Market Power and Joint 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

May 2002

Abstract:

Market power and joint dominance are examined in UK brewing, an industry that is periodically scrutinized by both UK and EU competition authorities. I distinguish between *unilateral* and *coordinated* effects, where the latter is equated with joint dominance, and I show how one can distinguish between the two econometrically. The application makes use of two classes of demand equations: the nested logit of Mc-Fadden (1978a) and the distance–metric of Pinkse, Slade, and Brett (2002). 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 characteristics 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 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 the firms possess market power and if the merger will lead to increases in that power. Furthermore, it is becoming increasingly common to supplement traditional competition–policy analysis with econometric evaluations of industry performance. European Union (EU) and North American authorities, however, take different approaches to the evaluation of market power.

To illustrate, 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 evaluation 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 particular, it discusses how rules of thumb based on unilateral and coordinated effects can lead to different choices of antitrust cases to investigate. In addition, if a case 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 (2002).

[3] See, e.g., Lexicon (1999a), 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 been scrutinized by both UK and EU authorities on numerous occasions. 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 and industry practices, 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 (2002) and Pinkse and Slade (2002). 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. In particular, the two–stage procedure that is used to estimate the demand equations and first–order conditions is described, the

---

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

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). 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, that power is due entirely to unilateral effects. In particular, there is no evidence of a dominant group. In addition, I find that whereas brand characteristics appear to be important determinants of margins, firm characteristics do not.

2 EU and North American Approaches

Competition authorities are responsible for policing violations of many sorts of antitrust laws. I limit attention, however, to abuse–of–dominance and merger cases. With those cases, the common goal of EU and North American authorities is to evaluate market power and how certain practices or acts contribute to that power. The approaches that they take to pursuing their common 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 determine 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 be used to distinguish between the two.

---

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

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 mergers to investigate. For example, in the application below, based on the HHI and average price/cost margins, the UK brewing industry is fairly concentrated and firms in the industry have substantial market power. I find, however, that that power is due entirely to unilateral effects. This means that North American guidelines would dictate that mergers in the industry 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 larger of the two acquires a smaller firm. The acquisition will cause the HHI to rise, and one might conclude that the increase in 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 for firms to reach agreements when they are less similar. Moreover, Compte, Jenny, and Rey (2002) show formally that, in a repeated pricing 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.

Simple 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, however, 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 predict 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.[7]

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, however, there has been a succession of mergers that have increased concentration

---

[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.

[7] Specifically, there have been about 30 reviews since 1969.

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.[8]

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,[9] are premium beers with relatively high alcohol contents.

With respect to production, the number of brewers has declined steadily. Indeed, in 1900, there were nearly 1,500 brewery companies, but this number fell dramatically

---

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

[9] 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.

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.

For many decades, there were six national brewers that jointly accounted for about 75% of sales. If a dominant group existed, therefore, the nationals were the natural members of that group. Furthermore, in the early 1990s, mergers reduced the number of nationals from six to four, and the increase in horizontal concentration that accompanied those mergers made the possibility of joint dominance still more likely.

This snapshot of the UK beer industry shows significant changes in tastes and consumption habits as well as a decline in the number of firms 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 thoseeffects, 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 nested multinominal logit (NML) and the distance metric (DM).[10]   Since the

---

[10] A third demand specification has been used to model 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 considerations

first is well known and the second is developed elsewhere (Pinkse, Slade, and Brett 2002 and Pinkse and Slade 2002), I simply reproduce the estimating equations and the equations for the own and cross–price elasticities of demand. Both specifications yield expressions for the elasticities that are simple and easy to interpret.

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 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 NML 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 NML 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 NML 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 is 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 NML elasticities are then

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

and cannot be chosen arbitrarily.

8

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

Equation (1) is slightly more flexible than the standard NML. 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.

4.2   The Distance Metric

Brands of a differentiated product can compete along many dimensions in product–characteristic space. For empirical tractability, 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 indirect–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. This utility function is in Gorman polar form and can therefore be aggregated to obtain brand-level demands. In particular, aggregation does not depend on the distribution of unobserved consumer heterogeneity or of income. However, simplicity in aggregation is not obtained costlessly, since one must assume that all consumers have the same constant marginal utility of income.

Since the indirect–utility function is quadratic, the aggregate–demand equations

are linear in normalized prices and income,[11] and brand sales can be written as

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

where $B = [b_{ij}]$ is an arbitrary $n \times n$ symmetric, negative–semidefinite matrix, and prices, $p = (p_1, \ldots, p_n)^T$, 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, $b_{ij} = g(d_{ij})$.[12] 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 NML, 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, an hypothesis being that, for example, the demand for high–alcohol beers might be systematically less elastic than that for low. In contrast to the NML, however, the off–diagonal elements, $b_{ij}, j \neq i$, that determine substitutability between brands depend on distance measures, an 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 $z_i = (x_i^T, y^T)$. If in addition there are unobserved brand and market variables ξ with typical element ξ_i, (3) can be written in matrix notation as

$$q = Z\beta + Bp + \xi, \qquad (4)$$

---

[11] By Roy's identity, demands must be divided by the marginal utility of income, which is a price index here that differs over time but not by consumer. In a short time series, this number can be set equal to 1.

[12] One can restrict the indirect–utility functions, the individual demand equations, or the aggregate demand equation (as is done here). The three possibilities are equivalent. In particular, the restrictions on aggregate demand do not imply that individuals have similar preferences.

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 NML, 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

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

A feature that distinguishes the DM from the NML 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, and a third that is due to collusion.[13]   Finally, the sum of the first two is the unilateral effect, whereas the third, if positive, is the coordinated effect.

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

---

[13] This decomposition is due to Nevo (2001).

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.[14]

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 *status–quo* partition). The margins that are associated with this game express the market power 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).

---

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

The third step is to estimate 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.[15]   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 that issue by examining differences in observed and predicted margins, which I call *excess margins*. Let the vector $\theta_\mathcal{K}$, denote those differences,

$$\theta_\mathcal{K} = L_o - \tilde{L}_\mathcal{K}. \tag{8}$$

One can think of the θs as parameters that measure the extent of the deviation from the null hypothesis of static Nash–equilibrium behavior for a given game.[16]

Since there are n first–order conditions (7) and n parameters, $\theta_{\mathcal{K}i}$, one can solve the first–order conditions to obtain a vector of implicit excess margins, $\tilde{\theta}_\mathcal{K}$. Alternatively, one can assume that excess margins depend on brand and firm characteristics, $\theta_{\mathcal{K}i} = f_\mathcal{K}(x_i)$, and estimate the first–order conditions. Let the econometrically estimated excess margins be $\hat{\theta}_\mathcal{K}$. $\tilde{\theta}_\mathcal{K}$ and $\hat{\theta}_\mathcal{K}$ can be used to test hypotheses concerning any static game (e.g., $E(\theta_\mathcal{K}) = 0$) as well as to assess systematic deviations from the Nash–equilibrium of that game. I use excess margins to assess the *status–quo* game.

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. Esti-

---

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

[16] Excess margins, however, are not the same as market–conduct parameters or 'conjectures,' which are obtained by allowing $\partial p_i / \partial p_j$ to be nonzero (see, e.g., Bresnahan 1989).

mation 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.[17]

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.[18]

6.2   Excess Margins

The first–order conditions (7) can be solved for a vector of excess margins, θ, that can be modeled as functions of the brand characteristics.[19]   In the absence of information on functional form, a simple linear relationship is used,

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

where x_i is the vector of observed firm and brand characteristics, and ϕ_i is an unobserved variable that affects markups. With the DM specification, the equation that is estimated is

---

[17] 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.

[18] 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.

[19] One can also estimate marginal costs econometrically, or one can retrieve the implicit marginal costs that rationalize the equilibrium assumption. This was done in an earlier version of the paper (Slade 2001). The principal conclusions do not change, however, and so those results are not shown.

$$Y \equiv L_o - E^{-1}e = X\gamma + \phi, \tag{10}$$

where e is an n vector of ones, $e = (1, 1, \ldots, 1)^T$, the $n \times n$ matrix, E, is defined by

$$E_{ij} = \begin{cases} -\epsilon_{ji}\frac{R_j}{R_i} & if \quad j \in \tilde{k}_i \\ 0 & if \quad j \notin \tilde{k}_i, \end{cases}$$

and $R_i = p_i q_i$. In other words, E is a weighted elasticity matrix where the weights are ratios of revenue shares. As is standard, ϕ_i 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 excess–margin function is substituted into (10), 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.[20] 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

---

[20] 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. The first set of instruments exploits the panel nature of the data. Specifically, I assume that prices in region one are valid instruments for prices in region two and vice versa. The brands in my sample are not brewed locally and thus have a common cost component. Moreover, it is likely that the error term in the demand equation principally reflects local promotional activity that is apt to be uncorrelated across regions.

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 instrument $W^1 p_{-r}$ 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. In particular, anything that is unobserved and is common to both demand equations is problematic. For example, national advertising can cause demand shocks to be correlated across regions. Fortunately, national advertising creates less

---

[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] Unobserved product characteristics can also cause problems. When one has a long time series, it is common to include brand fixed effects to capture unobserved time–invariant product characteristics.[23] When the data span a long time period, however, product characteristics such as package design are not likely to be invariant. This means that, even when brand fixed effects are included, price instruments are suspect. In my application, the time series is short, and brand characteristics are less apt to change. However, with a short time series, the inclusion of product fixed effects is impractical. For this reason, it is advisable to experiment with other sets of instruments and to develop formal tests of instrument validity.

The second set of instruments exploits the conditional–independence assumption; when $E[\xi_i|X] = 0$, rival characteristics can be used to form instruments. The use of rival characteristics here, however, is somewhat different from their use in much of the differentiated–products literature (e.g., Berry, Levinsohn, and Pakes 1995). There, rival characteristics are used as instruments for own price. Here, they are principally used as instruments for rival prices. Instruments are formed by premultiplying vectors of characteristics by weighting matrices. To illustrate, suppose that x^1 is the vector of alcohol contents of the brands (a column of the matrix X) and that W^1 is the same–product–type matrix. Then the instrument W^1x^1 has as ith element the average alcohol content of rival brands that are of the same type as i.

As with the first set of instruments, there are circumstances under which instruments formed from rival characteristics will not be valid. This will be the case, for example, if rival characteristics should enter the demand equation directly.

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

---

[22] Figures taken from the MMC cost study indicate that advertising and marketing expenditures are less than one percent of sales.

[23] See, e.g., Hausman, Leonard, and Zona 1994 and Nevo 2001.

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.

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. Not all variables were available for all observations, however. When data for an observation were incomplete, the corresponding observation was also dropped in the other region.[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

---

[24] The test and discussion appear in Pinkse and Slade (2002) and are reproduced here.

[25] Observations were dropped in both regions 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.

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.

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. Unfortunately, three brands — Tetley, Boddingtons, and John Smiths — have both real and keg–delivered variants. Since it is not possible to obtain separate data on the two variants, the classification that is used by StatsMR was adopted. Dummy variables that distinguish the four product types are denoted PROD_i, $i = 1, \ldots, 4$. The product types also form the basis of the groups for the NML specifications, and those specifications include an explanatory variable LGSHARE, 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, 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 = 1 for mul-

---

[27] The outside good here consists of all other products that individuals purchase. When the NML is estimated, the log of the share of the outside good is moved to the right–hand side of the equation, where it is captured by market (i.e., region/time-period) fixed effects.

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

tiples, the regions of the country, REGL = 1 for London, the nationals, NAT = 1 for national brewers, and the time periods, PER1 = 1 for the first period, 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 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, $PRICE_i \times ALC_i$.

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. 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.[31]

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, lagers have the highest alcohol contents, and stouts have the highest coverage rates. The last 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. However, the volume statistics must also be viewed with caution, 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 estab-

---

[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.

[30] The details of this construction can be found in Pinkse and Slade (2002).

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

lishments. In addition, both prices and volumes are higher in London.

7.2 The Metrics

Using the same data, Pinkse and Slade (2002) experimented 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, were also used.

They found 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, a 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×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.[32] 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.[33] 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

---

[32] Managed public houses are owned and operated by a brewer.
[33] Some company data are also available in a form that does not identify the companies.

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.

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.

If average variable costs in brewing are constant, these numbers are marginal costs, but if average variable costs vary with output, they either over or underestimate marginal costs. However, it is difficult to predict the direction of the bias. Indeed, due to the presence of fixed costs, average variable costs can increase with output even when there are increasing returns.

8   Empirical Results

8.1   Demand

Nested Logit Demand

A number of specifications of the nested–logit demand equation (2) are shown in table 4. The first is obtained by setting $\sigma = 0$ and $\alpha_i = \alpha$, which yields the standard logit (numbers 1 and 2). The second has $\sigma > 0$ and $\alpha_i = \alpha$, which is

the standard nested logit (numbers 3–9). The third is also nested but allows α to vary by brand. In other words, $\alpha_i = \alpha(x_i)$ (numbers 10 and 11). Furthermore, some equations are estimated by ordinary least squares (OLS) whereas others are estimated by instrumental variables (IV) with PRICE and LGSHARE endogenous.[34] The method of estimation is indicated in the last column of the table.

All specifications contain the log of coverage, LCOV, and time-period, regional, and product–type fixed effects. The specifications differ by the presence or absence of ALC, PREM, and brewer fixed effects. In particular, because ALC and PREM both measure a brand's strength, some specifications contain both of those variables, whereas others contain only one. Finally, with the equations that are nested, brands are partitioned into four groups according to product type — lager, stout, keg ale, and real ale.

A number of empirical regularities appear in the table. First, the coefficient of PRICE (the slope of the demand equation) is negative as predicted only when ALC is the sole measure of alcoholic strength.[35] Moreover, when this coefficient is negative, it is not significant at conventional levels. Second, when the same specification is estimated by OLS and IV, the slope is more negative with the IV estimation, as one would expect. Third, comparing ordinary and nested logits, the slope is always more negative with the former. Nevertheless, as discussed below, due to the presence of the group–share variable, LGSHARE, demand is more elastic with the latter. Finally, with the nested specifications, the coefficients of LGSHARE are positive, less than one, highly significant, and assume magnitudes that imply high within–group correlation of tastes.

The second half of table 4, which contains specifications in which prices are interacted with brand characteristics, shows estimates of $-\alpha$, the constant coefficient of price, and of $-\alpha_i = -\alpha(x_i)$, the slope of the demand equation evaluated at the mean of the product characteristics.[36] This section of the table shows that when prices are interacted with characteristics, slopes are neither more negative nor statistically more significant than when they are not.

When calculating elasticities and markups, only specifications with negative slopes are used, since otherwise demand would slope upwards and brands would be comple-

---

[34] Both sets of instruments were used in the estimations reported in the table.

[35] This is also true of other specifications of the logit that are not shown in the table.

[36] The characteristics that are included in this specification are the same as with the DM specifications that are presented in table 5. In particular, ALC appears in the intercept term, whereas PREM is interacted with price. For this reason, there are only two entries in this portion of the table.

ments. Furthermore, to give the NML the benefit of the doubt, the specification with the largest estimates of α and σ (# 7) is used.

Brand own–price elasticities are calculated holding the prices of all other brands constant. First, consider the ordinary logit elasticities. The mean and median own–price elasticities are -0.98 and -0.97 respectively, which are not reasonable values. Indeed, demand for individual brands should be highly elastic, since there are many close substitutes for a given brand. Turning to the NML, the mean own–price elasticity is -3.4, which is also the median elasticity. 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 NML own–price elasticities still seem low.

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 elasticity matrix are equal by assumption. At the mean of the data, this elasticity is 0.0005, which is also very low. The logit–demand specification is therefore not very satisfactory for this application. NML 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, those elasticities are 0.207 for the former and 0.0002 for the latter, an indication that most substitution is within groups, as the estimates of σ already suggested. Finally, there is substantial variation in partial cross–price elasticities across groups. Those 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 5 summarizes the estimated distance–metric–demand equations. The first two specifications in this table, however, are included only for comparison with the logit and NML. 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 NML, the slopes of those equations are not consistently negative and are not significant at conventional levels.

The third equation in table 5 is the DM specification. This equation is divided into three sections: the intercepts, $A_i = \beta^T z_i$, and the own–price terms, b_{ii}, are functions

of the characteristics. 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 terms b_{ij}, $j \neq i$, in contrast, depend on the distance measures.

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.[37] 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 here.

The second part of the equation, which assesses the determinants of brand substitutability, shows that the coefficient of the same–product–type variable, 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 significant 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 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 5 contains OLS estimates of the DM demand equation. The table shows that the OLS estimates of the coefficients of the endogenous variables tend to be smaller in magnitude than the GMM estimates but are similar in significance.

---

[37] Other specifications were estimated, and the principal conclusions were not affected.

Prior to using the DM demand equation, its identification and regularity were assessed. First, I determined that the instruments do indeed explain the endogenous variables. Second, I used the test of correlation between the residuals 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.

Third, I experimented with various sets of instruments. The equations shown in table 5 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.

Finally, the regularity conditions that are implied by economic theory were assessed. Symmetry, which implies that the system of demand equations is integrable, was imposed on the estimation *a priori*. With respect to curvature, all of the eigenvalues of the estimated matrix $\hat{B}$, which is the negative of the Hessian of the indirect–utility function, are nonpositive. This implies that $\hat{B}$ is negative semidefinite and that the indirect-utility function is convex.

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 NML 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 NML 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, which varies only by brand, averages 3.9.

As it is not practical to examine 63 own and approximately 4,000 individual cross–price elasticities, table 6 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 table 6 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 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 that 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.[38]

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 7 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.

---

[38] This is a common problem with flexible functional forms such as the normalized quadratic and the translog.

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 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 8 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* NML price is 224 pence per pint, which can be compared to the observed mean of 168. NML *status–quo* prices are thus on average 33% higher than observed prices, which means that, with the NML, observed behavior appears to be substantially more rivalrous than Bertrand. Furthermore, NML *status–quo* margins at the mean of the data are over 50%, which can be compared to the observed margins, that average 30%.

One must conclude that either this market is very competitive or that the NML underestimates price sensitivity in the beer data. Although it is possible that the market is very competitive, the second alternative seems more plausible. The end result is that the estimated NML elasticities are relatively small in magnitude and insignificantly different from zero. If those estimates were taken seriously, Bertrand decision makers would choose prices that are substantially higher than the ones that are observed.

Table 8 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).[39] 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 market power and results in margins of about 23%. Finally, *status–quo* prices and

---

[39] 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.

margins are extremely close to observed prices and margins.

Using the DM demand equation, one can decompose the observed margins of 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 (2001) 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 above decomposition is not a statistical test. One can use the implicit excess–margins, $\tilde{\theta}$, to test for coordinated effects more formally. First consider the implicit estimates obtained from the NML demand equation, $\tilde{\theta}_{NML}$. Both the mean and the median of those parameters are -0.4, and one can reject the hypothesis that the estimates are on average zero. This means that, as before, NML behavior is estimated to be significantly more rivalrous than Bertrand.

Next consider the implicit DM excess margins, $\tilde{\theta}_{DM}$. The mean of those parameters is -0.009, the median is -0.024, and the range is -0.30 to 0.56. Furthermore, the t statistic for the hypothesis that $1/n\Sigma_i\theta_{iDM} = 0$ is -0.66, which means that, on average, predicted and observed margins are equal and Bertrand behavior cannot be rejected.

It appears that although the choice between NML and DM specifications strongly influences the conclusions that can be drawn concerning firm behavior, with neither specification is there any evidence of collusion or 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. One can use an econometrically estimated excess–margin function to assess this possibility.

Given the insignificance of the logit and NML elasticities, joint estimation of a first–order condition with one of those equations does not seem worthwhile. In particular, since the first–stage estimates are imprecise, it is unlikely that the second–stage estimates, which build on the first, would be more accurate. The DM elasticities, however, are precisely measured, and a DM first–order condition is estimated.

Table 9 shows three specifications for a DM excess–margin function. The first contains only the binary variable NAT that equals one if the firm that produces the brand is a national brewer — a member of the hypothesized dominant group. The second two, which differ from one another according to the measure of brand strength that is used, also contain brand characteristics. The table shows that the coefficients of NAT are never significant at conventional levels, which implies that brands that are brewed by the nationals are not more collusively priced and that those firms do not form a dominant group. On the other hand, the second and third specifications show that more popular and higher–strength brands, as well as those that are sold in multiple establishments, are less competitively priced.[40] Finally, there appear to be few systematic differences across product types.

9   Concluding Remarks

No evidence of coordinated effects in UK brewing has been found. In particular, the national brewers do not appear to form a dominant group. Nevertheless, the industry is relatively concentrated and the firms possess substantial market power. The rules of thumb that are used by North American competition authorities would therefore dictate close scrutiny of mergers in the industry. In contrast, unless the merger resulted in a firm with more than 40% of the market, the rules that are used by EU authorities would dictate a more lenient policy. However, using the criterion of unilateral effects, Pinkse and Slade (2002) conclude that a proposed merger between two UK brewing firms (Bass and Whitbread) that would have resulted in a merged firm with less than 40% of the market would have caused substantial price increases.

Differences in policies across jurisdictions are somewhat disconcerting, at least when they lead systematically to different decisions. With respect to mergers, the guidelines that have been adopted by the European Commission (EC) seem to systematically discount mergers in consumer–product industries. There is some evidence, however, that merger policies on the two sides of the Atlantic are converging. Indeed,

---

[40] 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).

in the recent Airtours decision,[41] the EC stated that joint dominance is "not just about tacit collusion." Instead, it is "sufficient for oligopolists to act — independently — in ways which reduce competition."[42] Nevertheless, EC regulations still do not allow mergers to be prohibited unless they create or reinforce dominance. Furthermore, it is too soon to tell if the Airtours decision was an aberration or whether it represents an important new development.

Although brewer characteristics appear not to influence systematic departures from static Nash–equilibrium pricing, brand characteristic do. In particular, brands that have larger market shares, higher alcohol contents, or are sold in multiple establishments are found to be less competitively priced. Contrary to the UK Monopolies and Mergers Commission (1989) claim, however, pricing differences across product types appear to be small. In particular, I find no evidence that lagers are more collusively priced, in spite of the fact that they have higher margins. In other words, higher markups appear to be due to differences in elasticities rather than to less vigorous rivalry.

References Cited

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.

Bresnahan, T. (1989) "Empirical Studies of Industries With Market Power," in *The Handbook of Industrial Organization*, Vol. 1, R. Schmalansee and R. Willig (eds.) Amsterdam: Elsevier.

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.

---

[41] 22.09.99,IV-M.1524.

[42] This statement is quoted in Lexicon (1999b). For further analysis of the Airtours case, see Motta (1999).

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 Ltd. (1999a) "Joint Dominance," Competition Memo, London.

Lexicon Ltd. (1999b) "The Airtours Case," Competition Memo, London.

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.

Motta, M. (1999) "EC Merger Policy and the Airtours Case," European University Institute mimeo.

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. (2002) "Spatial Price Competition: A Semi-parametric Approach," *Econometrica*, 70: 1111-1155.

Pinkse, J. and Slade, M.E. (2002) "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. (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. (2001) "Assessing Market Power in UK Brewing," UBC D.P. # 01-04.

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

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 Estimation

Stage 1: *Estimation of the Demand Equation*

One can write the demand equation as

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

where X_1 is an $n x k_1$ matrix of endogenous and exogenous variables, β is a p_1 vector of parameters, and v is an n vector of errors. Let S be an $n x m_1$ matrix of instruments with $m_1 > p_1$.

The IV estimator of β is the minimum over β of

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

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 J closest neighbors and vice versa, one half if either i is one of j's or j is one of i's J 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 equation.

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

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

and $H_{1\beta}$ is the $n x p_1$ 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 $n x p_2$ matrix of exogenous variables, 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. 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 $n x p_1$ matrix $\partial f_2 / \partial \beta$,

evaluated at $\tilde{\beta}$. Then, if u and v are uncorrelated,[43]

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

Tests of Instrument Validity

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

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

has a limiting $N(0,1)$ distribution (see Pinkse, Slade and Brett 2002).

If one wants to test more than one instrument at a time, it is possible to use a matrix W instead of the vector w 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.

---

[43] 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:

Nested Logit Demand Equations
Dependent Variable: LSHARE

α *Constant*							
Equation	*PRICE (-α)*		*LGSHARE (σ)*	*ALC*	*PREM*	*Brewer Fixed Effects*	*Comment*
1	-0.0040 (-0.7)			yes	no	yes	Logit, OLS
2	-0.0059 (-1.1)			yes	no	yes	Logit, IV
3	-0.0022 (-0.8)		0.822 (29.8)	yes	no	yes	NML, OLS
4	-0.0005 (-0.2)		0.857 (16.4)	yes	no	no	NML, IV
5	0.0056 (2.4)		0.824 (15.4)	no	yes	no	NML, IV
6	0.0028 (1.1)		0.795 (15.0)	yes	yes	no	NML, IV
7	-0.0026 (-1.3)		0.881 (11.6)	yes	no	yes	NML, IV
8	0.0094 (2.9)		0.711 (7.9)	no	yes	yes	NML, IV
9	0.0051 (1.4)		0.729 (8.6)	yes	yes	yes	NML, IV
α *Variable (Price interacted with characteristics)*							
Equation	*PRICE (-α)*	*SLOPE (-α$_i$)[a]*	*LGSHARE (σ)*	*ALC*	*PREM*	*Brewer Fixed Effects*	*Comment*
10	-0.0029 (-1.2)	-0.0024 (-1.0)	0.776 (14.5)	yes	no	yes	NML, IV
11	0.0053 (1.5)	0.0047 (1.4)	0.658 (7.3)	yes	no	no	NML, IV

Ln(s0) in market/time-period fixed effects
Other explanatory variables: LCOV, PER1, REGL, and PROD$_i$, i = 1,...,4

Four groups, lager, stout, keg ale, and real ale
t statistics in parentheses
[a] Evaluated at the mean of the data

Table 5:

Distance-Metric Demand Equations
Dependent Variable: VOL

	1	2	3	4
Estimation Technique	IV	IV	IV	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

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

Table 6:

Own and Cross-Price Elasticities for Selected Brands
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 7:

Summary of Elasticity Estimates
Averages Across Brands

Demand Model	Own-Price Elasticity	Cross-Price Elasticity
Logit	- 0.97	0.0005
Nested Logit	- 3.4	0.0517
Distance Metric	- 4.6	0.0632
AIDS Hausman, Leonard, and Zona (1995)	- 5.0	0.12

Table 8:

Predicted Equilibrium Prices and Margins

Demand Equation	Equilibrium	Mean Price	Standard Deviation	% Difference[a]	Margins (L)
Nested Logit	Status Quo	223.7	40.4	33.3	55.1
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 & Margins		167.8	20.2		29.9

a) (Predicted - Observed)/Predicted.

Table 9:

Excess Margins, Actual - Status Quo
2-Step Estimates Using the DM Demand Equation

Equation	1[a]	2	3
NAT	0.005 (0.2)	-0.002 (-0.1)	-0.003 (-0.1)
LCOV		0.038 (3.7)	0.037 (3.4)
PREM		0.062 (3.8)	
ALC			0.047 (3.6)
PUBM		0.052 (2.9)	0.052 (2.9)
$PROD_1$ (lager)		-0.142 (-4.1)	-0.315 (-4.6)
$PROD_2$ (stout)		-0.028 (-0.6)	-0.218 (-2.3)
$PROD_3$ (keg ale)		-0.113 (-2.8)	-0.270 (-3.7)
$PROD_4$ (real ale)		-0.204 (-6.9)	-0.374 (-6.1)

[a] Includes a constant
t statistics in parentheses
Standard errors corrected for heteroskedasticity and spatial correlation of an unknown form.