Assessing Market Power and Market Dominance in UK Brewing

Margaret E. Slade[1]

Department of Economics
University of British Columbia
Vancouver, BC V6T 1Z1
Canada
Email: slade@econ.ubc.ca

January 2002

Abstract:

Market power and market dominance are examined in UK brewing, an industry that has witnessed a number of recent mergers and has been scrutinized by both UK and EU authorities. I contrast the North American approach to evaluating mergers and market power, which tends to be based on *unilateral effects*, with the European approach, which is based on dominance or *coordinated effects*, and show how one can distinguish between the two econometrically. The application makes use of two classes of demand equations: the nested logit of McFadden (1978a) and the distance–metric of Pinkse, Slade, and Brett (1998). The two equations yield very different predictions concerning elasticities and markups. Nevertheless, neither model uncovers evidence of collusion (coordinated effects). Using the distance–metric, I find that i) UK brewing firms have substantial market power, but that it is due entirely to unilateral effects, and ii) brand not firm characeristics determine margins.

Journal of Economic Literature classification numbers: L13, L41, L66, L81

Keywords: Market power, unilateral effects, coordinated effects, joint dominance, beer, mergers, differentiated products, multiproduct firms

[1] This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada. I would like to thank Joris Pinkse for collaboration on related papers and for suggestions for this one. I would also like to thank the following people for helpful comments on earlier versions of the paper: Steven Berry, David Genesove, Lars Mathiesen, Mark McCabe and participants at departmental seminars at the University of British Columbia and the Helsinki School of Economics and an NBER workshop on Industrial Organization in Palo Alto, CA.

1 Introduction

Industrial economists are frequently asked to assess the extent of market power that firms in an industry possess. For example, when two or more firms propose a merger, competition authorities must decide if that merger will lead to unacceptable increases in pricing power and thus prices. Furthermore, it is becoming increasingly common to supplement traditional merger analysis with econometric predictions of pre and post–merger industry performance. European Union (EU) and North American (NA) authorities, however, take very different approaches to the evaluation of market power.

In North America, merger policy tends to be based on the notion of *unilateral effects*. In other words, authorities attempt to determine if firms in an industry have market power and how a merger will affect that power, assuming that the firms act in an uncoordinated fashion (see, e.g., Shapiro 1996). In practice, this change is often evaluated as a move from one static Nash equilibrium to a second equilibrium with fewer players.[2]

European authorities, in contrast, tend to base their policy on the notion of dominance. In other words, they seek to determine if a single firm or group of firms occupies a dominant position and if the merger will strengthen that position. Traditionally, single–firm dominance was emphasized. However, the notion of joint dominance has assumed increasing importance due to high–profile merger cases such as Nestle/Perrier, Gencor/Lonrho, and Airtours. Joint dominance is usually taken to mean tacit collusion or *coordinated effects*.[3] In contrast to the evalution of unilateral effects, where econometric evidence has been introduced in court, the econometric evaluation of coordinated effects has received little attention.

In this paper, I propose an econometric technique that can be used to evaluate price/cost margins and to decompose those margins into unilateral and coordinated effects. The technique is based on the decomposition that is developed in Nevo (2001). The organization of the paper is as follows.

The next section compares the EU and North American approaches to evaluating abuse–of–dominance and merger cases. In paticular, it discusses how rules of thumb based on unilateral and coordinated effects can lead to different choices of mergers to investigate. In addition, if a merger is investigated, the two approaches can lead to different conclusions concerning competitive harm.

[2] See, e.g., Werden and Froeb (1994), Hausman, Leonard, and Zona (1994), Nevo (2000), Ivaldi and Verboven (2000), Jayaratne and Shapiro (2000), and Pinkse and Slade (2001).

[3] See, e.g., Lexicon (1999), Kuhn (2000), Kuhn and Motta (2001), and Compte, Jenny, and Rey (2002). The Airtours decision, however, has made this interpretation of joint dominance less clear.

Section 3 describes the UK brewing industry. This industry has witnessed a number of recent mergers of large firms and has been scrutinized by both UK and EU authorities. It is characterized by moderately high margins (approximately 30%), a relatively large number of producers (about 60), a much larger number of brands (many hundreds), and moderate to high horizontal concentration (Hirshman/Herfindahl index approximately 1800). Furthermore, the product — beer — is differentiated along several dimensions. For example, brands can be grouped into discrete classes, such as lagers, ales, and stouts, and they can be measured along continuous dimensions, such as alcohol content. Finally, in recent years, both the structure of the industry and consumers' demand for product characteristics have witnessed dramatic changes.

Section 4 discusses the demand side of the market. I use two classes of simple models in the analysis. Simple models are emphasized because time is an important factor in competition–policy cases. Indeed, since substantial efficiencies can accompany changes in market structure, authorities would like to decide cases in a matter of months, not years. In addition, if econometric evidence will be presented in court, it must be transparent, easy to interpret, and reproducible using standard software packages. The first and more familiar class of demand model, which includes the logit and nested logit of McFadden (1974 and 1978a), has been used extensively by economists to evaluate mergers.[4] With the logit class, own and cross–price elasticities depend on a brand's market and submarket shares. The second class, which I call the distance-metric, is developed in Pinkse, Slade, and Brett (1998) and Pinkse and Slade (2000). With a distance metric, own and cross–price elasticities depend on brand characteristics and a set of measures of the distance between those characteristics.

Section 5 describes how the estimated price/cost margins, which are summary statistics for the degree of market power that the firms possess, can be decomposed into unilateral and coordinated effects. Furthermore, since competition authorities have little control over product characteristics but can influence market structure, the unilateral effect is further decomposed into a portion that is due to differentiation or differences in product attributes and one that is due to concentration or multiproduct production. This decomposition is accomplished by considering pricing games that involve different ownership patterns.

Section 6 deals with estimation. Demand equations and first–order conditions are estimated by a two–step generalized–method–of–moments procedure. Since en-

[4] See, e.g., Werden and Froeb (1994), Werden (1999), Jayaratne and Shapiro (2000), and Ivaldi and Verboven (2000).

dogenous variables appear on the right–hand-side of those equations, the choice of instruments is discussed and tests of their validity are derived.

Section 7 describes the data. The demand data are a panel of brands of draft beers that constitute at least one half of one percent of a regional market. The panel includes 63 brands that are sold in two regions of the country (Greater London and Anglia) in two bimonthly time periods (Aug/Sept and Oct/Nov 1995) and in two types of public houses (multiples and independents). Marginal cost data come from the UK Monopolies and Mergers Commission, who performed a detailed study of brewing, wholesaling, and retailing costs.

Section 8 presents the empirical results, and section 9 concludes. To anticipate, I find that, although firms in the industry have substantial market power, it is due entirely to unilateral effects. In particular, there is no dominant group. In addition, I find that whereas brand characteristics are important determinants of margins, firm charactersitics are not.

2 The EU and North American Approaches

Competition authorities are responsible for policing violations of many sorts of antitrust laws. However, I limit attention to abuse–of–dominance and merger cases. With those cases, the common goal of EU and NA authorities is to evaluate market power and how certain practices or acts contribute to that power. The approaches that they take to pursuing that goal, however, can differ.

With abuse–of–dominance cases, authorities on both sides of the Atlantic must determine whether a single firm or group of firms occupies a dominant position before they can detemine whether that firm or group has abused its position. When a single firm is involved, although market share alone does not determine dominance, a firm's share must typically exceed 40% before it is considered dominant.[5] With joint dominance, in contrast, often no single firm has such a large market share, but some group of firms has a joint share that is large (perhaps in excess of 60%). It is clear, however, that regardless of industry concentration, *some* group of firms will have a large share of *any* market. For this reason, it must be demonstrated that the group behaves in a coordinated fashion to control the market. In other words, unilateral effects are not sufficient to establish joint dominance, and coordinated effects assume primary importance. It is therefore useful to have econometric techniques that can

[5] These numbers are not exact but are only indicative of common practice.

be used to distinguish between the two.

With merger cases, authorities must have some screening process to determine which cases will be investigated. Typically this involves assessing whether market power exists before determining whether the merger will increase that power. In performing that exercise, North American authorities have traditionally relied heavily on measures of industry concentration such as the Hirschman/Herfindahl index (HHI), whereas EU authorities, as well as many national competition bureaus within the EU, have relied more heavily on the notion of dominance. Unfortunately, these two approaches can lead to different choices of which mergers to investigate. For example, in the application below, I estimate that, using either the HHI or the price/cost margin, UK brewing firms have substantial market power. However, that power is due entirely to unilateral effects. This means that North American guidelines would indicate that mergers in the industry should be closely scrutinized whereas European guidelines would not.

The factors that are considered in screening mergers are indicative of the different approaches. When unilateral effects are the focus, market shares and entry barriers assume primary importance. When coordinated effects are the focus, in contrast, a number of additional factors are considered, including product homogeneity, stable and symmetric market shares, stagnant and inelastic demand, similar costs, and low levels of technical change. Those factors tend to discount mergers in consumer–product and high–tech industries.

If authorities decide to investigate a merger, they must perform a more formal assessment of whether that merger will increase market power. As with premerger evaluation, the two approaches can lead to very different conclusions concerning post–merger performance. To illustrate, suppose that two firms dominate a market and that the market shares of those firms become more unequal when, for example, the dominant firm acquires a smaller firm. The acquisition will cause the HHI to rise, and one might conclude that increased concentration will lead to higher prices. On the other hand, the conventional wisdom holds that asymmetric market shares are detrimental to collusion because it is harder to reach an agreement when firms are less similar. Moreover, Compte, Jenny, and Rey (2002) show formally that, in a repeated price game with capacity constraints, as market shares become more asymmetric, tacit collusion becomes more difficult to sustain. This means that mergers that increase unilateral effects can reduce coordinated effects.

Unfortunately, the predictions of game–theoretic models of how changes in market structure affect incentives to collude are very fragile. This means that, whereas it is easy to update concentration indices to reflect post–merger shares, it is difficult to predict how a merger will affect the firms' ability to collude.

Nevertheless, although proxies for market power such as the HHI that are based on market shares can be informative signals of performance in industries where products are homogenous,[6] when products are differentiated, the issue is more complex. For this reason, many recent studies of the effects of mergers bypass the process of defining a market and calculating shares within that market and rely instead on merger simulations. Those simulations, which are based on calibrated or estimated models of demand and cost, assess how prices and margins vary as the number of players in a static game changes, where players control some subset of industry physical assets (e.g., plants) or brands of a differentiated product. In other words, they assess changes in unilateral effects.

Given that it is relatively straight forward to use quantitative techniques to predict merger–related changes in unilateral effects but difficult to perdict changes in coordinated effects, I limit my analysis to techniques that can be used for premerger evaluation or for establishing joint dominance in abuse–of–dominance cases. The more difficult task of comparing post–merger predictions of industry performance across approaches is left to the future. In particular, there is a need for more robust theoretical models before the predictions of those theories can be quantified.

3 The UK Brewing Industry

The UK brewing industry is interesting for a number of reasons. In particular, it has recently undergone rapid change with respect to consumer tastes, product offerings, and market structure. In addition, both its horizontal and vertical organization have been subjected to numerous reviews by several levels of government.

Historically, the UK brewing industry developed in a very different fashion from those in, for example, the US, Canada, and France, which were dominated by a few large brewers that sold rather homogeneous national brands of lagers. Indeed, the UK industry, which was relatively unconcentrated, produced a large variety of ales, and regional variation in product offerings was substantial. Moreover, national advertising played a less important role than in many countries. In the last decade,

[6] To illustrate, when firms are engaged in a symmetric Cournot game, equilibrium market shares and margins move continuously from monopoly to perfect competition as the number of players increases. Furthermore, with a fixed number of asymmetric players, firms' price/cost margins are directly related to their market shares.

however, there has been a succession of mergers that have increased concentration in brewing and have caused the industry to move towards a more North–American style. Nevertheless, UK brewing is still less concentrated than its counterparts in the US, Canada, and France, where beer tends to be mass produced. It is substantially more concentrated, however, than its counterpart in Germany, where specialty beers predominate.

Among Western countries, the UK is not an outlier with respect to consumption of beer per head or the fraction of sales that are imported. It is very different, however, with respect to the ratio of draft to total beer sales. Indeed, draft sales in the UK, which in 1995 were just under 70% of total sales, accounted for almost three times the comparable percentages in France and Germany and about six times the percentages in North America.[7]

Substantial changes in both consumption and production have occurred in the industry in the last few decades. To illustrate, beers can be divided into three broad categories: ales, stouts, and lagers. Although UK consumers traditionally preferred ales to lagers, the consumption of lager has increased at a rapid pace. Indeed, from less than 1% of the market in 1960, lager became the dominant drink in 1990, when it began to sell more than ale and stout combined. Most UK lagers bear the names of familiar non–British beers such as Budweiser, Fosters, and Stella Artois. Almost all, however, are brewed under license in the UK and are therefore not considered to be imports.

A second important aspect of beer consumption is the popularity of 'real' or cask–conditioned ale. Real products are alive and undergo a second fermentation in the cask, whereas keg and tank products are sterilized. Although real products' share of the ale market has increased, as a percentage of the total beer market, which includes lager, they have lost ground.

A final trend in consumption is the rise in popularity of premium beers, which are defined as brands with alcohol contents in excess of 4.2%. Traditional ales are of lower strength than stouts and lagers. In addition, keg products tend to contain less alcohol than real products. Many of the more recently introduced brands, however, particularly the lagers and hybrid ales,[8] are premium beers with relatively high alcohol contents.

With respect to production, the number of brewers has declined steadily. Indeed,

[7] Only in Ireland was it higher, where draft sales accounted for over 80% of consumption.

[8] A hybrid is a keg ale that uses a nitrogen and carbon–dioxide mix in dispensing that causes it to be smoother and to more closely resemble a cask ale.

in 1900, there were nearly 1,500 brewery companies, but this number fell dramatically and is currently around sixty. In addition to incorporated brewers, however, there are approximately 200 microbreweries operating at very small scales. In spite of increases in industry concentration, most brewers are still small, and few produce products that account for more then 0.5% of local markets.

In the early 1990s, mergers reduced the number of national brewers from six to four. In addition, in 1997, a very large merger was proposed that involved the numbers two and three brewers, Bass and Carlsberg-Tetley, and would have created a new firm with 37% of the market. The UK Monopolies and Mergers Commission (MMC) estimated that, after the merger, the HHI would increase by about 650 points. Nevertheless, the MMC recommended that the firms be allowed to merge. The merger did not take place, however, because the president of the Board of Trade did not accept the MMC's advice. Still more recently, in 2000, the world's largest brewer, the Belgian firm Interbrew, acquired Whitbread, the fourth national brewer, and later in the same year, it acquired the brewing assets of Bass. This time the MMC did not approve the merger. Instead, it recommended that Interbrew be required to divest the UK business of Bass.

This snapshot of the UK beer industry shows significant changes in tastes and consumption habits as well as a decline in the number of companies that cater to those tastes. Nevertheless, there is still considerable variety in brand offerings and brand characteristics. Brewer market power could therefore result from fewness, differentiation, collusion, or from a combination of the three. To disentangle these effects, we turn to the econometric model.

4 Demand Models

Firms can possess market power because they have few competitors and thus operate in concentrated markets. Even when there are many producers of similar items, however, they can possess market power if their products have unique features that cause rival products to be poor substitutes. To evaluate power in markets where products are differentiated, it is therefore important to have good estimates of substitutability.

When a product is homogeneous, a single price prevails in the market. There is therefore just one price elasticity of demand — the own-price elasticity — to estimate. When products are differentiated, in contrast, the number of brands can be very large, often several hundred, and the number of price elasticities is formidable. One must therefore place some structure on the estimation.

A number of demand specifications have been developed recently to deal with the problem of an abundance of elasticities. I use two relatively simple ones here: the multinominal nested logit (MNL) and the distance metric (DM). Since the first is well known and the second is developed elsewhere (Pinkse, Slade, and Brett 1998 and Pinkse and Slade 2000), I simply reproduce the estimating equations and the equations for the own and cross–price elastcities of demand.

With both models, there are n brands of a differentiated product, $q = (q_1, \ldots, q_n)^T$ as well as an outside good q_0 that is an aggregate of all other products. In most of the discussion that follows, I assume that there is only one market with exogenous size. It is straight forward, however, to extend the demand models to encompass multiple markets, in which case the size of each market is an endogenous function of regional variables.

4.1 The Nested Logit

The MNL demand equation is based on the random–utility model in which an individual consumes one unit of the product that yields the highest utility, where products include the outside good. The MNL is distinguished from the ordinary logit by the fact that the n brands or products are partitioned into G groups, indexed by $g = 1, ..., G$, and the outside good is placed in group 0. The partition is chosen so that like products are in the same group. For example, when the differentiated product is beer, the groups might be lager, ale, and stout.

The MNL estimating equation is

$$ln(s_i) - ln(s_0) = \beta^T x_i - \alpha_i p_i + \sigma ln(\bar{s}_{i/g}) + \xi_i, \tag{1}$$

where s_i is product i's volume share of the entire market, x_i is a vector of observed characteristics of that product, p_i its price, $\bar{s}_{i/g}$ is its share of the group g to which it belongs, and ξ_i is an unobserved (by the econometrician) product characteristic. The parameter $\sigma(0 \leq \sigma \leq 1)$ measures the within-group correlation of utility, and the ordinary logit is obtained by setting σ equal to zero. When $\sigma = 0$, substitution possibilities are completely symmetric (e.g., all products belong to the same group). Finally, ξ_i is assumed to be mean independent of x_i.

Let ε_{ij} denote the price–elasticity of demand, $(\partial q_i/\partial p_j)(p_j/q_i)$. The MNL elasticities are then

$$\varepsilon_{ii} = \alpha_i p_i[s_i - 1/(1-\sigma) + \sigma/(1-\sigma)\bar{s}_{i/g}], \tag{2}$$

$$\varepsilon_{ij} = \begin{cases} \alpha_j p_j [s_j + \sigma/(1-\sigma)\bar{s}_{j/g}] & if \quad j \neq i \quad and \quad j \in g \\ \alpha_j p_j s_j & if \quad j \neq i \quad and \quad j \notin g. \end{cases}$$

Equation (1) is slightly more flexible than the standard MNL. In particular, the coefficient of p_i, α_i, is allowed to depend on the characteristics of that product. In other words, $\alpha_i = \alpha(x_i)$. Nevertheless, as equation (2) shows, the cross–price elasticity between i and j is independent of i. This means that the off–diagonal elements in a column of the elasticity matrix take on at most two values, depending on whether the rival product is in the same or a different group.[9]

4.2 The Distance Metric

Brands of a differentiated product can compete along many dimensions in product–characterisitc space. For empirical tractbility, however, one must limit attention to a small subset of those dimensions. Nevertheless, it is not desirable to exclude possibilities *a priori*. The distance–metric demand model allows the researcher to experiment with and determine the strength of competition along many dimensions. It can thus be used to construct an empirically tractable demand equation that relies on few *a priori* assumptions. In particular, virtually any hypothesis concerning the way in which products compete (any distance measure) can be assessed in the DM framework. However, only the most important measures are typically used in the final specification.

The DM is not a discrete-choice model. Instead, it is assumed that individuals have a systematic taste for diversity and thus might want to consume more than one brand. Furthermore, individuals are allowed to purchase variable amounts of each brand. Finally, all individuals consume the outside good. The DM model is based on a normalized–quadratic utility function (Berndt, Fuss, and Waverman 1977 and McFadden 1978b) in which the prices of the differentiated products as well as individual incomes have been divided (or normalized) by the price of the outside good.[10]

---

[9] A second class of demand equation has been used to model more flexible substitution patterns in antitrust cases. This class involves multi–stage budgeting in which the final stage (brand choice) is modeled as an Almost–Ideal–Demand System (AIDS) (see, Hausman, Leonard, and Zona 1994). That specification, however, is empirically intractable when there is a large number of brands in any group, as is the case here. Unfortunately, the number of brands per group must be determined by economic consideration and cannot be chosen arbitrarily.

[10] This utility function is flexible in prices. In other words, it is a second-order approximation

Since the utility function is quadratic, the demand equations are linear in normalized prices and incomes. One can write aggregate product demands as

$$q_i = a_i + \Sigma_j b_{ij} p_j - \gamma_i y, \tag{3}$$

where $p = (p_1, \ldots, p_n)^T$ and aggregate income y have been divided by p_0.

Equation (3) clearly has more parameters than can be estimated using a single cross section or short panel. It is therefore assumed that a_i and b_{ii}, $i = 1, \ldots n$, are functions of the characteristics of brand i, $a_i = a(x_i)$ and $b_{ii} = b(x_i)$. For example, when the product is beer, the characteristics might be the brand's alcohol content, product type (e.g., lager, ale, or stout), and brewer identity. Furthermore, the off–diagonal elements of B are assumed to be functions of a vector of measures of the distance between brands in some set of metrics, $b_{ij} = g(d_{ij})$. For example, when the product is beer, the measures of distance, or its inverse closeness, might be alcoholic-content proximity and dummy variables that indicate whether the brands belong to the same product type (e.g., whether both are stouts) and whether they are brewed by the same firm.

As with the MNL, since the intercepts depend on product characteristics, the equation is transformed from one in which consumers demand brands into one in which they demand the characteristics that are embodied in those brands. If the number of characteristics is less than the number of brands, the dimensionality of the problem is reduced. The own–price elasticity of demand also depends on the characteristics, the hypothesis being that, for example, the demand for high–alcohol beers might be systematically less elastic than that for low. In contrast to the MNL, however, the off–diagonal elements, $b_{ij}, j \neq i$, that determine substitutability between brands depend on distance measures, the hypothesis being that, for example, brands that have similar alcohol contents might be closer substitutes.

Let Z be the matrix of observed brand and market variables with typical row $z_i = (x_i^T, y^T)^T$.[11] If in addition there are unobserved brand and market variables ξ with typical element ξ_i, (3) can be written in matrix notation as

$$q = Z\beta + Bp + \xi, \tag{4}$$

that places no restrictions on substitution possibilities between brands of the differentiated product. Moreover, the function is in Gorman polar form and can therefore be aggregated to obtain brand-level demands (see Blackorby, Primont, and Russell 1978 for a discussion of the conditions that are required for consistent aggregation across households.) In particular, aggregation does not require one to specify the distribution of unobserved consumer heterogeneity.

[11] y can be a vector of market characteristics.

where β is a vector of parameters that must be estimated. The random variable ξ can be heteroskedastic and spatially correlated. However, as with the MNL, ξ is assumed to be mean independent of the observed characteristics, $\mathrm{E}[\xi_i|X] = 0$.

The own and cross-price elasticities that are implied by equation (4) are

$$\varepsilon_{ii} = \frac{p_i \lambda^T x_i}{q_i} \quad \text{and} \quad \varepsilon_{ij} = \frac{p_j g(d_{ij})}{q_i}. \tag{5}$$

As with the MNL, DM own–price elasticities depend on prices, market shares, and product characterisitcs. However, a feature that distinguishes the DM from the MNL is that, with the former, cross–price elasticities depend on attributes of both brands — i and j — whereas with the latter, they depend only on the characteristics of j. This means that cross–price elasticities can be modeled more flexibly. Indeed, by choosing appropriate distance measures, one can obtain models in which substitution patterns depend on *a priori* product groupings, as with the nested logit. There are, however, many other possibilities. For example, one can also obtain models in which cross-price elasticities depend on continuous distance measures, such as differences in alcohol contents, and models that use common–market-boundary measures, as in Feenstra and Levinsohn (1995). Finally, hybrid models that include more than one distance measure are possible.

5 Evaluating Unilateral and Coordinated Effects

The term market power usually denotes the ability of firms to charge prices in excess of marginal costs. The most common measure of market power is the Lerner index or price/cost margin, $L_i = (p_i - c_i)/p_i$.

If one has exogenous estimates of marginal costs, one can calculate n price/cost margins L_i, one for each brand, and it is possible to decompose those margins into unilateral and coordinated effects. Furthermore, since market structure can to some extent be controlled by competition authorities, whereas the degree of product differentiation cannot, it is useful to further decompose the unilateral effect. There are then three components: one that is due to differentiation, one that is due to market structure or multibrand production, and the third that is due to collusion.[12] Finally, the sum of the first two is the unilateral effect, whereas the third, if positive, is the coordinated effect.

[12] This decomposition is due to Nevo (2000).

This procedure involves solving first–order conditions to obtain equilibrium prices of different games and calculating the associated margins. To illustrate, suppose that there are K sellers of the differentiated product and that player k, $k = 1, \ldots, K$, controls a set of prices p_i with $i \in \tilde{k}$, where $\mathcal{K} = [\tilde{1}, \tilde{2}, \ldots, \tilde{K}]$ is a partition of the integers $1, \ldots, n$. Let $p_{\tilde{k}}$ be the set of prices that k controls. Assume also that sellers of the differentiated product play a game, whereas the outside good is competitively supplied. For given $\mathcal{K}$ and prices p_j with $j \notin \tilde{k}$, player k chooses $p_{\tilde{k}}$ to

$$max_{p_{\tilde{k}}} \ \pi_k = \Sigma_{i \in \tilde{k}} (p_i - c_i) q_i - F_k, \tag{6}$$

where c_i is the constant marginal cost of producing brand i and F_k is the fixed cost for firm k.[13]

The ith first-order condition is

$$q_i + \Sigma_{j \in \tilde{k}_i} (p_j - c_j) \frac{\partial q_j}{\partial p_i} = 0, \tag{7}$$

where $\tilde{k}_i$ is the element of the partition to which p_i belongs.

Equation (7) nests the following games:

i) *Bertrand behavior with single-product firms:* $K = n$.

ii) *Bertrand behavior with multiproduct firms:* $1 < K < n$.

iii) *Joint-profit-maximizing behavior:* $K = 1$.

Given a partition $\mathcal{K}$ and a set of marginal costs, one can solve the first-order conditions (7) for equilibrium prices and margins, $\tilde{p}_{\mathcal{K}i}$ and $\tilde{L}_{\mathcal{K}i} = (\tilde{p}_{\mathcal{K}i} - c_i)/\tilde{p}_{\mathcal{K}i}$, of the corresponding Bertrand game. Moreover, with the DM demand equation, this calculation normally involves only matrix inversion.

The first step in the decomposition is to evaluate the market power that results from differentiation alone. One does this by solving game i). With this game, each element of the partition, $\mathcal{K}$, is a singleton, and there are n Bertrand players or decision makers, one for each brand. The margins that correspond to the equilibrium prices of this game express the market power that is due to differentiation. The implicit comparison here is with marginal–cost pricing or $L = 0$.

The second step is to evaluate the market power that results from concentration, or equivalently, fewness or multibrand ownership. To do this, one solves game ii), where the partition $\mathcal{K}$ with $1 < K < n$ corresponds to the observed brand–ownership pattern. The margins that are associated with this game express the market power

[13] With this specification, economies of scope enter only though the firms' fixed costs.

that is due to a combination of differentiation and fewness. Furthermore, differences in the margins that are associated with the two games measure the additional power that is due to fewness (i.e., to the fact that there are K rather than n firms).

The third step involves estimating the coordinated effect. If one interprets tacit collusion in the game–theoretic sense — as obtaining an outcome that is preferred by the players to the Nash equilibrium of the one–shot game — then the residual market power that has not been explained is the coordinated effect.[14] Let the vector of Lerner indices evaluated at the exogenous cost estimates, $c = (c_1, \ldots, c_n)^T$, and observed prices, p, be L_o, where o stands for observed. Then differences between L_o and the margins of the second game, if positive, can be attributed to collusion. One cannot distinguish, however, between tacit and overt collusion. Furthermore, if collusion is believed to be tacit, one cannot determine the sort of dynamic game that underlies that collusion, at least not using the methods that are described here.

This procedure can be used to test for the existence of and to estimate the average magnitude of coordinated effects, but it does not indicate which firms price less competitively or which brands are less competitively priced. It is possible to investigate this issue by augmenting the first–order condition (7) to include a vector of market–conduct parameters, $\theta_i,\ i = 1, \ldots, n$,

$$q_i + \Sigma_{j\in\tilde{k}_i}\Big\{(p_j - c_j)\Big[\frac{\partial q_j}{\partial p_i} + \theta_i \Sigma_{m\notin\tilde{k}_i}\frac{\partial q_j}{\partial p_m}\Big]\Big\} = 0, \tag{8}$$

One can think of the θs as parameters that measure the extent of the deviation from the null hypothesis of static Nash–equilibrium behavior.[15]

Since there are n first–order conditions (8) and n market–conduct parameters, θ_i, one can solve the first–order conditions to obtain a vector of implicit market–conduct parameters, $\tilde{\theta}$. Alternatively, one can assume that market–conduct is a function of brand and firm characteristics, $\theta_i = f(x_i)$, and estimate equation (8) jointly with the demand equation. I call the econometrically estimated parameter vector $\hat{\theta}$. Both $\tilde{\theta}$ and $\hat{\theta}$ can be used to test hypotheses concerning systematic deviations from static Nash–equilibrium.

[14] Tacit collusion can be due to one or more of many dynamic factors, such as repetition of the one-shot game.

[15] The term 'conjecture' is often used because the parameters are often interpreted as conjectured responses, $\Theta_{ji} = E(\partial p_j/\partial p_i),\ j \notin \tilde{k}_i$. This interpretation, however, is often not useful. Nevertheless, the first–order conditions are obtained by allowing these partial derivatives to be nonzero and then setting $\Theta_{ji} = \theta_i$ for all j.

6 Estimation

6.1 Demand

The demand equations (1) and (3) contain endogenous right–hand–side variables (prices) and are therefore estimated by instrumental-variables (IV) techniques. Estimation of the nested logit is entirely straight forward.

The DM equation (3) can be estimated by either parametric or semiparametric methods. With the parametric estimator used here, $g(\cdot)$ is a parametric function of the distance measures d_{ij}. One must also choose a stochastic specification for ξ. The covariance–matrix estimator that is used, which is nonparametric, is similar to the one that is proposed in Newey and West (1987) in a time-series context. In particular, as discussed in the appendix, observations that are 'close' to one another are assumed to have nonzero covariances, where closeness is measured by one or more of the distance measures. Our estimator, however, which involves correlation in space rather than time, can be used when the errors are nonstationary, as is more apt to be the case in a spatial context.[16]

The issue of identification is complicated by the fact that the Z variables can enter both the linear part of the model, $Z\beta$, and the g function. In particular, it is not immediately obvious that g is identified, even by functional form. However, if the discrete distance measures (such as product groupings) are used in g, but no corresponding product dummies are included in Z, which is the case with the results reported later, g can be identified. In general, this procedure will not work well if price distributions and/or locations in taste space do not vary much across categories. Fortunately, with the application, there is substantial variation in both across product types.[17]

6.2 First-Order Conditions

The first–order conditions (8) contain a vector of market–conduct parameters, θ, that can be modeled as functions of the brand characteristics.[18] In the absence of information on functional form, a simple linear relationship is used,

[16] Stationarity is used here to mean that the joint distribution can depend on locations, not just on distance between locations, and not to denote a unit root.

[17] If the price distributions are different across regions and/or time periods, and g is the same across regions and time periods, g can also be identified.

[18] One can also estimate marginal costs. This was done in an earlier version of this paper (Slade 2001). The principal conclusions do not change, however, and so those specifications are not shown.

$$\theta_i = \gamma^T x_i + \phi_i, \tag{9}$$

where x_i is the vector of observed characteristics, and ϕ_i is an unobserved variable that affects conduct. With the DM specification, the equation that is estimated is $Y_i = \gamma^T x_i + \phi_i$, where

$$Y_i = \frac{-q_i - \Sigma_{j\in\tilde{k}_i}(p_j - c_j)b_{ji}}{\Sigma_{j\in\tilde{k}_i}\{(p_j - c_j)[\Sigma_{m\notin\tilde{k}_i} b_{jm}]\}}. \tag{10}$$

As is standard, ϕ is assumed to be mean independent of x.

I use a two-step generalized-method-of-moments (GMM) procedure to estimate the first–order conditions. In the first step, the parameters of the demand equation are estimated as in the previous subsection. In the second step, the estimated demand parameters and the postulated market–conduct function is substituted into the first–order condition, and the remaining parameters are estimated. The only complication is that the standard errors of the second–stage parameters must be adjusted to reflect the fact that the demand equation was itself estimated. The method that is used to do this, which is described in the appendix, is based on suggestions of Newey (1984) and Murphy and Topel (1985). An advantage of a two–step procedure is that misspecification of the first–order condition does not contaminate the demand estimates, in which one typically has more confidence.

6.3 Hypothesis Tests

Hypotheses concerning the econometrically estimated parameters, $\hat{\theta}$, can be tested using standard techniques. In addition, the implicit variables, $\tilde{\theta}$, which are nonlinear functions of estimated parameters, are themselves random variables that can be the subject of tests. Two methods of testing hypotheses concerning implicit variables are used. The first and simpler of the two is based on the fact that any sequence of i.i.d. variates with uniformly bounded moments greater than two, whether they are estimates or not, have a limiting normal distribution.[19] Unfortunately, the notion that the estimates, $\tilde{\theta}_i$, $i = 1, \ldots, n$, are independent across i, even in the limit, is questionable. If they are dependent, their standard errors will in general be larger and rejection of the null less likely. When the null is not rejected, only this test is used. The second test, which is used when the null is rejected by the first, is a parametric bootstrap. In particular, repeated draws from the estimated joint distribution of

[19] This follows from the Lindberg theorem (e.g., Doob 1953, theorem 4.2).

the parameters are performed, the desired quantity is calculated, and a bootstrap distribution is generated.

6.4 The Choice of Instruments and Tests of Their Validity

An important issue is the choice of instruments. In particular, one needs instruments that vary by brand and market. The exogenous demand and cost variables, Z and c, are obvious candidates, and some of them (e.g., coverage) vary by brand and market. A number of other possible instruments have been discussed in the differentiated–products literature. For example, Hausman, Leonard, and Zona (1994), assume that systematic cost factors are common across regions and that short–run shocks to demand are not correlated with those factors. This allows them to use prices in one city as instruments for prices in another. Berry, Levinsohn, and Pakes (1995), in contrast, point out that, since a given product's price is affected by variations in the characteristics of competing products, one can use rival–product characteristics as instruments.

The identifying assumptions made here involve a combination of the two suggestions. In particular, the first set of instruments exploits the panel natur e of the data. The brands in my sample are not brewed locally and thus have a common cost component.[20] Furthermore, brands that are sold in one region are not substitutes for those that are sold in another. Profit-maximizing decision makers will therefore not coordinate their price choices across regions. For these reasons, I assume that prices in region one are valid instruments for prices in region two and vice versa.

Price in the other region, p_{-r}, can enter the instrument set directly. Moreover, it can be used to construct additional instruments. This is done by premultiplying the price vector by weighting matrices W, where each W is an element of the distance vector, d. To illustrate, suppose that W^1 is the same–product–type matrix (i.e., the matrix whose i, j element is one if brands i and j are the same type of product — both lagers for example — and zero otherwise). The product $W^1 p_{-r}$ has as ith element the average in the other region of the prices of other brands that are of the same type as i.[21]

Unfortunately, there are circumstances under which price instruments, p_{-r} will not be valid. For example, national advertising campaigns could cause the shocks in the two regions to be correlated. Fortunately, national advertising creates less

[20] Many brands are brewed in just one or two national breweries.

[21] The weighting matrices are normalized so that the rows sum to one.

of a problem here than with, for example, US beer, which is much more heavily promoted.[22] Nevertheless, it is advisable to experiment with other instruments.

The second set of instruments exploits the conditional–independence assumption; when $\mathrm{E}[\xi_i|X] = 0$, rival characteristics can be used to form instruments. This is done by premultiplying the vectors of characteristics by weighting matrices W. To illustrate, suppose that x^1 is the vector of alcohol contents of the brands (a column of the matrix X). The product W^1x^1 has as ith element the average alcohol content of rival brands that are of the same type as i.[23]

As with the first set of instruments, there are circumstances under which instruments formed from rival characteristics will not be valid. This would be the case, for example, if rival characteristics entered the demand equation directly, a possibility that can be assessed econometrically.

One check on instrument validity is to determine if the results obtained are sensitive to the set chosen. In other words, since there are two sets of instruments used in the application — those constructed from prices in the other region and those constructed from characteristics of rival products — it is possible to use each set separately as well as the two together and to compare the implied elasticities and markups.

It is also important to assess the validity of the instruments (i.e., that they are uncorrelated with the errors in the estimating equations) more directly. In particular, the exogeneity of price in the other region is questionable. Moreover, many other instruments are created from that variable and thus might also be suspect. A formal test of exogeneity, one that is valid in the presence of heteroskedasticity and spatial correlation of an unknown form, is derived in the appendix.[24] Intuitively, the test involves assessing correlation between instruments and residuals, taking into account the fact that the residuals are not errors but are estimates of errors.

[22] Figures taken from the MMC cost study indicate that advertising and marketing expenditures are less than one percent of sales. Moreover, variations in advertising by firm (but not by brand) are captured by firm fixed effects.

[23] The use of rival characteristics is somewhat different here from their use in much of the differentiated–products literature (e.g., Berry, Levinsohn, and Pakes 1995), where rival characteristics are used as instruments for own price. Here they are principally used as instruments for rival prices.

[24] The test and discussion appear in Pinkse and Slade (2000) and are reproduced here.

7 The Data

7.1 Demand Data

The data are a panel of brands of draft beer sold in different markets, where a market is a time-period/regional pair. The panel also includes two types of establishments. Brands that are sold in different markets are assumed not to compete, whereas brands that are sold in the same market but in different types of establishments are assumed to compete.

Most of the demand data were collected by StatsMR, a subsidiary of A.C. Nielsen Company. An observation is a brand of beer sold in a type of establishment, a region of the country, and a time period. Brands are included in the sample if they accounted for at least one half of one percent of one of the markets. There are 63 brands. Two types of establishments are considered, multiples and independents, two regions of the country, London and Anglia, and two bimonthly time periods, Aug/Sept and Oct/Nov 1995. There are therefore potentially 504 observations. Some brands, however, were not sold in a particular region, time period, and type of establishment. When this occurred the corresponding observation was dropped in both regions of the country.[25] This procedure reduced the sample to 444 observations.

Establishments are divided into two types. Multiples are public houses that either belong to an organization (a brewer or a chain) that operates 50 or more public houses or to estates with less than 50 houses that are operated by a brewer. Most of these houses operate under exclusive–purchasing agreements (ties) that limit sales to the brands of their affiliated brewer.[26] Independents, in contrast, can be public houses, clubs, or bars in hotels, theaters, cinemas, or restaurants. Independent establishments are usually not tied to a particular brewer.

For each observation, there is a price, sales volume, and coverage. Price, which is measured in pence per pint, is the average for that brand, type of establishment, region, and time period. This variable is denoted PRICE. Volume, which is measured in 100 barrels, is total sales of the brand in the region, time period, and type of establishment. This variable is denoted VOL. Finally, coverage, which is the percentage of outlets in the region, time period, and type of establishment that stocked the brand, is denoted COV.

[25] Dropping an observation in both regions of the country is necessary because prices in one region are used as instruments for prices in the other.

[26] Many tied houses also sell brands that are brewed by firms that do not have tied estates (e.g., Guiness) as well as a 'guest' cask–conditioned ale.

VOL is the dependent variable in the distance–metric demand equation. With the nested–logit specifications, in contrast, the dependent variable is LSHARE — the natural logarithm of the brand's overall market share — where the market includes the outside good.[27]

In addition, there are data that vary by brand but not by region, establishment type, or time period. These variables are product type, brewer identity, and alcohol content.

Brands are classified into four product types, lagers, stouts, keg ales, and real ales. Two types of ales are distinguished because real or cask–conditioned ales undergo a second fermentation in the cask, whereas keg ales are sterilized. Unfortunately, three brands — Tetley, Boddingtons, and John Smiths — also have keg–delivered variants. Since it is not possible to obtain separate data on the two variants of these brands, the classification that is used by StatsMR was adopted. Dummy variables that distinguish the four product types are denoted $\mathrm{PROD}_i,\ i = 1, \ldots, 4$. These product types also form the basis of the groups for the MNL specifications, and those specifications include an explanatory variable LGRSHARE, the natural logarithm of the brand's share of the group to which it belongs. Finally, as explained below, product types also play a role in determining one of the metrics or distance measures for the DM specifications.

There are ten brewers in the sample, the four nationals, Bass, Carlsberg–Tetley, Scottish Courage, and Whitbread, two brewers without tied estate,[28] Guiness and Anheuser Busch, and four regional brewers, Charles Wells, Greene King, Ruddles, and Youngs. Brewers are distinguished by dummy variables, $\mathrm{BREW}_i,\ i = 1, \ldots, 10$.

Each brand has an alcohol content that is measured in percentage. This continuous variable is denoted ALC. Moreover, brands whose alcohol contents are greater than 4.2% are called premium, whereas those with lower alcohol contents are called regular beers. A dichotomous alcohol–content variable, PREM, that equals one for premium brands and zero otherwise, was therefore created.

Dummy variables that distinguish the establishment types, PUBM and PUBI for multiples and independents, regions of the country REGL and REGA, for London and Anglia, and time periods, PER1 and PER2 were also created.

Finally, a variable, NCB, was created as follows. First, each brand was assigned a spatial market, where brand i's market consists of the set of consumers whose most

[27] The outside good here consists of all other products that individuals purchase.

[28] Brewers without tied estate are not vertically integrated into retailing.

preferred brand is closer to i in taste space than to any other brand.[29] Euclidean distance in alcohol/coverage space was used in this calculation. Specifically, i's market consists of all points in alcohol/coverage space that are closer to i's location in that space than to any other brand's location. NCB_i is then the number of brands that share a market boundary (in the above sense) with i, where boundaries consist of indifferent consumers (i.e., loci of points that are equidistant from the two brands).[30]

A number of interaction variables are also used. Interactions with price are denoted PRVVV, where VVV is a characteristic. To illustrate, PRALC_i denotes price times alcohol content, $\mathrm{PRICE}_i \times \mathrm{ALC}_i$.

The set of endogenous variables consists of prices, volumes, and any variables that were constructed from prices or volumes. Coverage, in contrast, is considered to be weakly exogenous.[31] Whereas coverage would be endogenous in a longer-run model, according to people in the industry, there is considerable inertia in brand offerings. This is partially due to the existence of contracts between wholesalers and retailers and partially due to the need to change taps when brands are changed.[32]

Table 1 shows summary statistics by product type. 1A divides observations into the three major product groups: lagers, stouts, and ales, whereas 1B gives statistics for the two types of ales. In these tables, total volume is the sum of sales for that product type, whereas average volume is average sales per establishment. 1A shows that stouts are the most expensive than lagers, and that lagers have the highest alcohol contents. In addition, average coverage is highest for stouts. This statistic, however, is somewhat misleading, since it is due to the fact that Guiness is an outlier that is carried by a very large fraction of establishments. Finally, cask–conditioned ales have higher prices and sell larger volumes than keg ales. The volume statistics must be viewed with caution, however, since some of the most popular brands have keg variants.

Table 2 contains summary statistics by establishment type and region of the country. This table shows that prices are higher and volumes are lower in multiple establishments. In addition, both prices and volumes are higher in London.

[29] This construction does not rely on a discrete–choice assumption. Consumers can have a most–preferred brand and still consume more than one brand. Moreover, they can consume brands in variable amounts.

[30] The details of this construction can be found in Pinkse and Slade (2000).

[31] This assumption is tested below.

[32] The 'guest' beer is an exception. With such beers, a plaque with the name of the brand is merely hung over the tap.

7.2 The Metrics

Using the same data, Pinkse and Slade (2000) experiment with a number of metrics or measures of similarity of beer brands. These include several discrete measures: same product type, same brewer, and various measures of being nearest neighbors or sharing a market boundary in product–characteristic space. Two continuous measures of closeness, one in alcohol–content and the other in coverage space, are also used.

They find that one metric stands out in the sense that it has the greatest explanatory power, both by itself and in equations that include several measures. That metric, WPROD, is the same–product–type measure that is set equal to one if both brands are, for example real ales, and zero otherwise, and then normalized so that the entries in a row sum to one. A second measure, the similar–alcohol–content measure, is also included in their final specification. That metric is calculated as $\mathrm{WALC}_{ij} = 1/(1 + 2 \mid \mathrm{ALC}_i - \mathrm{ALC}_j \mid)$. I use the same metrics here.

To create average rival prices, the vector, PRICE, is premultiplied by each distance matrix, W, and the product is denoted RPW. For example, RPPROD is WPROD$\times$PRICE, which has as ith element the average of the prices of the other brands that are of the same type as i.

7.3 Cost Data

The Monopolies and Mergers Commission performed a detailed study of brewing and wholesaling costs by brand and company. In addition, they assessed retailing costs in managed public houses.[33] A summary of the results of that study is published in MMC (1989). Although the assessment of costs was conducted on a brand and company basis, only aggregate costs by product type are publicly available.[34] The MMC used volume weights to calculate average unit costs, where the volumes were based on the sales of each brand in managed houses.

Brewing and wholesaling costs include material, delivery, excise, and advertising expenses per unit sold. Retailing costs include labor and wastage. Finally, combined costs include VAT. Table 3 summarizes those costs by product type. Two changes to the MMC figures were made. First, their figures include overhead, which is excluded here because it is a fixed cost. Second their figures do not include advertising and marketing costs. Nevertheless, several of the companies report advertising expenditures per unit sold, and the numbers in the table are averages of those figures.

[33] Managed public houses are owned and operated by a brewer.

[34] Some company data are also available in a form that does not identify the companies.

The table shows that margins in brewing average 30%, which is moderately high. Retail margins, however, are considerably lower, which causes the combined margins to be modest, on average 14%. There are reasons to believe, however, that 30% is a more representative figure. Indeed, most retail establishments are not operated by a brewer (are not managed), and wholesale prices to other types of establishments are higher than transfer prices to managed public houses.

The last row of the table contains the updated cost figures in 1995 pence per pint. Updating was performed to reflect inflation. To do this, the closest available price index for each category of expense was collected and expenditures in each category were multiplied by the ratio of the appropriate price index in 1995 to the corresponding index in 1985.

When I interviewed brewers and asked questions concerning their costs, I uncovered a number of factors that could cause the updated costs to be inaccurate. In particular, advertising–to–sales ratios have increased in recent years, particularly for best–selling lagers. In addition, a higher fraction of the stout that is consumed is now brewed in the UK. Finally, all brewers that were interviewed claimed that retailing is now at least as profitable as brewing and perhaps more so. In the absence of better numbers, however, the updated MMC figures are used as $\check{c}$.

If brewing were subject to constant returns to scale, these would be marginal costs. Under increasing returns, however, which could be a more reasonable assumption, unit costs overestimate marginal costs. Unfortunately, the MMC produced no quantitative information on economies of scale.

8 Empirical Results

8.1 Demand

Nested Logit Demand

Three specifications of the nested–logit equation (2) are shown. The first is obtained by setting $\sigma = 0$ and $\alpha_i = \alpha$, which yields the standard logit. The second has $\sigma > 0$ and $\alpha_i = \alpha$, which is the standard nested–logit. The third allows α to vary by brand (i.e., p_i is interacted with with z_i).

Table 4 summarizes the estimated logit–demand equations ($\sigma = 0$). All specifications contain the log of coverage, LCOV, and time-period, regional, and product fixed effects. The specifications differ by the presence or absence of ALC, PREM, and brewer fixed effects. In particular, because ALC (alcohol content in percentage) and PREM (a dummy for alcohol content $> 4.2\%$) both measure a brand's strength, some specifications contain both of these variables, whereas others contain only one.

The table shows that the coefficient of PRICE is negative as predicted in only two of the six specifications. Moreover, when this coefficient is negative, it is not significant at conventional levels. Nevertheless, the negative estimate (-0.0007) is used in the calculation of the logit elasticities, since otherwise demand would slope upwards, and brands would be complements.

Brand own–price elasticities are calculated holding the prices of all other brands constant. At the mean of the data, the logit own–price elasticity is -0.115, which is not a reasonable value. Indeed, demand for individual brands should be highly elastic, since there are many close substitutes.

Brand cross–price elasticities are calculated allowing the price of a single rival brand to increase, holding own price and the prices of all other rivals constant. With the logit, those elasticities vary only by brand, since off–diagonal entries in a column of the logit–elasticity matrix are equal by assumption. At the mean of the data, this elasticity is 0.0001, which is also very low. The logit–demand specification is therefore not very satisfactory for this application.

Table 5 summarizes the estimated nested–logit–demand equations ($\sigma \neq 0$). In the first half of 5A, α, the coefficient of price, is constant. In contrast to the logit, however, brands are partitioned into four groups according to product type — lager, stout, keg ale, and real ale. As with the logit, most of the estimated coefficients of PRICE are positive (4 out of 6). The magnitudes of the negative estimates, however, are greater, and their significance is somewhat higher. Nevertheless, when they are negative, the estimated price coefficients are still not significant at conventional levels. The estimated coefficients of LGRSHARE — the log of a brand's share of the group to which it belongs — in contrast, are positive, less than one and significant at conventional levels.

The second half of table 5A shows specifications in which prices are interacted with brand characteristics. The table shows estimates of α, the constant coefficient of price, as well as $\alpha_i = \alpha(x_i)$, the slope of the demand equation evaluated at the mean of the product characteristics.[35] This section of the table shows that when prices are interacted with characteristics, slopes are neither larger in magnitude nor

[35] The characterisitcs that are included in this specification are the same as with the DM specifications that are presented below. In particular, ALC appears in the intecept term, whereas PREM is interacted with price. For this reason, there are only two entries in this portion of the table.

statistically more significant than when they are not.

To give the nested logit the benefit of the doubt, the specification with the slope that is largest in magnitude (# 5 with $\alpha = 0.0026$) is used in the calculation of elasticities and the evaluation of market conduct. This equation, which is shown in full in table 5B, also has the highest estimate of σ (0.83), a value that implies that within–group correlation of tastes is very high.

Table 5B, which contains the full specification for the MNL, shows that a brand's share is higher when its coverage is higher. In addition, shares are lower in London, which simply reflects small regional differences in consumption per head. Finally, all else equal, when a brand is a lager (stout or keg ale) its share is higher (lower), where comparisons are made with respect to real ales.

At the mean of the data, the brand own–price elasticity is -2.4, which is also the median elasticity. The range is -0.7 to -3.2. Demand is therefore substantially more elastic with the nested logit than with the logit. Compared to estimates reported in Hausman, Leonard, and Zona (HLZ, 1994), however, where own–price elasticities for US brands average -5.0, the MNL own–price elasticities still seem low.

MNL cross-price elasticities take on two values per brand, one for brands in the same group and one for brands in different groups. At the mean of the data, these elasticities are 0.137 for the former and 0.0002 for the latter, an indication that most substitution is within groups, as the estimate of σ already suggested. There is, however, substantial variation in partial cross–price elasticities across groups. Indeed, average samegroup cross–price elasticities for lagers, stouts, keg ales, and real ales are 0.08, 0.52, 0.19, and 0.10, respectively. These differences, however, are driven almost entirely by differences in same–group shares (i.e., by differences in the number of brands in each group).

Distance–Metric Demand

Table 6 summarizes the estimated distance–metric–demand equations. The first two equations in this table, however, are included only for comparison with the logit and MNL. Recall that the coefficients of price in those equations were often positive and, when negative, not significant at conventional levels. To demonstrate that this finding is not simply due to functional form, linear equations are shown in which prices are not interacted with characteristics and distance–weighted rival prices are not included. As with the logit and MNL, the slopes of these equations are not consistently negative and are not significant at conventional levels.

The third equation in table 6 is the DM specification. This equation is divided into three sections: the intercept terms, $A_i = \beta^T z_i$, and the own-price terms, b_{ii}, are functions of the characteristics, z_i. The characteristics in b_{ii} however, have been interacted with price to allow the own-price elasticities to vary with those characteristics. The rival-price term b_{ij}, $j \neq i$, in contrast, is a function of the distance measures, d_{ij}.

In theory, all characteristics that are included in z_i could enter both A_i and b_{ii}. In practice, however, each characteristic is highly correlated with the interaction of that characteristic with price. For this reason, the variables that appear in A_i and those that appear in b_{ii} are never the same. An attempt was made to allocate the variables in a sensible fashion. Nevertheless, the allocation is somewhat arbitrary. In addition, since coverage was found to be an important determinant of both brand–market size and own–price elasticity, coverage is included in both parts of the equation. To avoid collinearity, different functional forms are used in the two parts, with LCOV = log(COV) and COVR = 1/COV.

First, consider the own–price effect, b_{ii}, in the third specification. In contrast to the earlier findings, this slope is both negative and significant. Moreover, this is true not only of the coefficient of price, but also of most of the interaction terms. In particular, premium and popular brands have steeper (i.e., more negative) slopes (recall that COVR is an inverse measure of coverage), and when a brand has a large number of neighbors, its sales are more price sensitive. Allowing the slope to vary with the characteristics is therefore important.

The second part of the equation, which assesses the determinants of brand substitutability, shows that the coefficient of the same–product–type rival–price measure, RPPROD, is both positive and significant at 1%. This implies that competition is stronger among brands that are in the same group. The coefficient of the similar–alcohol–content variable, RPALC, is positive and siginificant at 10%. The DM demand equation is thus similar to a nested logit, where the nests are product types. In addition to the product groupings, however, beers with similar alcohol contents tend to compete regardless of type, but the strength of that rivalry is less pronounced.

Finally, consider the intercepts, A_i. In all three specifications, high coverage is associated with high sales. In addition, sales are higher in independent establishments and in London. Furthermore, a high alcohol content has a positive but weak effect on sales.

For comparison purposes, the last column of table 6 contains OLS estimates of the DM demand equation. The table shows that the OLS estimates of the coefficients of the endogenous variables are somewhat smaller in magnitude than the GMM estimates but are similar in significance.

As a check on the DM demand equation, its identification was assessed. First, I used the test of correlation between the residuals in that equation and various groups of instruments that is discussed in the appendix. This process uncovered no evidence of endogeneity. In particular, when price in the other region was investigated by itself, the p value for the test was 0.20, and when the instruments as a group were assessed, the value was 0.38.

Second, I experimented with various sets of instruments. The equations shown in table 6 were estimated using both sets of instruments — those constructed from prices in the other region and those constructed from characteristics of rival brands. When the demand equation was estimated with either set by itself, results were similar.

With respect to curvature, all of the eigenvalues of the estimated matrix B, which is the Hessian of the indirect–utility function, are nonnegative. This must be the case if $\hat{B}$ is negative semidefinite, and it shows a close adherence to quasi-convexity of the indirect-utility function.

Turning to the elasticities, with the DM specification, brand own-price elasticities vary with the characteristics of each brand. The mean own–price elasticity, however, is -4.6. Demand is therefore considerably more elastic than with either the logit or the MNL specifications. Furthermore, it is similar to, but slightly smaller in magnitude than, the Hausman, Leonard, and Zona (1994) average of -5.0. The median own–price elasticity is -4.1, which reflects an asymmetric distribution with a few large values in the upper tails.

Unlike the logit and MNL cross–price elasticities, DM cross–price elasticities vary with each brand pair. One can, however, define a total cross–price elasticity, which is the percentage change in one brand's sales due to a 1% increase in the prices of all of its rivals. This elasticity averages 3.9.

As it is not practical to examine 63 own and approximately 4,000 cross–price elasticities, table 7 contains elasticities for a selected subsample of brands. This subsample contains one regular lager, Tennants Pilsner, two premium lagers, Stella Artois and Lowenbrau, two keg ales, Toby and Websters Yorks Bitter, two real ales, and one stout. One of the real ales, Courage Best, is a best-selling brand brewed by a national brewer, whereas the other, Greene King IPA, is a small–sales brand brewed by a regional brewer. Finally, the stout, Guiness, is an outlier with a coverage that is substantially higher than that of any other brand in the sample.

In addition to identifying the type of each brand, the first row of the table shows the brand's alcohol content and number of neighbors, where a neighbor shares a market boundary with the given brand, and markets are delineated in characteristic space (see subsection 7.1).

The table shows that there is substantial variation in own–price elasticities, and that most of the magnitudes are plausible. In particular, if one ranks the the reciprocals of the (absolute values) of the own–price elasticities and ranks the estimated price/cost margins, the rankings are very similar. Nevertheless, the own–price elasticity of the small–sales brand, Greene King IPA, seems unrealistically high. This is due to the fact that elasticity estimates are inversely related to sales. Furthermore, the own–price elasticity for Guiness is very low, which is due to the fact Guiness has very high sales (as well as the fact that it has few neighbors). It therefore seems likely that the model over (under) estimates magnitudes of elasticities for brands with very small (large) market shares.[36]

Turning to the brand cross-price elasticities, the table illustrates that, as expected, these are greater when brands are of the same type and have similar alcohol contents. To illustrate, the three lagers are closer substitutes for one another than for the other brands in the table, and the two premium lagers, Stella and Lowenbrau, are closer substitutes for one another than for the regular lager, Tennants. The table also shows that Guiness is not a close substitute for any of the other brands. In addition, the cross–price elasticities for the small–coverage brand, Greene King IPA, seem high relative to the other estimates, which is a further indication that the model over predicts substitution possibilities for brands with small market shares.

All own-price elasticities are significant at 1%. Cross–price elasticities for brands of the same type (e.g., two lagers) are also significant at 1%. When brands are of different types (e.g., a lager and a stout), however, their cross–price elasticities are not significant at 5% but are at 10%.

Finally, table 8 compares average own and cross–price elasticities across models. It shows that as one moves from the logit to the nested–logit to the DM specification, the magnitudes of the elasticities increase. For comparison purposes, the table also contains the average elasticities for US brands of beer that were estimated by Hausman, Leonard, and Zona (1995), which are somewhat larger than the DM estimates.

8.2 Decomposition of Market Power

Corresponding to any demand equation and partition $\mathcal{K}$ that determines brand ownership, there is a set of static Nash–equilibrium prices and margins. Moreover, those

[36] This is a common problem with flexible functional forms such as a translog.

margins can be decomposed in the manner that is described in section 5. In particular, one can assess unilateral and coordinated effects and can further decompose the former into differentiation and concentration effects.

Table 9 summarizes the equilibrium prices and margins that are associated with various demand equations and games, where margins or Lerner indices are calculated using exogenous costs and predicted prices. Each of the predictions can be compared to the observed prices and margins that are summarized in the last row of the table.

For the nested logit, only *status–quo* prices are computed, where the *status–quo* game corresponds to the actual brand–ownership pattern. The table shows that the mean *status–quo* MNL price is 245 pence per pint, which can be compared to the observed mean of 168. MNL *status–quo* prices are thus on average about 50% higher than observed prices, which is an indication that, with the MNL demand equation, behavior is estimated to be substantially more rivalrous than Bertrand. Furthermore, MNL *status–quo* margins at the mean of the data are nearly 90%, which can be compared to the observed margins of30%. One must conclude that either this market is very competitive or that the MNL model of demand underestimates price sensitivity in the beer data.

Although it is possible that this market is very competitive, to me it seems more likely that the above finding is due to the inability of the MNL demand specification to uncover significant price responsiveness in the beer data. The end result is that the estimated MNL own and cross–price elasticities are relatively small in magnitude and insignificantly different from zero. If those estimates were taken seriously, Bertrand decision makers would choose substantially higher prices than the ones that are observed.

Table 9 also shows three hypothetical equilibria that were calculated using the distance–metric–demand equation: marginal–cost pricing, Bertrand pricing with single–product firms, and Bertrand pricing with multiproduct firms (the *status–quo* game).[37] The first results in prices that are on average 40 pence per pint lower than observed prices and in margins that are everywhere zero. Single–product prices, in contrast, which average 159 pence per pint, are only 9 pence lower than observed prices. This means that differentiation by itself endows the firms in this market with substantial pricing power and results in margins of over 23%. Finally, *satatus–quo* prices and margins are extremely close to observed prices and margins.

Using the DM demand equation, one can decompose the observed margins of

[37] Joint–profit maximizing prices and margins are not shown. Indeed, since industry demand is estimated to be inelastic, the monopoly markup model does not perform well.

30% into three factors. The first — the differentiation effect — is due to the fact that brands of beer are not identical and consumers differ in their tastes for beer characteristics. This effect accounts for about three quarters of the total margin. The second — the concentration effect — is due to the fact that there are 10 rather than 63 brewers in the sample. This effect accounts for the remaining quarter, which means that there is nothing left over to be explained by tacit or overt collusion. In other words, whereas substantial market power is uncovered, all of it is due to unilateral effects, and no evidence of coordinated effects can be found. Although this conclusion might have been unanticipated, it is similar to results reported in Nevo (2002) for the US breakfast–cereal industry, an industry where margins are substantially higher than in UK brewing. The estimated margins for these branded products can be contrasted with the situation that would prevail if the products were homogenous. With homogeneous commodities, Bertrand decision makers set prices equal to marginal costs and margins are zero.

8.3 Further Analysis of Coordinated Effects

The decomposition uncovers no evidence of coordinated effects, but it is not a statistical test. One can use the implicit market–conduct parameters, $\tilde{\theta}$ to test for coordinated effects more formally. First consider the implicit estimates obtained from the MNL demand equation, $\tilde{\theta}_{MNL}$. Both the mean and the median of those parameters are -0.6. Such low estimates imply that the market is very competitive. In particular, one can reject the hypothesis that $E(\theta_{MNL}) = 0$, which means that, as before, MNL behavior is estimated to be significantly more rivalrous than Bertrand.

Next consider the implicit DM market–conduct parameters, $\tilde{\theta}_{DM}$. The mean of those parameters is 0.014, the median is -0.011, and the range is -0.8 to 2.[38] Furthermore, the p value for the hypothesis that $1/n\Sigma_i\theta_{iDM} = 0$ is 0.46, which means that Bertrand behavior cannot be rejected. It appears that although the choice between MNL and DM specifications for demand strongly influences the conclusions that can be drawn concerning firm behavior, with neither specification is there any evidence of collusion.

The analysis thus far indicates that, on average, there are no coordinated effects. This finding, however, does not rule out the possibility that some firms (i.e., the dominant group) behave in a collusive fashion while others behave more competitively.

[38] All implicit market–conduct parameters are less than one except for one brand, Guiness, which has a value of two. Guiness, however, is an outlier with extremely high coverage.

One can use an econometrically estimated market–conduct function to assess this possibility.

It is possible to estimate a market–conduct function jointly with any of the demand equations. Given the unsatisfactory nature of the logit and MNL elasticities , however, joint estimation with one of those equations does not seem worthwhile. In particular, since the first–stage estimates are not significantly different from zero, it is unlikely that the second–stage estimates, which build on the first, would be more accurate.

In contrast to the MNL estimates, the DM elasticities are precisely estimated, and a DM market–conduct function is therefore presented. Table 10 shows three specifications for this function. The first contains only a variable DNAT that equals one if the firm that produces the brand is a national brewer (a member of the hypothesized dominant group) and zero otherwise. The second two, which differ according to the measure of brand strength that is used, also contain other explanatory variables. The table shows that the coefficient of DNAT is never significant at conventional levels, which imples that brands that are brewed by the nationals are not less competitively priced, and those firms do not form a dominant group. On the other hand, the second and third specifications show that more popular, higher–strength, and multiple–establishment beers are less competitively priced.[39] Finally, there is no evidence of conduct differences across product types.

9 Concluding Remarks

[39] The third regularity is perhaps due to the fact that vertical relationships between brewer and retailer are more complex when public houses are owned by brewers or retail chains (see Slade 1998).

References

Berndt, E.R., Fuss, M.A., and Waverman, L. (1977) "Dynamic Models of the Industrial Demand for Energy," Electric Power Institute, EA-580, Palo Alto, CA.

Berry, S., Levinsohn, J., and Pakes, A. (1995) "Automobile Prices in Market Equilibrium," *Econometrica*, 63: 841-890.

Blackorby, C. Primont, D. and Russell, R. (1978) *Duality, Separability, and Functional Structure: Theory and Economic Applications*, Amsterdam: North Holland.

Compte, O., Jenny, F., and Rey, P. (2002) "Capacity Constraints, Mergers, and Collusion," *European Economic Review*, 46: 1-29.

Doob, J.L. (1953) *Stochastic Processes*, Wiley: New York.

Feenstra, R.C. and Levinsohn, J.A. (1995) "Estimating Markups and Market Conduct with Multidimensional Product Attributes," *Review of Economic Studies*, 62: 19-52.

Hausman, J., Leonard, G., and Zona, D. (1994) "Competitive Analysis with Differentiated Products," *Annales d'Econometrie et de Statistique*, 34: 159-180.

Ivaldi, M. and Verboven, F. (2000) "Quantifying the Effects from Horizontal Mergers: The European Heavy Trucks Market," University of Toulouse mimeo.

Jayaratne, J. and Shapiro, C. (2000) "Simulating Partial Asset Divestitures to 'Fix' Mergers," *International Journal of the Economics of Business*, 7: 179-200.

Kuhn, K.-U. (2000) "An Economists' Guide Through the Joint Dominance Jungle," mimeo.

Kuhn, K.-U. and Motta, M. (2001) "The Economics of Joint Dominance," mimeo.

Lexicon (1999) "Joint Dominance," Competition Memo.

McFadden, D. (1974) "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, P. Zarembka (ed.) Academic: New York, 105-142.

McFadden, D. (1978a) "Modeling the Choice of Residential Location," in *Spatial Interaction Theory and Planning Models*, A. Karlgvist et. al. (eds.) Amsterdam: North Holland.

McFadden, D. (1978b) "The General Linear Profit Function," in *Production Economics: A Dual Approach to Theory and Applications*, Vol. 1, M. Fuss and D. McFadden (eds.) Amsterdam: North Holland, 269-286.

Monopolies and Mergers Commission (1989) *The Supply of Beer*, London: Her Majesty's Stationary Office.

Murphy, K.M. and Topel, R.H. (1985) "Estimation and Inference in Two-Step Models," *Journal of Business and Economic Statistics*, 3: 370-379.

Nevo, A. (2000) "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *RAND Journal of Economics*, 31: 395-421.

Nevo, A. (2001) "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 69: 307-342.

Newey, W.K. (1984) "A Method of Moments Interpretation of Sequential Estimators," *Economics Letters*, 14: 201-206.

Newey, W.K. and K.D. West (1987) "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 55: 703-708.

Pinkse, J., Slade, M.E., and Brett, C. (1998) "Spatial Price Competition: A Semiparametric Approach," *Econometrica*, forthcoming, http://www.econ.ubc.ca/slade/price.pdf.

Pinkse, J. and Slade, M.E. (2000) "Mergers, Brand Competition, and the Price of a Pint," http://www.econ.ubc.ca/slade/beer.pdf.

Shapiro, C. (1996), "Mergers with Differentiated Products," *Antitrust*, Spring: 23-30.

Slade, M.E. (1994) "What Does an Oligopoly Maximize?" *Journal of Industrial Economics*, 57: 45-62.

Slade, M.E. (1998) "Beer and the Tie: Did Divestiture of Brewer-Owned Public Houses Lead to Higher Beer Prices?" *Economic Journal*, 108: 565-602.

Slade, M.E. (1998) "Assessing Market Power in UK Brewing," UBC mimeo.

Werden, G.J., and Froeb, L.M. (1994) "The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy," *Journal of Law, Economics, and*

Organization, 10: 407-426.

Werden, G.J. (1999) "Expert Report in *United States v. Interstate Bakeries Corp. and Continental Baking Co.*, US Department of Justice mimeo.

APPENDIX

Estimation and Testing

The Two-Step GMM Estimation

Stage 1: GMM Estimation of the Demand Equation

One can write the demand equation as

$$f_1(X_1, \beta) = v,$$

where X_1 is an nxk_1 matrix of endogenous and exogenous variables, β is a p_1 vector of parameters, and v is an n vector of errors. Let Z be an nxm_1 matrix of instruments with $m_1 > p_1$.

The GMM estimator of β is the minimum over β of

$$v^T Z(Z^T \Omega_1 Z)^{-1} Z^T v,$$

where Ω_1 is a matrix that corrects for heteroskedasticity and spatial correlation of an unknown form. Specifically, Ω_1 has i, i element $\hat{v}_i^2$ and i, j element $\iota_{ij}\hat{v}_i\hat{v}_j$, where $\hat{v}$ is the vector of two-stage least-squares residuals, and ι_{ij} equals one if j is one of i's L closest neighbors and vice versa, one half if either i is one of j's or j is one of i's L closest neighbors (but not both), and zero otherwise. Closeness between i and j is measured here by $\text{WPROD}_{ij} \times \text{WALC}_{ij}$, where WPROD and WALC are the metrics that appear in the demand equatio n.

This yields $\tilde{\beta}$ and $\tilde{\Sigma}_\beta$, the GMM estimates of β and Var(β), where

$$\tilde{\Sigma}_\beta = [H_{1\beta}^T Z(Z^T \Omega_1 Z)^{-1} Z^T H_{1\beta}]^{-1},$$

and $H_{1\beta}$ is the nxp_1 matrix $\partial f_1/\partial\beta$ evaluated at $\tilde{\beta}$.

Stage 2: Estimation of the First-Order Condition

One can write the first-order condition as

$$y(\beta) = X_2\gamma + u \quad \text{or} \quad f_2(y, X_2, \beta, \gamma) = u,$$

where y is an endogenous variable, X_2 is an nxp_2 matrix of exogenous variables,[40] γ is a p_2 vector of parameters, and u is an n vector of errors. Since this equation is exactly identified, the GMM estimates, $\tilde{\gamma}$, can be obtained by simply solving the moment conditions.[41] The standard errors of $\tilde{\gamma}$, however, must be corrected to reflect the fact that β was estimated in a prior stage.

Let Ω_2 be defined like Ω_1 with $\hat{v}$ replaced by $\hat{u}$ and $H_{2\beta}$ be the nxp_1 matrix $\partial f_2/\partial\beta$, evaluated at $\tilde{\beta}$. Then, if u and v are uncorrelated,[42]

$$\tilde{\Sigma}_\gamma = (X_2^T X_2)^{-1}(X_2^T \Omega_2 X_2)(X_2^T X_2)^{-1} + (X_2^T X_2)^{-1}(X_2^T H_{2\beta}\tilde{\Sigma}_\beta H_{2\beta}^T X_2)(X_2^T X_2)^{-1}.$$

Tests of Instrument Validity

Suppose that the estimating equation is $y = R\delta + \epsilon$ and that $\{(z_i, \epsilon_i, Q_i, R_i)\}$ is i.i.d., where z_i is the suspect instrument, Q_i is the set of nonsuspect instruments, R_i is the set of explanatory variables, which includes at least one endogenous regressor, and ϵ_i is the error for observation i. For z to be a valid instrument, ϵ and z must be element-wise uncorrelated (i.e., $E(z_i\epsilon_i) = 0$). Let $P_Q = Q(Q^T Q)^{-1}Q^T, \Omega = Var(\epsilon \mid R, z, Q), M = I - R(R^T P_Q)^{-1}R^T P_Q, \tilde{V} = z^T M\tilde{\Omega}M^T z$, where $\tilde{\Omega}$ is our estimate of Ω, and $\hat{\epsilon}$ be the residuals from an IV estimation using Q (but not z) as instruments. Then, under mild regularity conditions on $\tilde{\Omega}$,

$$\tilde{V}^{-1/2} z^T \hat{\epsilon} = \tilde{V}^{-1/2} z^T M\epsilon$$

has a limiting $N(0, 1)$ distribution (see Pinkse, Slade and Brett 2000).

If one wants to test more than one instrument at a time, it is possible to use a matrix Z instead of the vector z to get a limiting $N(0, I)$ distribution. Taking the squared length, one has a limiting χ^2 distribution whose number of degrees of freedom is equal to the number of instruments tested.

40 When $\delta \neq 1$ in the cost function, instruments are used.

41 This is simply OLS with correction for heteroskedasticity and spatial correlation of an unknown form.

42 v is an unobserved demand factor, whereas u is an unobserved cost factor. The assumption that they are uncorrelated in thus not unreasonable. The formula is similar to equation (8) in Newey (1984) for the uncorrelated case. The difference is that his first stage estimation is exactly identified.

Table 1:

Summary Statistics by Product Type[a)]
London and Anglia Draft Beer
Brands in Sample

1A: Three Major Groups

Variable	*Units*	*Lager*	*Stout*	*Ale*
Average Price	Pence per pint	175.3	184.0	154.6
Total Volume	100 barrels	8732	1494	4451
Average Volume	100 barrels	47.5	67.9	18.7
Market Share	%	59	10	31
Average Coverage	%	10.1	31.3	6.3
Alcohol Content	%	4.3	4.1	3.9
Number of Brands		25	4	34

1B: Ales

Variable	*Units*	*Cask Conditioned (‘Real’)*	*Keg*
Average Price	Pence per pint	158.3	148.2
Total Volume	100 barrels	3092	1359
Average Volume	100 barrels	20.3	15.8
Market Share	%	21.5	9.5
Average Coverage	%	7.0	5.2
Alcohol Content	%	4.1	3.7
Number of Brands		21	13

[a)] Averages taken over brands, regions, and time periods

Table 2:

Summary Statistics by Establishment Type and Region[a)]
Draft Beer
Brands in Sample

2A: London

Establishment Type	*Average Price*	*Average Volume*	*Average Coverage*
Multiples	174.5	42.7	11.3
Independents	160.9	58.0	7.2

2B: Anglia

Establishment Type	*Average Price*	*Average Volume*	*Average Coverage*
Multiples	168.5	10.4	10.5
Independents	155.6	20.4	7.7

[a)] Averages taken over brands, regions, and time periods

Table 3:

Brewer Costs and Margins[a)]

	Lager	*Stout*	*Real Ale*	*Keg Ale*
Brewing and Wholesaling				
Duty	16.4	0.0	17.0	16.9
Materials	2.3	0.0	2.7	2.5
Other	5.0	0.0	3.9	5.3
Bought-in-Beer	1.5	39.2	-	-
Delivery	5.6	0.0	4.2	4.4
Advertising and Marketing[b)]	0.9	0.0	0.8	0.8
B&W Cost	31.7	39.2	28.6	29.9
Transfer Price	45.4	54.2	41.6	41.1
B&W Profit	13.7	15.0	13.0	11.2
B&W Margins (%)	30.2	27.7	31.3	27.3
Retailing				
Transfer Price	45.4	54.2	41.6	41.1
Wastage	1.1	1.4	1.0	1.0
Labor	33.4	35.0	34.0	32.6
Retail Cost	79.9	90.6	76.6	74.7
Takings	94.1	104.7	82.4	81.1
Retail Profit	14.2	14.1	5.8	6.4
Retail Margins (%)	15.1	13.5	7.0	7.9
Combined				
VAT	12.3	13.7	10.8	10.6
Combined Cost	78.5	89.3	74.4	74.1
Combined Profit	15.6	15.4	8.0	7.0
Combined Margins (%)	16.6	14.7	9.7	8.6
Updated Costs				
Brewing, Wholesaling, and Retailing	132	147	125	124

a) Excludes overhead.
b) 1% of takings.

Source: MMC (1989)

Table 4:

Logit Demand Equations, IV Estimates

($\sigma = 0$)

Dependent Variable: LSHARE

Equation	*PRICE* *($-\alpha$)*	*ALC*	*PREM*	*Brewer Fixed Effects*
1	0.014 (1.7)	yes	yes	yes
2	-0.0007 (-0.1)	yes	no	yes
3	0.017 (2.4)	no	yes	yes
4	0.012 (1.7)	yes	yes	no
5	-0.0007 (-0.1)	yes	no	no
6	0.016 (2.6)	no	yes	no

Other explanatory variables: LCOV, PER1, REGL, and $PROD_i$, $i = 1,\ldots,4$
t statistics in parentheses

Table 5:

Nested Logit Demand Equations, IV Estimates

($\sigma > 0$)

Dependent Variable: LSHARE

5A: Various Specifications

α Constant						
Equation	*PRICE* *($-\alpha$)*	*SLOPE* *(α_j)*	*LGRSHARE* *(σ)*	*ALC*	*PREM*	*Brewer Fixed Effects*
1	0.0081 (1.9)		0.554 (6.6)	yes	yes	yes
2	-0.0011 (-0.3)		0.691 (10.0)	yes	no	yes
3	0.0116 (3.1)		0.546 (6.2)	no	yes	yes
4	0.0041 (1.2)		0.644 (7.5)	yes	yes	no
5	-0.0026 (-1.0)		0.830 (12.3)	yes	no	no
6	0.0076 (2.4)		0.664 (7.6)	no	yes	no
α Variable (Price interacted with characteristics)						
Equation	*PRICE* *($-\alpha$)*	*SLOPE* *($-\alpha_j$)[a]*	*LGRSHARE* *(σ)*	*ALC*	*PREM*	*Brewer Fixed Effects*
7	0.0053 (1.5)	0.0047 (1.4)	0.658 (7.3)	yes	no	yes
8	-0.0029 (-1.2)	-0.0024 (-1.0)	0.776 (14.5)	yes	no	no

Other explanatory variables: LCOV, PER1, REGL, and $PROD_i$, $i = 1,\ldots,4$
Four groups, lager, stout, keg ale, and real ale
t statistics in parentheses
[a] Evaluated at the mean of the data

Table 5:

Nested Logit Demand Equations, IV Estimates

($\sigma > 0$)

Dependent Variable: LSHARE

5B: Final Specification

Variable	*Coefficient*	*t Statistic*
PRICE *($-\alpha$)*	-0.0026	-1.0
LGRSHARE *(σ)*	0.830	12.3
LCOV	0.161	2.0
ALC	0.031	0.6
PER1	0.024	0.8
REGL	-0.135	-4.2
$PROD_1$ *(lager)*	0.832	14.6
$PROD_2$ *(stout)*	-0.760	-8.0
$PROD_3$ *(keg ale)*	-0.695	-9.1
Constant	-2.541	-4.1

Four groups, lager, stout, keg ale, and real ale

Table 6:

Distance-Metric Demand Equations[a)]

Dependent Variable: VOL

	1	2	3[b)]	4
Estimation Technique	IV	IV	GMM	OLS
Own Price *(b_{ii})*				
PRICE	0.348 (1.3)	-0.811 (-1.2)	-1.125 (-2.9)	-0.871 (-2.6)
PRCOVR			0.165 (7.8)	0.153 (7.4)
PRPREM			-0.030 (-0.1)	-0.025 (-0.7)
PRNCB			-0.117 (-2.7)	-0.106 (-2.5)
Rival Price *(b_{ij})*				
RPPROD			0.712 (2.6)	0.747 (2.9)
RPALC			0.215 (1.6)	0.172 (1.0)
Intercept *(A_i)*				
LCOV	30.64 (11.9)	32.27 (11.4)	60.29 (11.7)	56.81 (13.6)
ALC	9.145 (1.4)	6.660 (0.5)	8.801 (0.7)	8.36 (0.7)
PUBM	-25.93 (-4.4)	-10.47 (-1.1)	-10.97 (-1.9)	-16.03 (-3.1)
PER1	2.229 (0.5)	-0.221 (-0.1)	3.806 (0.8)	3.886 (0.8)
REGL	30.22 (6.1)	36.60 (6.2)	31.49 (6.4)	31.13 (6.4)
Product Fixed Effects	no	yes	no	no

a) t statistics in parentheses
b) Standard errors corrected for heteroskedasticity and spatial correlation of an unknown form

Table 7:

Own and Cross-Price Elasticities for Selected Brands[a)]
Using the DM Demand Equation

Evaluated at observed prices and quantities

Brand *Alcohol Content* *Product Type* *# Neighbors*	***Tennants Pilsner*** *3.2%* *Reg. Lager* *12*	***Stella Artois*** *5.2%* *Prem. Lager* *8*	***Lowenbrau*** *5.0%* *Prem. Lager* *8*	***Toby Bitter*** *3.3%* *Keg Ale* *12*	***Websters Yorks Bitter*** *3.5%* *Keg Ale* *8*	***Courage Best*** *4.0%* *Real Ale* *15*	***Greene King IPA*** *3.6%* *Real Ale* *9*	***Guiness*** *4.1%* *Stout* *2*
Tennents Pilsner	-4.80	0.189	0.181	0.021	0.018	0.011	0.013	0.012
Stella Artois	0.068	-2.49	0.085	0.002	0.003	0.004	0.003	0.005
Lowenbrau	0.091	0.119	-3.10	0.003	0.004	0.006	0.004	0.007
Toby Bitter	0.030	0.009	0.009	-4.87	0.457	0.015	0.018	0.016
Websters Bitter	0.013	0.006	0.006	0.227	-3.20	0.010	0.013	0.010
Courage Best	0.009	0.007	0.008	0.010	0.011	-2.79	0.124	0.021
Greene King IPA	0.064	0.038	0.041	0.061	0.090	0.852	-12.62	0.081
Guiness	0.002	0.002	0.002	0.001	0.002	0.004	0.002	-0.93

Table 8:

Summary of Elasticity Estimates
Averages Across Brands

Demand Model	*Own-Price Elasticity*	*Cross-Price Elasticity*
Logit	- 0.12	0.0001
Nested Logit	- 2.4	0.0344
Distance Metric	- 4.6	0.0632
AIDS *Hausman, Leonard, and Zona (1995)*	- 5.0	0.12

Table 9:

Predicted Equilibrium Prices and Margins

Demand Equation	*Equilibrium*	*Mean*	*Standard Deviation*	*% Difference*	*Margins*[a)]
Nested Logit	*Status Quo*	244.7	44.2	45.8	89.5
Distance Metric	*Marginal-Cost Pricing*	129.1	5.2	-23.1	0.0
	Single-Product Firms	159.4	19.8	-5.1	23.5
	Status Quo	168.4	29.5	0.4	30.4
Observed Prices		167.8	20.2		29.9

a) Calculated using the exogenous cost estimates.

Table 10:
Market-Conduct Equations[a) b)]
2-Step GMM Estimates Using the DM Demand Equation

Equation	1	2	3
DNAT	0.082 (1.3)	0.077 (1.4)	0.081 (1.4)
LCOV		0.017 (2.4)	0.016 (2.2)
PREM		0.059 (3.2)	
ALC			0.052 (4.0)
PUBM		0.079 (4.5)	0.080 (4.6)
$PROD_1$ *(lager)*		0.008 (0.3)	0.003 (0.1)
$PROD_2$ *(stout)*		0.003 (0.1)	-0.016 (-0.8)
$PROD_3$ *(keg ale)*		0.003 (1.2)	0.047 (1.3)
Constant	-0.078 (-1.3)	-0.173 (-2.8)	-0.366 (-4.0)

a) t statistics in parentheses
b) Standard errors corrected for heteroskedasticity and spatial correlation of an unknown form.