k . Similar to the own price elasticities, I present the average cross price elasticities across all 928 markets. All cross price elasticities have been calculated according to equation 4.9 with a basic simulation method. They are all positive and have a median value of 0.0685 and are similar in dimension to the ones in Pinkse and Slade (2004) and Rojas (2008).

Table 7
Sample of average cross price Elasticities

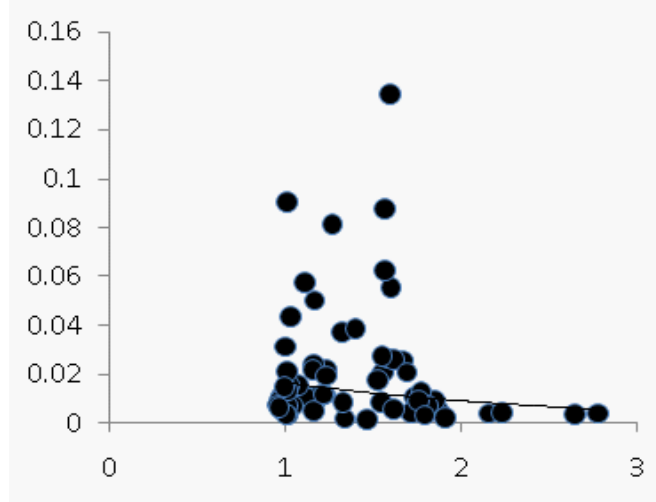
	Price	Size	Brands										
			Bud	Bud Lt	Busch	Coors	Coors Lt	Lone Star	Old Style	Old Vi-enna	MHL	Miller Lt	Old Mil.
AB Bud	1.56	3.3	-3.960	0.230	0.201	0.072	0.242	0.024	0.032	0.056	0.090	0.324	0.098
AB Bud Dry	1.66	3.5	0.608	0.258	0.227	0.080	0.270	0.026	0.035	0.065	0.102	0.359	0.109
AB Bud Light	1.60	3.4	0.509	-5.121	0.188	0.067	0.225	0.022	0.039	0.056	0.086	0.301	0.091
AB Busch	1.27	3.5	0.825	0.347	-5.802	0.109	0.365	0.037	0.048	0.078	0.139	0.484	0.147
AB Busch Light	1.32	3.8	0.797	0.338	0.297	0.106	0.356	0.035	0.046	0.086	0.134	0.469	0.143
AB Michelob Golden Draft	1.74	3.5	0.782	0.330	0.291	0.104	0.348	0.035	0.045	0.079	0.132	0.458	0.139
AC Coors	1.55	3.3	0.619	0.260	0.230	-6.558	0.274	0.028	0.036	0.059	0.105	0.363	0.110
AC Coors Light	1.56	3.5	0.574	0.243	0.215	0.076	-5.382	0.026	0.033	0.060	0.097	0.337	0.102
AC Keystone Light	1.24	3.7	0.767	0.328	0.286	0.102	0.343	0.033	0.043	0.086	0.130	0.451	0.137
B Henry Weinhardt Ale	1.77	2.6	0.589	0.245	0.229	0.078	0.255	0.028	0.032	0.053	0.096	0.332	0.097
B Lone Star	1.20	3.2	0.769	0.352	0.287	0.100	0.332	-7.124	0.042	0.071	0.132	0.452	0.134
B Old Style	1.18	3.4	0.746	0.324	0.288	0.101	0.334	0.032	-6.875	0.078	0.126	0.435	0.131
B Rainier	1.52	2.7	0.638	0.267	0.221	0.090	0.288	0.027	0.035	0.054	0.109	0.371	0.109
GE Genessee	1.40	2.7	0.573	0.245	0.207	0.077	0.254	0.023	0.033	0.056	0.100	0.342	0.103
M Old Vienna	1.58	4.1	0.620	0.296	0.239	0.090	0.299	0.019	0.031	-8.479	0.105	0.364	0.111
P Olympia	1.01	3.8	0.783	0.330	0.289	0.105	0.349	0.035	0.044	0.081	0.133	0.456	0.137
P Pabst Blue Ribbon	1.16	3.5	0.808	0.342	0.303	0.107	0.357	0.036	0.046	0.079	0.137	0.471	0.144
PM Genuine Draft	1.61	2.9	0.531	0.223	0.195	0.070	0.235	0.023	0.031	0.056	0.090	0.314	0.095
PM Miller High Life (MHL)	1.55	3.2	0.618	0.252	0.228	0.082	0.273	0.028	0.036	0.060	-6.548	0.363	0.110
PM Miller Lite	1.57	3.5	0.549	0.232	0.203	0.073	0.243	0.024	0.032	0.058	0.093	-4.917	0.098
PM Milwaukee's Best	1.01	3.4	0.823	0.348	0.306	0.110	0.365	0.036	0.047	0.081	0.139	0.482	0.147
S Old Milwaukee	1.11	3.5	0.832	0.348	0.308	0.111	0.368	0.038	0.048	0.075	0.141	0.487	-6.052
S Old Milwaukee Light	1.16	4.0	0.783	0.334	0.290	0.104	0.349	0.032	0.044	0.089	0.133	0.459	0.141
S Piel's	1.01	3.0	0.537	0.240	0.173	0.244	0.244	0.014	0.029	0.089	0.098	0.325	0.103
Outside good	-	-	0.134	0.055	0.081	0.020	0.062	0.011	0.010	0.027	0.028	0.087	0.057

Recall from section 4.2.1, that the Logit model restricts substitution patterns in such a way that a price increase of one product will lead to an equal percentage rise in market shares across all other products, i.e. the numbers within each column of table 7 would be identical in the Logit model. A diagnostic as to how well the full model overcomes this Logit restriction is by constructing ratios of the maximal values to the minimal values within each column and observing how they differ from the ratio resulting from the Logit model (unity). These ratios have a median of 3.5 and range from 2.4 to 11.2 and are therefore considerably different than one. The cross price substitution patterns generally indicate that consumers will switch to products that are more similar following a price increase. For example, a price increase in Bud Light leads to consumers switching to (among others) Busch Light, Keystone Light or Old Milwaukee Light. Similarly, a price rise of Old Vienna (a product sold in large container sizes) will lead to preferable switches to beers with larger container sizes such as Old Milwaukee Light, Olympia or Busch Light. Note however that the substitution patterns are not as straightforward throughout the table. In some cases a price increase might lead to a switch to products that are not necessarily 'closest' in Euclidean product space. This is likely to be due to the fact that price plays a significant factor in the purchasing decision. A similar product in this model is therefore a combination of the included product characteristics and prices.

In addition, I have included the cross price elasticity between respective products and the outside good in Table 7. In other words, the estimations provide an answer to the question what the percentage increase in the outside market share is with an increase in prices of a particular product. For the full data set, these range from 0.135 to 0.001 with a mean of 0.021. To test for the validity and sensibility of these estimates, one can simply assess the correlation between prices and the outside elasticities. One would generally expect a negative correlation, since consumers of highly priced products will be less likely to stop their consumption when prices increase marginally. Figure 3 shows that the economic rationale is weakly confirmed in my demand estimation.

Figure 3

Correlation of η_{0k} (y-axis) and *prices* (x-axis) for all 64 brands



8.2 Collusion game results

This section describes the results obtained from the simulation of the merger between G. Heileman and Stroh. All results in this section are based on pricing rules and profit functions derived in section 3. For ease of interpretation I focus my attention on the results of the six largest brewers. These firms supply 88% of total supermarket beer consumption in my data set. Furthermore, I have not listed all individual products in this section to make my results and the subsequent interpretation more accessible. The effect of the merger on the prices and payoffs of the remaining firms and products can be found in the appendix.

8.2.1 Presentation of merger simulation results

a. Price dynamics

Table 8 shows the price dynamics under three different modes of competition pre and post merger. It lists the price changes for the two most profitable brands of each brewer and the average price change across all of its products.

Table 8
Changes in prices of the two most profitable brands caused by merger of GH and S

Brewer	market share	Brands	Pre-merger					Post-merger				
			$p^{pre,NE}$	$p^{pre,cl}$	Δ	$p^{pre,def}$	Δ	$p^{post,NE}$	Δ	$p^{post,def}$	Δ	
AB	35.43%	Budweiser	1.59	1.83	15.01%	1.58	-13.91%	1.59	0	1.58	0	
		Budweiser Light <i>all AB brands</i>	1.60	1.98	23.91%	1.80	-9.09%	1.60	0	1.80	0	
PM	23.69%	Miller Light	1.56	2.17	38.45%	1.82	-16.00%	1.56	0	1.82	0	
		Genuine Draft <i>all PM brands</i>	1.61	1.83	13.58%	1.80	-1.71%	1.61	0	1.80	0	
AC	11.18%	Coors Light	1.56	1.99	27.61%	1.79	-10.16%	1.56	0	1.79	0	
		Coors <i>all AC brands</i>	1.55	1.77	13.97%	1.69	-4.61%	1.55	0	1.69	0	
S	9.33%	Old Milwaukee	1.11	1.46	31.15%	1.42	-2.73%	1.21	8.60%	1.38	-3.11%	
		Piels <i>all S brands</i>	1.03	1.22	19.19%	1.20	-2.37%	1.03	-0.03%	1.18	-1.62%	
GH	4.35%	Henry Weinhard Ale	1.77	2.26	27.67%	1.97	-13.00%	1.80	1.42%	1.97	0.23%	
		Weideman <i>all GH brands</i>	1.04	1.40	34.41%	0.96	-31.24%	1.04	-0.31%	0.97	1.22%	
P	4.17%	Olympia	1.01	1.38	36.10%	0.98	-28.52%	1.01	0	0.98	0	
		Pabst Blue Ribbon <i>all P brands</i>	1.16	1.47	27.20%	1.33	-9.37%	1.16	0	1.33	0	
			-	-	26.88%	-	-	-	-	0		

The firms are listed in order of size and the first column indicates their average national market share in 1992. All prices are calculated except for the competitive prices (columns (1) and (4)), which are *observed* prices. Focusing on the pre merger case first, the results hint at a substantial difference in pricing strategies among the firms in the industry. In particular one can observe a strong relationship between firm size and price changes: in a collusive setting, smaller firms are able to increase their prices by a higher proportionate amount than large firms. When defecting however, the smaller firms reduce their prices more than do their larger rivals. This results in a higher variation in price cost margins for smaller firms (Table 9).

Turning now to the post merger equilibrium prices, one can see that for the non merging parties the competitive prices remain constant compared to the pre merger case¹⁸. Post merger competitive prices for the merged entity fall on average (-2.2%), however prices of the most profitable brands rise. The post merger defection prices for the merged firms increase slightly by 3.35% .

Table 9

Price Cost Margins and price changes

Brewer	Pre-merger		
	$PCM^{pre,NE}$	PCM^{cll}	$PCM^{pre,def}$
Anheuser-Busch	27.6%	38.0%	31.7%
Philip Morris	22.3%	34.2%	25.8%
Adolph Coors	15.8%	30.7%	13.3%
Stroh	18.1%	33.6%	19.8%
G. Heileman	26.2%	37.1%	6.1%
Pabst	18.7%	36.8%	13.2%

b. Changes in static payoffs

Table 10 summarizes the static payoffs and the critical discount factors pre and post merger resulting from the price changes for the 6 largest brewers.

¹⁸Note however that they are not identical. The calculated price change from pre to post merger is marginal (of the order of 0.00001%) and are far inferior to the degree of detail of the reported results. Davis and Huse (forthcoming) find similarly small changes in prices in their application.

Table 10
static payoffs (000's) and critical discount factors

Brewer	Pre-merger				Post-merger				
	$\pi^{pre,NE}$	$\pi^{pre,cll}$	$\Delta(NE - > cll)$	$\pi^{pre,def}$	$\Delta(cll - > def)$	δ^{pre}	$\pi^{post,NE}$	$\pi^{post,cll}$	δ^{post}
Anheuser-Busch	69,499	77,110	11.0%	81,784	6.1%	0.3805	69,702	77,110	0.3869
Philip Morris	61,700	74,525	20.8%	79,273	6.4%	0.2702	61,767	74,525	0.2712
Adolph Coors	20,261	27,456	35.5%	29,020	5.7%	0.1785	20,291	20,291	0.1791
Stroh	3,314	7,288	119.9%	7,961	9.2%	0.1447	10,288	15,902	0.0794
G. Heileman	6,340	8,614	35.9%	9,236	7.2%	0.2143	10,288	15,902	0.0794
Pabst	283	530	86.9%	570	7.6%	0.1409	283	530	0.1410

The profit figures in Table 10 are lower than actual profit figures, due to the data only covering supermarket purchases in 58 selected cities. They therefore only constitute a part of the overall beer sales that will generally include on-premise and Liquor store sales and exports. Pre merger the smaller firms enjoy a far higher proportional profit increase from competitive to collusive payoffs resulting from the higher collusive prices. The percentage increase from collusive to defection payoffs are similar across firms (between 5.7 % and 9.2%).

All payoffs post merger are identical to those pre merger with the exception of competitive profits and defection profits for the merged entity. As anticipated in section 3.4, $\pi_f^{post,NE} > \pi_f^{pre,NE}$ for all firms. Interestingly, the defection payoffs of the merged firms post merger sum up to less than the two individual defection payoffs of S and GH pre merger.

c. Differences in critical discount factors pre merger

The resulting pre merger critical discount rates follow logically from the above interpretations. For the two largest firms (PM and AB), the gains from collusion are relatively modest compared to their gains of defection. Consequently, their critical discount factors are higher than for any other firm (0.38 and 0.27), making the two largest firms least likely to collude.

S and P are the two firms that are most likely to collude (0.14 and 0.14), as their additional gains from defecting are very small compared to the payoffs they would receive from collusion.

GH and AC lie in between with discount factors of 0.21 and 0.17 respectively. Both firms seem to gain in similar proportions from collusion. However, the fact that GH drops its prices to very low levels (PCM is only 6%) and thereby reaching a much greater market share when defecting, makes defection slightly more interesting for GH than for AC.

Except for GH, it seems that the smaller firms will be more likely to collude than the larger firms. The main reason is that the smaller firms can raise their prices more easily in a collusive agreement without losing too much of their demand. When it comes to defecting however, the smaller firms need to drop their prices drastically to reach more consumers and make defection worthwhile. This makes their collusive payoffs more appealing than their defection payoffs. This is a rather striking result and clearly very different from the homogeneous product case, where the smaller firms have less incentive to collude, as they could potentially capture the whole market (Tirole, 1988).

d. Changes in critical discount factors resulting from merger

The last column of table 10 shows the critical discount factors post merger. The critical discount factors for the non merging parties are marginally higher post merger, implying a lower likelihood of collusion. The reason is that the 'punishment', in form of reversion to the competitive profits, after defection has been detected is slightly less severe than pre merger. The small change in δ suggests that the merger has almost no effect on the likelihood of collusion of the non merging parties, as their true discount factor would have to be between δ^{pre} and δ^{post} to change the competitive strategy of a firm.

For the merging firms, the effect is rather different. The critical discount rate post merger falls quite considerably to 0.0794, suggesting that the merging parties post merger are actually more likely to collude than before the merger.

8.2.2 Interpreting the results

The above presentation of results shows that the merger simulation led to two key insights. Firstly, the fact that smaller firms have lower critical discount factors than larger firms. Secondly, the fact that a higher concentration in the market leads to different changes in the likelihood of collusion among the firms. In particular that the merging parties have a higher likelihood of colluding and non merging parties are almost unaffected by the merger. Both insights stand in stark contrast to the classic homogeneous goods collusion model first introduced by Friedman (1971).

The first result is due to the ability of smaller firms to raise their collusive prices higher above competitive levels than their larger counterparts can. Because the defection payoffs are similar to those of larger firms, a collusive outcome is favoured by smaller firms. The exact reason as to why smaller firms can raise their prices to such extents in a collusive agreement will not be investigated further in this paper. Due to a lack of robust theoretical research and no clear evidence in my data, I feel that further interpretations of this result would be too speculative .

The second insight leads to the question as to why critical discount factors of the merging parties fall post merger? It appears to be due to the unilateral effects (increase in $\pi_f^{post,NE}$) being cancelled out by considerably lower defection payoffs (π_f^{def}). Recall that the critical discount factor is established by

$\delta_f \geq \frac{\pi_f^{def} - \pi_f^{cll}}{\pi_f^{def} - \pi_f^{NE}}$. Given that $\pi_f^{cll} > \pi_f^{NE}$, a fall in π_f^{def} will reduce the critical discount rate. Table 11 shows that considerably lower defection payoffs for Stroh's products seem to be the reason for the fact that $\pi^{pre,def} > \pi^{post,def}$ for the merged entity.

Table 11

Payoffs post and pre merger for merged entity

Brewer	profits (000's)			
	$\pi^{pre,NE}$	$\pi^{post,NE}$	$\pi^{pre,def}$	$\pi^{post,def}$
Stroh	3,314	3,468	7,961	7,106
G. Heileman	6,340	6,819	9,236	9,280

When considering product specific defection payoffs of Stroh (Table 12), it becomes clear that the two most profitable brands Piels and Old Milwaukee experience a sharp drop in defection payoffs. So why do the defection payoffs of Piels and Old Milwaukee fall so drastically post merger?

Table 12

Defection payoffs for Stroh

Product	profits (000's)	
	$\pi^{pre,def}$	$\pi^{post,def}$
Goebel	13	8
Old Milwaukee	3,004	2,961
Old Milwaukee Light	141	155
Piels	4,383	3,563
Schaefer	397	389
Schlitz	18	25
Stroh	1	2

The reason is that the two most profitable products of Stroh, Piels and Old Milwaukee, are highly substitutable with the products of GH¹⁹. In fact, for 11 of the 15 products of GH the cross price elasticities rank among the top ten cross price elasticities of the 64 products. And since the above collusion model restricts the firms to defect with all their products, these products will effectively compete for the same consumers when defecting. So consumers that

¹⁹ see cross price elasticities in appendix

before the merger switched to Piels and Old Milwaukee when S defected, now might be switching to GH products or indeed remain GH consumers. As a result Piels and Old Milwaukee are scrambling for demand in the post merger defection scenario, even with lower post merger defection prices (see Table 8). The products of the new merged firm are therefore cannibalizing each other when the firm is defecting. For the firm as a new entity the defection profits are therefore less than they were for both firms individually, making collusion a more attractive option than defection.

When simulating defection behavior such that firms defect with all their products, a merger between two firms with highly substitutable products is likely to lead to a fall in defection payoffs post merger.

9 Conclusion

This paper has analyzed the collusive effects of a merger between the fifth (GH) and fourth (S) largest brewer in the US beer industry in 1996 by firstly estimating a random coefficient demand system and applying its results to Friedman's (1971) theoretic collusive game. The results of the demand estimation are precise, reliable and realistic and suggest a slightly higher price elasticity among beer consumers than previous studies. The product characteristics chosen as product 'differentiators' capture the difference in market shares of each brand very well. Furthermore, the chosen empirical demographic distribution are precise measures of the heterogeneity in preferences and prices across the population.

The application of the merger simulation method has shown two insights. Firstly, the likelihood of collusion was increased for merging parties and remained approximately constant for non merging parties. The reason being that defection payoffs were lower post merger for the merging parties due to product cannibalization when defecting. Secondly, the smaller firms were more prone to collusion in the US beer industry because of their ability to raise collusive prices far above competitive levels.

On the basis of these results, the FTC was right to approve the merger, since collusive behavior was not drastically altered for more than one player in the market. The overall likelihood of collusion was therefore not significantly changed by the merger since it takes more than one firm to collude. Another implication of the results is that having dominant players in an industry does not necessarily lead to a higher intensity of collusion. In fact the dominant players are less prone to collusion than smaller firms under a scenario where all firms are simulated colluders.

The merger simulation methodology and its results have several limitations and extensions to which I want to direct the reader's attention to. Firstly, I have employed a basic collusive model and a variety of different extensions could be applied that might change its results. These could include the modelling of entry and exit, capacity constraints, efficiency gains and/or the risk of being fined by an antitrust agency. The model is sufficiently flexible to deal with these extensions and they could easily be incorporated in future research or in antitrust applications.

Secondly, due to the nature of my data, I have to opt for quarterly interaction

among the players in the industry. This seems an unlikely periodicity for an industry with a relatively high frequency of interactions such as the brewing industry.

Thirdly, the critical discount rates are far below what would usually be expected from firm specific discount rates²⁰. The downward bias of the critical discount factor is due to excessively high collusive payoffs. They are a direct result of (1) modelling collusion by all firms, which raises collusive prices and (2) allowing firms to collude at full collusive prices. It can be argued that both assumptions are somewhat unrealistic since antitrust agencies might become suspicious at such high prices. In reality, firms would agree to collude at lower prices. Identifying these price levels is difficult, since they would be determined by a variety of unmeasurable factors. One could however create a range of collusive price levels in subsequent applications of this methodology.

Despite these limitations, the benchmark case as analyzed in this paper provides valuable insights into cross firm comparisons within the industry and helps identify those firms that are more prone to collusion. The application of this methodology in further antitrust cases could therefore provide an additional puzzle-piece to a mainly qualitative assessment of coordinated effects.

²⁰This seems to be a reoccurring problem in the evaluation of critical discount factors (see Davis and Huse).

References

- [1] Abreu, D. (1988) 'On the Theory of Infinitely repeated Games with Discounting', *Econometrica*, Vol. 56, No2, pp 383-396
- [2] Akerberg, D., L. Benkard, S. Berry and A. Pakes (2006) 'Econometric Tools for Analyzing Market Outcomes', *Chapter in Handbook of Econometrics*
- [3] Berry, S. (1994) 'Estimating Discrete-choice models of product differentiation', *RAND Journal of Economics*, Vol. 25, pp 242-262
- [4] Berry, S., J. Levinsohn, A. Pakes (1995) 'Automobile Prices in Market Equilibrium', *Econometrica*, Vol. 63(4), pp 841-90
- [5] Bresnahan, T. (1982) 'The Oligopoly Solution concept is identified', *Economics Letters*, Vol. 10, pp 87-92
- [6] Bresnahan, T. (1987) 'Competition and Collusion in the American Automobile Market: The 1955 Price War', *Journal of Industrial Economics*, Vol. 35 (4), pp 457-482
- [7] Bresnahan, T.,S. Stern and M. Trajtenberg (1997) 'Market segmentation and the source of rents from innovation:Personal Computers in the late 1980's', *RAND Journal of Economics*, Vol. 28, pp S17-S44
- [8] Case, G. ,A.S. Distefano and B.K. Logan (2000) 'Tabulation of alcohol content of beer and malt beverages', *Journal of Analytical Toxicology*, Vol. 24, pp 202-210
- [9] Chamberlain, G. (1982) 'Multivariate Regression Models for Panel Data', *Journal of Econometrics*, Vol 18 (1), pp 5-42
- [10] Compte, O., F. Jenny and P. Rey (2002) 'Capacity Constraints, Mergers and Collusion', *European Economic Review*, Vol. 46, pp 1-29
- [11] Davis, P. (2006) 'Coordinated Effects Merger Simulation With Linear Demands', Competition Commission
- [12] Davis,P. and Huse, C. (2009) 'Estimating the Coordinated Effects of mergers', in progress, LSE

- [13] Deaton and Muellbauer (1980) 'An Almost Ideal Demand System', *American Economic Review*, Vol. 70, pp 312-336
- [14] Friedman, J.W. (1971) 'A Noncooperative Equilibrium for Supergames', *Review of Economic Studies*, Vol. 38, pp 1-12
- [15] Gorman, T. (1971) 'Two Stage Budgeting', mimeo University of North Carolina
- [16] Hausman, J., G. Leonard and D. Zona (1994) 'Competitive Analysis with Differentiated Products', *Annales d' Economies et de Statistiques*, Vol. 34, pp 159-180
- [17] Kovacic, W. R. Marshall, L. Marx and S. Schulenberg (2008) 'Coordinated Effects in Merger Review: Quantifying the Payoffs from Collusion', mimeo Duke University
- [18] Kuhn, K.-U.(2002), 'Technology choice and capital structure under rate regulation: Rejoinder', *International Journal of Industrial Organization*, Vol. 20(2), pp 283-284
- [19] Kuhn, K-U (2004) 'The Coordinated Effects of Mergers in Differentiated Product Markets', CEPR Discussion paper No. 4769
- [20] Lancaster, K. (1971) 'Consumer Demand: A New Approach', New York: Columbia University Press.
- [21] Mcfadden, D. (1974) 'Conditional Logit Analysis of Qualitative Choice behavior ' in *Frontiers of Econometrics*
- [22] Nevo, A. (1998) 'Identification of the Oligopoly Solution Concept in a Differentiated Products Industry', *Economics Letters*, Vol. 59 (3), pp 391-395
- [23] Nevo, A. (2000) 'A Practitioner's Guide to the Estimation of Random Coefficients Logit Models of Demand', *Journal of Economics and Management Strategy*, Vol. 9 (4), pp 513-548
- [24] Nevo, A. (2001) 'Measuring Market Power in the Ready to Eat Cereal Industry', *Econometrica*, Vol. 69, pp 307-342
- [25] Pinkse, J. , M. Slade and M. Brett (2002) 'Spatial Price Competition: A semiparametric Approach', *Econometrica*, Vol. 70, pp 1111-1155

- [26] Pinkse, J. and M. Slade (2004) 'Mergers, brand Competition and the price of a pint', *European Economic Review*, Vol. 48 (3), pp 617-643
- [27] Sabbatini, P. (2004) 'How to Simulate the Coordinated Effect of a Merger', Working paper, Italian Antitrust Agency, November 2004
- [28] Scherer, F. (1996) 'Industry Structure, Strategy, and Public Policy', HarperCollins
- [29] Tirole, J. (1988) 'The Theory of Industrial Organization', MIT Press, Cambridge
- [30] Train, K. (2003) 'Discrete Choice Methods with Simulation', Cambridge University Press
- [31] Tremblay, V. (2005) 'The U.S. Brewing Industry' MIT Press, Cambridge
- [32] Rojas, C. (2008) 'Demand for Differentiated Products: price and advertising evidence from the US beer market', *International Journal of Industrial Organization*, Vol. 26,pp 288-307
- [33] Rasmussen, E. (2007) 'The BLP Method of Demand Curve Estimation in Industrial Organization', mimeo University of Tokio
- [34] Werden, G. and L. Froeb(1994) 'The Effects of Mergers in Differentiated products Industries: Logit Demand and Structural Merger Policy' Papers 93-4, U.S. Department of Justice - Antitrust Division.

Appendix

Solving for δ_{jt} in the Logit model:

Setting equation 4.5 equal to the observed market share $s_{jt} = \frac{\exp(x_{jt}\beta - \alpha p_{jt} + \xi_{jt})}{1 + \sum_{k=1}^J \exp(x_{kt}\beta - \alpha p_{kt} + \xi_{kt})} =$

S_{jt} ; taking natural logarithms yields $(x_{jt}\beta - \alpha p_{jt} + \xi_{jt}) - \ln(1 + \sum_{k=1}^J \exp(x_{kt}\beta - \alpha p_{kt} + \xi_{kt})) = \ln S_{jt}$. Subtracting the log of outside market share $\ln(S_{0t}) = \ln(\exp(0)) - \ln(1 + \sum_{k=1}^J \exp(x_{kt}\beta - \alpha p_{kt} + \xi_{kt}))$ results in $\delta_{jt} = (x_{jt}\beta - \alpha p_{jt} + \xi_{jt}) = \ln S_{jt} - \ln S_{0t}$.

Table A
All cities in the IRI data set with start and finish date

City	State	City	State
Albany	New York	Louisville	Kentucky
Albuquerque	New Mexico	Memphis	Tennessee
Atlanta	Georgia	Miami/Ft. Lauderdale	Florida
Baltimore	Maryland	Milwaukee	Wisconsin
Birmingham	Alabama	Nashville	Tennessee
Boise	Wisconsin	New Orleans	Louisiana
Boston	Massachusetts	New York	New York
Buffalo/Rochester	New York	Oklahoma City	Oklahoma
Charleston	Florida	Omaha	Nebraska
Charlotte	North Carolina	Orlando	Florida
Chicago	Illinois	Peoria	Illinois
Cincinnati/Dayton	Ohio	Phoenix/Tucson	Arizona
Cleveland	Ohio	Portland	Maine
Columbus	Ohio	Portland	Oregon
Dallas/Ft Worth	Texas	Raleigh/Greensborough	North Carolina
Denver	Colorado	Richmond/Norfolk	Virginia
Des Moines	Iowa	Roanoke	Virginia
Detroit	Michigan	Sacramento	California
El Paso	Texas	Salt Lake City	Utah
Grand Rapids	Michigan	San Antonio	Texas
Green Bay	Wisconsin	San Diego	California
Hartford/Springfield	Connecticut	San Francisco	California
Houston	Texas	Seattle	Washington
Indianapolis	Indiana	Shreveport	Louisiana
Jacksonville	Texas	St. Louis	Missouri
Kansas City	Kansas	Syracuse	New York
Knoxville	Tennessee	Tampa	Florida
Little Rock	Kansas	Toledo	Ohio
Los Angeles	California	Wichita	Kansas

Table B
Products and product characteristics, Average over all markets

Brand	Prices		Alcohol		Coverage		Size	Brand	Prices		Alcohol		Coverage		Size
AB BUDWEISER	1.60	4.90	0.96	3.32	GE KOCHS	1.08	3.90	0.49	2.78						
AB BUDWEISER DRY	1.67	5.00	0.89	3.49	GM CORONA	2.65	4.60	0.76	2.04						
AB BUDWEISER LIGHT	1.60	4.20	0.95	3.42	G O GOYA	1.46	0.50	0.35	1.88						
AB BUSCH	1.27	4.60	0.91	3.49	H HEINEKEN	2.78	5.01	0.82	2.23						
AB BUSCH LIGHT	1.33	4.20	0.82	3.84	LB LABATT	2.16	5.01	0.60	2.28						
AB MICHELOB	1.84	5.00	0.94	2.35	LB LABATTS BLUE	1.76	5.01	0.45	4.49						
AB MICHELOB DRY	1.87	4.90	0.84	2.19	LB ROLLING ROCK	1.80	4.50	0.59	2.18						
AB MICHELOB GOLDEN	1.74	4.70	0.87	3.32	M M OLSON	1.91	5.01	0.59	3.52						
AB MICHELOB LIGHT	1.85	4.30	0.92	2.37	M M OLSON GOLDEN	2.23	5.01	0.66	2.23						
AB NATURAL LIGHT	1.17	4.20	0.73	3.54	M OLD VIENNA	1.59	4.77	0.70	4.13						
ABODOULS	1.72	0.40	0.79	2.09	P FALSTAFF	0.99	4.77	0.29	2.97						
ACCOORS	1.56	5.00	0.93	3.31	P HAMMS	1.02	4.70	0.54	3.59						
ACCOORS EXTRA GOLL	1.55	5.00	0.75	3.34	P HAMMS LIGHT	1.01	4.10	0.43	3.92						
ACCOORS LIGHT	1.57	4.20	0.95	3.47	P OLYMPIA	1.01	4.80	0.59	3.80						
AC KEYSTONE	1.22	4.80	0.72	3.60	P PABST BLUE RIBBON	1.16	5.01	0.72	3.50						
AC KEYSTONE LIGHT	1.24	4.20	0.76	3.71	P RED WHITE & BLUE	1.02	5.15	0.33	3.67						
GH BLACK LABEL	1.04	4.68	0.58	3.03	PM GENUINE DRAFT	1.62	5.01	0.95	2.91						
GH BLATZ	0.98	4.89	0.47	3.49	PM MEISTER BRAU	1.02	4.50	0.58	3.68						
GH HEIDELBERG	0.97	5.00	0.50	3.46	PM MEISTER BRAU LIGI	1.01	4.50	0.35	4.09						
GH HENRY WEINHARD	1.77	5.25	0.90	2.63	PM MILLER GENUINE D	1.69	4.50	0.87	3.19						
GH HENRY WEINHARD	1.81	4.46	0.74	2.28	PM MILLER HIGH LIFE	1.55	5.01	0.95	3.21						
GH KINGSBURY	1.34	0.45	0.52	2.08	PM MILLER LITE	1.57	4.50	0.95	3.49						
GH LONE STAR	1.21	4.70	0.60	3.16	PM MILWAUKEES BEST	1.01	4.51	0.89	3.43						
GH LONE STAR LIGHT	1.62	3.49	0.50	2.49	S GOEBEL	0.97	4.30	0.47	3.30						
GH OLD STYLE	1.18	5.01	0.54	3.37	SOLD MILWAUKEE	1.11	4.90	0.85	3.54						
GH OLD STYLE LIGHT	1.13	4.10	0.52	3.89	SOLD MILWAUKEE LIGI	1.16	3.90	0.70	4.07						
GH RAINIER	1.53	4.91	0.64	2.74	SPIELS	1.03	4.84	0.75	3.02						
GH SCHMIDTS	1.12	4.40	0.38	2.77	SSCHAEFER	1.00	4.40	0.73	3.45						
GH STERLING	1.16	4.77	0.54	2.60	SSCHLITZ	1.23	4.60	0.62	2.34						
GH WEIDEMANN	1.04	4.77	0.62	2.50	S STROHS	1.33	4.60	0.64	4.00						
GH WHITE STAG	0.96	5.20	0.38	3.30	MA MATTS	1.22	4.77	0.54	5.09						
GE GENESEE	1.40	5.01	0.58	2.72	MA UTICA CLUB	1.00	5.02	0.56	3.27						

Table C
Marginal cost (MC), Price Cost margins (PCM) and own price elasticities (η)

Brand	MC	PCM	η	Brand	MC	PCM	η
AB BUDWEISER	0.94	41.05%	-3.960	GE KOCHS	0.88	18.13%	-5.999
AB BUDWEISER DRY	1.27	23.58%	-6.688	GM CORONA	2.38	10.22%	-9.416
AB BUDWEISER LIGHT	1.14	28.88%	-5.121	GO GOYA	1.22	16.94%	-6.682
AB BUSCH	0.82	35.26%	-5.802	H HEINEKEN	2.51	9.64%	-11.081
AB BUSCH LIGHT	0.96	28.04%	-6.540	LB LABATT	1.91	11.34%	-9.049
AB MICHELOB	1.41	23.64%	-5.210	LB LABATTS BLUE	1.42	19.42%	-9.444
AB MICHELOB DRY	1.43	23.61%	-5.876	LB ROLLING ROCK	1.56	13.36%	-8.210
AB MICHELOB GOLDEN I	1.34	23.00%	-6.096	M MOLSON	1.58	17.45%	-8.519
AB MICHELOB LIGHT	1.41	23.61%	-4.983	M MOLSON GOLDEN	1.98	11.24%	-8.266
AB NATURAL LIGHT	0.79	32.08%	-6.069	M OLD VIENNA	1.37	13.63%	-8.479
AB ODOULS	1.35	21.49%	-6.169	P FALSTAFF	0.83	16.18%	-6.010
AC COORS	1.30	16.38%	-6.553	P HAMMS	0.81	20.63%	-6.043
AC COORS EXTRA GOLD	1.34	13.38%	-7.444	P HAMMS LIGHT	0.80	20.94%	-5.772
AC COORS LIGHT	1.27	19.21%	-5.383	P OLYMPIA	0.83	18.15%	-5.747
AC KEYSTONE	1.05	14.34%	-6.755	P PABST BLUE RIBBOI	0.99	15.14%	-6.377
AC KEYSTONE LIGHT	1.04	15.74%	-6.580	P RED WHITE & BLUE	0.80	21.43%	-6.114
GH BLACK LABEL	0.83	19.86%	-6.167	PM GENUINE DRAFT	1.30	19.59%	-5.886
GH BLATZ	0.73	25.11%	-5.899	PM MEISTER BRAU	0.79	22.68%	-5.995
GH HEIDELBERG	0.67	31.53%	-5.012	PM MEISTER BRAU LI	0.72	28.31%	-5.392
GH HENRY WEINHARD A	1.49	15.79%	-6.951	PM MILLER GENUINE	1.37	19.17%	-6.130
GH HENRY WEINHARD P	1.48	18.42%	-7.591	PM MILLER HIGH LIFE	1.28	17.65%	-6.548
GH KINGSBURY	1.07	20.25%	-6.345	PM MILLER LITE	1.20	23.65%	-4.917
GH LONE STAR	0.86	28.84%	-7.125	PM MILWAUKES BES	0.75	25.59%	-5.133
GH LONE STAR LIGHT	1.30	19.36%	-6.380	S GOEBEL	0.72	25.55%	-5.558
GH OLD STYLE	0.84	29.04%	-6.785	S OLD MILWAUKEE	0.93	16.50%	-6.052
GH OLD STYLE LIGHT	0.79	29.76%	-6.539	S OLD MILWAUKEE LI	0.96	17.13%	-6.585
GH RAINIER	0.96	37.23%	-6.624	S PIELS	0.83	19.74%	-5.491
GH SCHMIDTS	0.84	24.60%	-6.436	S SCHAEFFER	0.83	16.93%	-5.918
GH STERLING	0.73	37.08%	-6.663	S SCHLITZ	1.06	14.35%	-6.514
GH WEIDEMANN	0.72	31.12%	-5.410	S STROHS	1.11	16.73%	-7.069
GH WHITE STAG	0.71	26.05%	-5.320	MA MATTS	0.99	19.21%	-6.529
GE GENESEE	1.06	24.63%	-6.959	MA UTICA CLUB	0.79	21.19%	-6.346

Table E
Prices under different modes of competition

Brand	p NE,pre	p NE,post	p cll	Δ NE, pre > cll	p def,pre	p def,post	Δ cll> def, pre
AB BUDWEISER	1.599	1.599	1.839	15.01%	1.583	1.583	-13.91%
AB BUDWEISER DRY	1.665	1.665	1.824	9.57%	1.810	1.810	-0.78%
AB BUDWEISER LIGHT	1.601	1.601	1.984	23.91%	1.804	1.804	-9.09%
AB BUSCH	1.270	-	-	-	-	-	-
AB BUSCH LIGHT	1.327	1.327	1.560	17.56%	1.491	1.491	-4.43%
AB MICHELOB	1.842	1.842	1.822	-1.08%	1.964	1.964	7.76%
AB MICHELOB DRY	1.874	1.874	1.811	-3.40%	1.714	1.714	-5.36%
AB MICHELOB GOLDEN DRAFT	1.741	1.741	1.760	1.07%	1.772	1.772	0.66%
AB MICHELOB LIGHT	1.851	1.851	1.823	-1.49%	1.874	1.874	2.79%
AB NATURAL LIGHT	1.170	1.170	1.406	20.21%	1.189	1.189	-15.43%
AB ODOULS	1.719	-	-	-	-	-	-
AC COORS	1.556	1.556	1.774	13.97%	1.692	1.692	-4.61%
AC COORS EXTRA GOLD	1.548	1.548	1.792	15.75%	1.371	1.371	-23.48%
AC COORS LIGHT	1.566	1.566	1.998	27.61%	1.795	1.795	-10.16%
AC KEYSTONE	1.223	1.223	1.540	25.85%	1.087	1.087	-29.40%
AC KEYSTONE LIGHT	1.237	1.237	1.560	26.06%	1.136	1.136	-27.15%
GH BLACK LABEL	1.036	0.930	1.311	26.55%	0.843	0.868	-35.72%
GH BLATZ	0.976	0.931	1.222	25.21%	0.748	0.772	-38.84%
GH HEIDELBERG	0.972	0.948	1.180	21.36%	0.711	0.731	-39.71%
GH HENRY WEINHARD ALE	1.775	1.800	2.266	27.67%	1.971	1.976	-13.00%
GH HENRY WEINHARD PRIVATE RE	1.813	1.718	2.117	16.81%	1.715	1.731	-18.99%
GH KINGSBURY	1.338	1.240	1.459	9.04%	1.092	1.111	-25.15%
GH LONE STAR	1.205	1.113	1.511	25.38%	0.874	0.906	-42.15%
GH LONE STAR LIGHT	1.616	1.438	1.776	9.89%	1.311	1.335	-26.18%
GH OLD STYLE	1.181	1.097	1.361	15.21%	0.859	0.886	-36.85%
GH OLD STYLE LIGHT	1.129	1.043	1.349	19.42%	0.811	0.838	-39.86%
GH RAINIER	1.526	1.516	1.942	27.28%	0.980	1.029	-49.53%
GH SCHMIDTS	1.119	0.979	1.333	19.16%	0.853	0.880	-36.01%
GH STERLING	1.161	1.021	0.991	-14.71%	0.759	0.772	-23.37%
GH WEIDEMANN	1.044	1.041	1.403	34.41%	0.965	0.977	-31.24%
GH WHITE STAG	0.960	0.840	1.057	10.13%	0.725	0.743	-31.44%
GE GENESEE	1.401	1.401	1.544	10.19%	1.084	1.084	-29.80%
GE KOCHS	1.079	1.079	1.241	14.97%	1.298	1.298	4.64%
GM CORONA	2.650	2.650	2.750	3.79%	2.699	2.699	-1.87%
GO GOYA	1.464	1.464	1.465	0.08%	1.463	1.463	-0.09%
H HEINEKEN	2.780	-	-	-	-	-	-
LB LABATT	2.158	2.158	2.280	5.66%	1.933	1.933	-15.22%
LB LABATTS BLUE	1.757	1.757	1.841	4.75%	1.657	1.657	-9.99%
LB ROLLING ROCK	1.795	1.795	1.962	9.26%	2.026	2.026	3.26%
M MOLSON	1.909	1.909	1.904	-0.29%	1.833	1.833	-3.70%
M MOLSON GOLDEN	2.233	2.233	2.327	4.24%	2.024	2.024	-13.03%
M OLD VIENNA	1.585	1.585	1.800	13.52%	2.833	2.833	57.45%
P FALSTAFF	0.993	0.993	1.269	27.72%	0.870	0.870	-31.41%
P HAMMS	1.021	1.021	1.377	34.95%	0.874	0.874	-36.54%
P HAMMS LIGHT	1.012	1.012	1.244	22.83%	0.915	0.915	-26.43%
P OLYMPIA	1.015	1.015	1.381	36.10%	0.987	0.987	-28.52%
P PABST BLUE RIBBON	1.161	1.161	1.477	27.20%	1.338	1.338	-9.37%
P RED WHITE & BLUE	1.016	1.016	1.143	12.46%	0.917	0.917	-19.80%
PM GENUINE DRAFT	1.620	1.620	1.840	13.58%	1.808	1.808	-1.71%
PM MEISTER BRAU	1.016	1.016	1.197	17.74%	0.887	0.887	-25.91%
PM MEISTER BRAU LIGHT	1.007	1.007	1.160	15.20%	0.803	0.803	-30.82%
PM MILLER GENUINE DRAFT LIGH	1.695	1.695	1.986	17.19%	2.016	2.016	1.53%
PM MILLER HIGH LIFE	1.552	1.552	1.789	15.27%	2.109	2.109	17.89%
PM MILLER LITE	1.569	1.569	2.172	38.45%	1.825	1.825	-16.00%
PM MILWAUKEES BEST	1.011	-	-	-	-	-	-
S GOEBEL	0.965	0.913	1.067	10.52%	1.054	1.258	-1.16%
S OLD MILWAUKEE	1.114	1.210	1.466	31.55%	1.426	1.381	-2.73%
S OLD MILWAUKEE LIGHT	1.160	1.123	1.435	23.72%	1.115	1.192	-22.33%
S PIELS	1.031	1.031	1.229	19.19%	1.200	1.181	-2.37%
S SCHAEFFER	1.004	1.057	1.356	35.08%	1.054	1.173	-22.22%
S SCHLITZ	1.232	1.204	1.617	31.18%	1.100	1.121	-31.98%

Table F
Payoffs under different modes of competition (US \$)

Brand	π NE,pre	π NE,post	π cll	π def,pre	π def,post
AB BUDWEISER	12480717	12487162	14705065	17456473	17456473
AB BUDWEISER DRY	5150717	5149507	7822057	7228701	7228701
AB BUDWEISER LIGHT	38711329	38916181	34329047	37822642	37822642
AB BUSCH	-	-	-	-	-
AB BUSCH LIGHT	3748564	3729518	6883655	6623308	6623308
AB MICHELOB	4201539	4214173	5481155	4933446	4933446
AB MICHELOB DRY	200292	199084	374405	321286	321286
AB MICHELOB GOLDEN DRAFT	1337842	1338023	2086310	1910763	1910763
AB MICHELOB LIGHT	1894541	1903757	2406824	2182641	2182641
AB NATURAL LIGHT	1774293	1765587	3021595	3305630	3305630
AB ODOULS	-	-	-	-	-
AC COORS	2299568	2287608	4924931	4632977	4632977
AC COORS EXTRA GOLD	5664	5629	12938	2551	2551
AC COORS LIGHT	17622113	17664952	21788378	23981268	23981268
AC KEYSTONE	26342	26232	53103	16416	16416
AC KEYSTONE LIGHT	308081	306594	677128	386890	386890
GH BLACK LABEL	14725	13409	15396	3543	9081
GH BLATZ	23213	25486	16114	8251	17330
GH HEIDELBERG	98535	105196	87045	94687	118502
GH HENRY WEINHARD ALE	4718597	5350952	6132616	6753136	6707721
GH HENRY WEINHARD PRIVATE RE	780838	613363	1228479	1115800	1135993
GH KINGSBURY	4906	4720	8303	2113	3392
GH LONE STAR	10486	12122	10869	3908	9460
GH LONE STAR LIGHT	114	100	176	15	55
GH OLD STYLE	27866	30365	41121	11939	23231
GH OLD STYLE LIGHT	12735	14977	12059	5105	10832
GH RAINIER	1278	1293	1282	864	2104
GH SCHMIDTS	2666	3071	2378	557	1811
GH STERLING	4537	7005	9528	4389	5858
GH WEIDEMANN	632710	629118	1042618	1230043	1230364
GH WHITE STAG	7029	8567	6137	2408	4692
GE GENESEE	7160	7118	13131	13353	13353
GE KOCHS	12000	11974	20108	20967	20967
GM CORONA	346	346	586	606	606
GO GOYA	111	112	162	170	170
H HEINEKEN	-	-	-	-	-
LB LABATT	728	722	1613	1953	1953
LB LABATTS BLUE	1212	1207	2151	2099	2099
LB ROLLING ROCK	10368	10291	24140	25075	25075
M MOLSON	2111	2097	4138	3932	3932
M MOLSON GOLDEN	1781	1768	3738	4755	4755
M OLD VIENNA	183747	182758	430940	467863	467863
P FALSTAFF	1399	1399	862	841	841
P HAMMS	15818	15811	12375	14643	14643
P HAMMS LIGHT	6181	6186	4604	7122	7122
P OLYMPIA	63004	62928	68012	88580	88580
P PABST BLUE RIBBON	197187	196164	444190	459306	459306
P RED WHITE & BLUE	1219	1219	896	1426	1426
PM GENUINE DRAFT	14104322	14083294	23429570	21766429	21766429
PM MEISTER BRAU	40743	40708	47654	47966	47966
PM MEISTER BRAU LIGHT	4169	4172	2891	5236	5236
PM MILLER GENUINE DRAFT LIGH	6835652	6882270	7161557	6434100	6434100
PM MILLER HIGH LIFE	5314982	5292842	10514328	7782643	7782643
PM MILLER LITE	35400756	35464528	33369504	43237485	43237485
PM MILWAUKEES BEST	-	-	-	-	-
S GOEBEL	13134	14364	13708	13914	8530
S OLD MILWAUKEE	1327433	1490116	2974047	3004376	2961795
S OLD MILWAUKEE LIGHT	91689	87006	150342	141568	155969
S PIELS	1629157	1619077	3735049	4383685	3563727
S SCHAEFER	225056	231879	355576	397322	389415
S SCHLITZ	25434	23359	55546	18917	25219
S STROHS	3011	2470	4503	1502	2144
MA MATTS	6850	6849	6000	6027	6027

Table G
Discount rates for all firms

Brewer	δ pre	δ post	change (%)
Anheuser Busch	0.381	0.387	1.7%
Adolph Coors	0.179	0.179	0.3%
G. Heileman	0.215	0.079	-63.1%
Genessee	0.071	0.071	-0.4%
Grupo Modelo	0.040	0.040	0.0%
Goya	0.050	0.050	0.0%
Heineken	-	-	-
Labatt	0.073	0.073	0.3%
Molson	0.131	0.130	-0.4%
Pabst	0.141	0.141	0.3%
Phillip Morris	0.270	0.271	0.4%
Stroh	0.145	0.079	-45.1%
Matts FX	0.089	0.089	0.6%

FOOD MARKETING POLICY CENTER RESEARCH REPORT SERIES

This series includes final reports for contract research conducted by Policy Center Staff. The series also contains research direction and policy analysis papers. Some of these reports have been commissioned by the Center and are authored by especially qualified individuals from other institutions. (A list of previous reports in the series is available on our web site.) Other publications distributed by the Policy Center are the Working Paper Series, Journal Reprint Series for Regional Research Project NE-165: *Private Strategies, Public Policies, and Food System Performance*, and the Food Marketing Issue Paper Series. Food Marketing Policy Center staff contribute to these series. Individuals may receive a list of publications in these series and paper copies of older Research Reports are available for \$20.00 each, \$5.00 for students. Call or mail your request at the number or address below. Please make all checks payable to the University of Connecticut. Research Reports can be downloaded free of charge from our web site given below.

Food Marketing Policy Center
1376 Storrs Road, Unit 4021
University of Connecticut
Storrs, CT 06269-4021

Tel: (860) 486-1927
FAX: (860) 486-2461
email: fmpc@uconn.edu
<http://www.fmpc.uconn.edu>