ASSESSING MARKET POWER IN UK BREWING

by

Margaret E. Slade

MARCH 2001

Discussion Paper No.: 01-04

DEPARTMENT OF ECONOMICS
THE UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, CANADA V6T 1Z1

http://web.arts.ubc.ca/econ

Assessing Market Power in UK Brewing

Margaret E. Slade

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

March 2001

Abstract:

Market power in UK brewing, an industry that has witnessed a number of recent mergers and has been scrutinized by both UK and EU authorities, is examined. To do this, quantitative methods that can be used to estimate price/cost margins and to decompose those margins into economic factors are assessed. Two classes of demand equations are estimated: the nested logit of McFadden (1978a) and the distance–metric method of Pinkse, Slade, and Brett (1998). Marginal costs are approximated in three ways. Finally, various notions of industry equilibrium are compared. With this application, the most important decision from the point of view of market–power assessment turns out to be the choice of demand model. Different classes of demand equations yield very different predictions concerning elasticities and markups, whereas, within a demand–model class, all methods of assessing market power result in similar predictions concerning industry performance. With a distance–metric demand equation, a static Nash equilibrium in which players set the prices of the brands that they own receives greatest empirical support. Furthermore, both differentiation and fewness endow the firms in the brewing industry with the power to charge prices in excess of marginal costs, but no evidence of collusion is uncovered.

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

Keywords: Market power, market conduct, beer, mergers, differentiated products, multiproduct firms

This research was supported by grants from the Social Sciences and Humanities Research Council of Canada. I would like to thank Joris Pinkse for collaboration on related papers and for comments on this one. I would also like to thank the following people for helpful suggestions: Steven Berry, David Genesove, Lars Mathiesen, Mark McCabe and participants at the EARIE meetings in Lausanne, a departmental seminar at the University of British Columbia, and an NBER workshop in Industrial Organization in Palo Alto, CA.

1 Introduction

Industrial economists are frequently asked to assess the extent of market power that firms in an industry possess. For example, when two or more firms propose a merger, competition authorities must decide if that merger will lead to unacceptable increases in pricing power and thus prices. It is clear that if we cannot evaluate price/cost margins under the existing industry configuration and ownership structure, we have no hope of assessing the situation that will exist under hypothetical future conditions.

Since direct measurement of market power is complex, in evaluating mergers economists have traditionally relied on defining geographic and product markets and calculating market shares and indices of concentration within those markets. This exercise is performed in the hope that the structure of the market will provide useful information concerning the conduct of the firms and thus the industry's performance.

Proxies that are based on market shares can be informative signals of performance in industries where products are homogenous. 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.

When products are differentiated, however, the issue is more complex. Indeed, margins generally depend on product characteristics, and firms that have small shares can set high margins if substitution possibilities are poor. Market–share-based indices are therefore thought to be less relevant when firms produce an entire spectrum of products and market power is determined to a large extent by the degree of differentiation of those products.

In this paper, I assess a number of quantitative methods that can be used to estimate price/cost margins in industries where products are differentiated. Those techniques are used to assess market power in UK brewing, an industry that has witnessed a number of recent mergers of large firms and has been scrutinized by both UK and EU authorities. That industry 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). I therefore emphasize quantitative methods that are best suited to dealing with multiproduct firms that manufacture branded retail products.[1] Furthermore, due to data constraints, I limit attention to horizontal-market power and ignore the possibly complex vertical links that can exist among the firms, which can also influence margins.

It is important to have good estimates of market power, but those estimates by themselves are often insufficient from a policy point of view, since some factors that determine margins (e.g., concentration) can be controlled by competition authorities, whereas others (e.g., differentiation) cannot. I therefore decompose the estimated margins into various economic factors, such as the concentration (or lack thereof) of

[1] Similar techniques can be used when products are differentiated in geographic rather than characteristic space and when they are purchased by firms rather than individuals.

production into the hands of a few firms, the degree of differentiation of the products that those firms produce, and the extent of tacit or overt collusion among industry participants.

The organization of the paper is as follows. The next section describes the UK brewing industry. This industry is characterized by a few large national brewers that market throughout the country, a somewhat larger number of smaller regional brewers, and a still larger number of even smaller brewers that sell only locally. 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 3 discusses the demand side of the market. In contrast to studies of homogeneous–product industries, estimation of demand has been the focus of attention of empirical studies of differentiated products. The reason is that obtaining good estimates of cross-price elasticities is a prerequisite for understanding the extent of the market and the strength of competition across market segments. I describe two very different classes of demand models that have been used to assess market power. The first and more familiar class, which includes the logit and nested logit of McFadden (1974 and 1978a), is particularly important because it has been used extensively by economists to evaluate mergers and other public–policy issues.[2] To illustrate, Werden and Froeb (1994, p. 408) state that "the logit model has direct policy relevance, since the 1992 Horizontal Merger Guidelines use it as the base case for the analysis of mergers in differentiated–products industries." 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 method, is derived in Pinkse, Slade, and Brett (1998) and Pinkse and Slade (2000). With a distance metric, own and cross–price elasticities depend on brand characteristics and a set of measures of the distance between those characteristics.

Section 4 assesses one aspect of the supply side of the market — the estimation of marginal costs. Costs can be obtained from independent sources (e.g., engineering data),[3] they can be estimated implicitly by solving the firms' first–order conditions,[4] or they can be estimated econometrically from those conditions.[5] All three methods are used here.

In section 5, which deals with the second aspect of supply — the market game — the two methods of assessing market conduct that are most prevalent in the literature are considered. With the first, conduct is summarized by a set of parameters that capture deviations from static Nash–equilibrium behavior or, alternatively, from

[2] Recent empirical papers that use a variant of a nested logit to estimate the demand for differentiated products include Trajtenberg (1989), Goldberg (1995), Verboven (1996), Fershtman and Gandal (1998 and 1999), Werden (1999), and Ivaldi and Verboven (2000).

[3] See Nevo (1997), Genesove and Mullin (1998), Wolfram (1999), and Pinkse and Slade (2000).

[4] See, e.g., Nevo (2000)

[5] See, e.g., Berry, Levinsohn, and Pakes (1995) and Petrin (1998).

marginal–cost or monopoly pricing.[6] With the second, *a priori* assumptions are made concerning market conduct (e.g., price-taking behavior), models are estimated under those assumptions, and a choice among assumptions is based on the performance of the estimated models.[7]

Section 6 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 various economic factors. This decomposition, which is due to Nevo (1997), is accomplished by considering games that involve different ownership patterns. For example, suppose that the prior analysis revealed that joint–profit maximization best characterizes the market under consideration. The market power that is due solely to differentiation can be determined by calculating the margins that are associated with a static pricing game in which each brand is owned by a different player. The additional power that is due to fewness or multiproduct production can then be determined by calculating the margins that are implied by a static game under the observed multibrand–ownership pattern. Finally, the residual market power that has not been explained can be attributed to tacit or overt collusion.[8]

Section 7 deals with estimation. Demand equations and first–order conditions are estimated by a two–step generalized–method–of–moments procedure. Since endogenous variables appear on the right–hand-side of these equations, the choice of instruments is discussed, and tests of their validity are derived.

The data, which are discussed in section 8, 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).

Section 9 presents estimates of demand and market power under various specifications for the former and methods of assessing the latter. To anticipate, I find that, with this application, the most important decision from the point of view of market–power assessment is the choice of demand model. Indeed, different classes of models yield very different predictions concerning elasticities and markups, whereas, within a demand–model class, the various methods of assessing marginal costs and market equilibrium result in similar predictions concerning industry performance. Insensitivity to marginal–cost specification, however, is a feature of the application and not a general finding.

Finally, with the distance–metric specification, a static Nash equilibrium in which each player sets the prices of the brands that he owns receives greatest empirical support. Furthermore, both differentiation and fewness endow the firms in this market with the power to charge prices in excess of marginal costs, but no evidence of collusion is uncovered.

[6] Much of this literature is summarized in Bresnahan (1989).

[7] See Bresnahan (1987), Gasmi, Laffont, and Vuong (1992), Feenstra and Levinsohn (1995), and Nevo (1997).

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

2 The UK Brewing Industry

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

Historically, the UK brewing industry developed in a very different fashion from those in, for example, the US, Canada, and France, which were dominated by a few large brewers that sold rather homogeneous national brands of lagers. Indeed, the UK industry, which was relatively unconcentrated, produced a large variety of ales, and regional variation in product offerings was substantial. Moreover, national advertising played a less important role than in many countries. In the last decade, however, there has been a succession of mergers that have increased concentration in brewing and have caused the industry to move towards a more North–American style. Nevertheless, UK brewing is still less concentrated than its counterparts in the US, Canada, and France, where beer tends to be mass produced. It is substantially more concentrated, however, than its counterpart in Germany, where specialty beers predominate.

Among Western countries, the UK is not an outlier with respect to consumption of beer per head or the fraction of sales that are imported. It is very different, however, with respect to the ratio of draft to total beer sales. Indeed, draft sales in the UK, which in 1995 were just under 70% of total sales, accounted for almost three times the comparable percentages in France and Germany and about six times the percentages in North America.[9]

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,[10] are premium beers with relatively high

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

[10] A hybrid is a keg ale that uses a nitrogen and carbon–dioxide mix in dispensing that causes it

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 and is currently around sixty. In addition to incorporated brewers, however, there are approximately 200 microbreweries operating at very small scales. In spite of increases in industry concentration, most brewers are still small, and few produce products that account for more then 0.5% of local markets.

In the 1990s, mergers reduced the number of national brewers from six to four.[11] In addition, in 1997, a very large brewing merger was proposed.[12] This involved the numbers two and three brewers, Bass and Carlsberg-Tetley, and would have created a new firm, BCT, with an overall market share of about 37%. The UK Monopolies and Mergers Commission (MMC) estimated that, after the merger, the Hirshman/Herfindahl index of concentration would rise from 1,678 to 2,332. Having made those calculations, the MMC still recommended that the merger be allowed to go forward.[13] In spite of the MMC's favorable recommendation, the merger did not take place because the president of the Board of Trade did not accept the MMC's advice.

Still more recently, in May of 2000, the world's largest brewer, the Belgian firm Interbrew, acquired Whitbread, the fourth national brewer. This acquisition did not change the number of national brewers in the UK, but it transferred the ownership of some brands. More importantly, in August of the same year, Interbrew also acquired the brewing assets of Bass, which gave it a UK market share of approximately 36%. This time the MMC did not approve the merger. Instead, it recommended that Interbrew be required to divest the UK business of Bass to a buyer approved by the Director General of Fair Trading. At the time of writing, however, it is not clear who will acquire those assets.

The MMC's attitude towards earlier mergers in the brewing industry is puzzling. It seems that either they concluded that brewers had little market power or that large increases in concentration would not change that power. However, their own calculations show that brewing margins were approximately 30%, which is moderately high. For example, margins of approximately 20% are more common in the food sector.

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

---

to be smoother and to more closely resemble a cask ale.

[11] Courage and Grand Met merged to form Courage, Allied Lyons and Carlsberg merged to form Carlsberg-Tetley, and the merged Courage merged with Scottish&Newcastle to form Scottish Courage.

[12] For an analysis of this merger, see Pinkse and Slade (2000).

[13] The one economist on the Commission, David Newbery wrote a dissenting opinion.

3 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 to estimate — the own-price elasticity — and a relatively short time series or cross section can be used for this purpose. 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. For example, when there are $n = 200$ brands, which is not an unusual situation, there are $n^2 = 40{,}000$ own and cross-price elasticities, and it is clear that even with a very large data set, one cannot estimate each as a free parameter.[14] One must therefore place some structure on the estimation.

A number of demand specifications have been develo-ped recently to deal with the problem of an abundance of elasticities. Most of them are based on a random-utility, discrete–choice model.[15] In this paper, I assess a special case of that broad class — the nested logit — primarily because it is easy to estimate and is therefore often used in merger cases. In addition, I discuss the distance-metric method that my coauthors and I have used to evaluate competition in differentiated–product industries. 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, M. It is straight forward, however, to extend the demand models to encompass multiple markets, in which case the size of each market would be an endogenous function of regional variables.

3.1 The Nested Logit

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

Individual h receives utility u_{hi} when consuming product i. The resulting utility function is

$$u_{hi} = \beta^T x_i - \alpha_i p_i + \xi_i + (1 - \sigma)\mu_{hi} + \zeta_{hg}. \tag{1}$$

[14] A functional form that places no restrictions on the elasticities is said to be 'flexible'. See Diewert (1971).

[15] See Berry (1994) for detailed discussions of this class of demand model.

In equation (1), x_i is a vector of observed characteristics of product i, p_i is that product's price, ξ_i is an unobserved (by the econometrician) product characteristic, and μ_{hi} captures individual deviations from the mean valuation of product i. The unobserved taste parameter ζ is common to all products of group g but specific to individual h. Finally, ξ_i is assumed to be mean independent of x_i, and μ_{hi} is assumed to have an extreme-value distribution.

The parameter $\sigma (0 \leq \sigma \leq 1)$ measures the within-group correlation of utility. The ordinary logit, which is a special case of (1), 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).

It is well known that the estimating equation corresponing to (1) is

$$ln(s_i) - ln(s_0) = \beta^T x_i - \alpha_i p_i + \sigma ln(\bar{s}_{i/g}) + \xi_i, \tag{2}$$

where $\bar{s}_{i/g}$ is brand i's share of the group g to which it belongs. This equation, which can be easily estimated by an instrumental–variables technique, is highly restrictive. To illustrate, let ε_{ij} denote the price–elasticity of demand, $(\partial q_i/\partial p_j)(p_j/q_i)$. With the MNL–demand equation, these elasticities take the form

$$\varepsilon_{ii} = \alpha_i p_i [s_i - 1/(1-\sigma) + \sigma/(1-\sigma)\bar{s}_{i/g}], \tag{3}$$

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

Equation (1) is slightly more flexible than the standard MNL. In particular, the coefficient of p_i, α_i, is allowed to depend on the characteristics of that product. In other words, $\alpha_i = \alpha(x_i)$. Nevertheless, as equation (3) shows, the cross–price elasticity between i and j is independent of i. Clearly, the substitution patterns that are implied by (3) are unappealing. In particular, 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.

To obtain more general substitution patterns, it has become standard to estimate random–coefficients demand equations that are derived from a random–utility model in which the coefficients, β, vary by consumer (see, e.g., Berry, Levinsohn, and Pakes 1995 and McFadden and Train 1996). I, however, choose a different approach, an approach that is computationally less burdensome but still offers considerable modeling flexibility.

A second class of demand model has also been used to obtain more flexible substituion patterns. This class involves continuous choice in a setting in which consumers have a systematic taste for variety. The distance–metric demand model developed below more closely resembles the Almost–Ideal–Demand System (AIDS) with multistage budgeting that is used by Hausman, Leonard, and Zona (1994). Their specification, however, is empirically intractable when there is a large number of brands in any group, as is the case here.[16] In particular, the product groups must make sense from an economic point of view, which implies that the number of brands per group cannot be chosen arbitrarily.

3.2 The Distance Metric

Brands of a differentiated product can compete along many dimensions in product–characterisitc space. For empirical tractbility, however, one must limit attention to a small subset of those dimensions. Nevertheless, it is not desirable to exclude possibilities *a priori*. The distance–metric (DM) demand model, which is developed in Pinkse, Slade, and Brett (1998) and Pinkse and Slade (2000), 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 model 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.

A feature that distinguishes the DM from the MNLis 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. To achieve this dependence, one must interact prices with characterisitcs. When there is a large number of brands, however, it is clearly impractical to include all rival prices on the right–hand side of the estimating equation and even less practical to interact those prices with own and rival characteristics. In what follows, I describe how one can use measures of distance between brands to condense this information.

The DM model is based on a normalized–quadratic utility function (Berndt, Fuss, and Waverman 1977 and McFadden 1978b) in which the prices of the differentiated products as well as individual incomes have been divided (or normalized) by the price of the outside good.[17] 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.

Let $\tilde{p}_i$ be the nominal price of the ith brand, $\tilde{y}_h$ be the nominal income of the hth individual, and $p = p_0^{-1}\tilde{p}$ and $y_h = p_0^{-1}\tilde{y}_h$ be relative (or normalized) prices and incomes, where $\tilde{p} = (\tilde{p}_1, \ldots, \tilde{p}_n)^T$. The normalized–quadratic utility function that is used here is

$$u_h(p, y_h) = -[a_{0h} + a_h^T p - p_0 y_h(\gamma_0 + \gamma^T p) + \frac{p_0}{2} p^T B_h p], \tag{4}$$

where each B_h is an arbitrary $n \times n$ symmetric, negative-semidefinite matrix.

[16] Hausman, Leonard, and Zona have five brands per group, whereas I have, for example, 50 lagers in a group, 25 in each of two sorts of establishments. Moreover, specifications with finer groupings were rejected.

[17] This and other flexible functional forms were developed to model substitution among broad classes of products, such as food, clothing, and housing. However, they can equally well be used to model substitution among brands of a differentiated product.

This utility function is flexible in prices. In other words, it is a second-order approximation that places no restrictions on substitution possibilities between brands of the differentiated product. Moreover, the function is in Gorman polar form and can therefore be aggregated to obtain brand-level demands.[18] Finally, aggregation does not require one to specify the distribution of unobserved consumer heterogeneity.

Unfortunately, the matrix $B_h = [b_{hij}]$ alone has $n(n+1)/2$ parameters, and equation (4) must be simplified. I assume that a_{hi} and b_{hii}, $i = 1, \ldots n$, the linear and quadratic coefficients of p_i in the utility function, are individual–specific functions of the characteristics of brand i, $a_{hi} = a_h(x_i)$ and $b_{hii} = b_h(x_i)$. 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_h, which are the coefficients of the interaction terms, $p_i \times p_j$ for $j \neq i$, are assumed to be functions of a vector of measures of the distance between brands in some set of metrics, $b_{hij} = g_h(d_{ij})$. For example, when the product is beer, the measures of distance, or its inverse closeness, might be alcoholic-content proximity and dummy variables that indicate whether the brands belong to the same product type (e.g., whether both are stouts) and whether they are brewed by the same firm. These functions can also vary by individual.

By Roy's identity, individual demands for the differentiated product are

$$q_{hi} = \frac{a_{hi} + \Sigma_j b_{hij} p_j - \gamma_i y_h}{p_0(\gamma_0 + \gamma^T p)}. \tag{5}$$

In equation (5), $p_0(\gamma_0 + \gamma^T p) = \gamma_0 p_0 + \gamma^T \tilde{p}$ is a price index that can, without loss of generality, be set equal to one in a cross section or very short time series. After this normalization, aggregate product demands become

$$q_i = \Sigma_h a_{hi} + \Sigma_j (\Sigma_h b_{hij}) p_j - \gamma_i (\Sigma_h y_h) = a_i + \Sigma_j b_{ij} p_j - \gamma_i y, \tag{6}$$

where

$$a_i = \Sigma_h a_h(x_i), b_{ii} = \Sigma_h b_h(x_i), b_{ij} = \Sigma_h g_h(d_{ij}) = g(d_{ij}), y = \Sigma_h y_h, \tag{7}$$

and y is aggregate income.

Equation (7) shows that the demand intercept, which determines the size of the market for brand i, depends on product and market characterisitcs, x_i and y.[19] This assumption transforms the model from one in which consumers demand brands into one in which they demand the characteristics that are embodied in those brands, as in a hedonic study. If the number of characteristics is less than the number of brands, the dimensionality of the problem is reduced.

[18] See Gorman (1953, 1961) or Blackorby, Primont, and Russell (1978) for discussions of the conditions that are required for consistent aggregation across households.

[19] A market is a regional/time-period pair with zero cross-price elasticities across markets. More generally, in addition to per–capita income, market characteristics might include, for example, population and the number of firms that supply the product in the region. The form of the utility function, however, (i.e., the fact that it is in Gorman polar form) implies that only average consumer characteristics matter.

The own–price elasticity of demand also depends on the characteristics, the hypothesis being that, for example, the demand for high–alcohol beers might be systematically less elastic than that for low. Off–diagonal elements, $b_{ij}, j \neq i$, in contrast, which determine substitutability between brands, depend on distance measures, the hypothesis being that, for example, brands that have similar alcohol contents might be closer substitutes. Finally, notice that although the aggregate function $g(.)$ is common to all brands, this does not imply that household–substitution patterns are identical (i.e., $g(.)$ is a sum of functions that vary by household).

Let X be the matrix of observed brand and market variables with typical row $(x_i^T, y)^T$. If in addition there are unobserved brand and market variables ξ with typical element ξ_i, (6) can be written in matrix notation as

$$q = \beta_0 + X\beta + Bp + \xi, \tag{8}$$

where β_0 is a vector of intercepts that are treated as random effects and β is a vector of parameters that must be estimated. The matrix B has two parts: b_{ii} is a parametric function of the characteristics, $b_{ii} = \lambda^T x_i$, and $b_{ij} = g(d_{ij}), i \neq j$. The function $g(\cdot)$ can be estimated by parametric or semiparametric methods. Finally, the random variable ξ, which captures the influence of unobserved product and market characteristics, can be heteroskedastic and spatially correlated. However, as with the MNL, ξ is assumed to be mean independent of the observed characteristics, $\mathrm{E}[\xi_i|X] = 0$.[20]

The own and cross-price elasticities that are implied by equation (8) are

$$\varepsilon_{ii} = \frac{p_i \lambda^T x_i}{q_i} \quad \text{and} \quad \varepsilon_{ij} = \frac{p_j g(d_{ij})}{q_i}. \tag{9}$$

As with the MNL, DM own–price elasticities depend on prices, market shares, and product characterisitcs. However, cross–price elasticities can be modeled very 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) that depend on the prices and locations of all brands.[21] Finally, hybrid models that include more than one distance measure are possible. Clearly, substitution possibilities are more flexible with a DM model than with a logit, nested or otherwise. Furthermore, the distributional assumptions that underlie the construction of the DM are not restrictive.

[20] Although not always reasonable, virtually all researchers in the area make this assumption (e.g., Berry Levinsohn, and Pakes, 1995, p. 854). I test the assumption below.

[21] The elasticities in (9) always depend on the prices of all brands, since q_i does. However, one can construct distance measures that also depend on all prices.

4 Marginal Costs

An appropriate demand specification is one of the building blocks that is used to assess market power. In addition, equilibrium calculations require estimates of marginal cost, $c_i,\ i = 1, \ldots, n$. There are at least three methods that can be used to obtain marginal-cost estimates.

With the first method, researchers obtain independent information (e.g., engineering data) concerning the cost function. This information is then substituted into the first-order conditions that are solved to obtain equilibrium prices and markups. The advantage of this method is its simplicity. The disadvantage is that, unless the cost data are very accurate, it is difficult to distinguish between average–variable and marginal costs. Exogenous estimates of marginal costs are denoted $\check{c}_i,\ i = 1, \ldots, n$.

With the second method, which involves estimating marginal costs implicitly, researchers assume that firms are engaged in a particular game (e.g., Bertrand) and write down the first-order conditions for that game. Those conditions typically include marginal costs as well as demand parameters. One can therefore substitute the estimated demand parameters into the first-order conditions and solve those conditions for implicit costs. In other words, implicit costs are the estimates that rationalize the observed prices and the equilibrium assumption. To illustrate, consider a simple pricing game where each firm produces one brand and sets its price so that $(p_i - c_i)/p_i = -1/\varepsilon_{ii}$. There are n first-order conditions of this form that can be solved for the n unknowns, c_i. With the example, given estimates of demand elasticities and observed prices, an implicit estimate of marginal cost is $p_i(1 + 1/\hat{\varepsilon}_{ii})$. This method is valid if the firms are indeed playing the assumed game. If they are playing a different game, however, the estimates of marginal cost so obtained are biased. The implicit estimates are denoted $\tilde{c}_i,\ i = 1, \ldots, n$.

The third method involves estimating marginal costs econometrically from first-order conditions that can involve a vector of market-conduct parameters θ.[22] Those parameters summarize the outcome of the market game without specifying that game. To illustrate, the above first-order conditions can be altered so that $(p_i - c_i)/p_i = -(1+\theta_i)/\varepsilon_{ii}$. If $\theta_i > 0 (< 0)$ the outcome is more (less) collusive than the outcome of a static pricing game, whereas the finding that $\theta_i = 0$ is evidence of Bertrand behavior. One can replace c_i in this equation with a function of standard cost variables, such as factor prices and product attributes, and estimate the marginal-cost function along with θ. A disadvantage of the third method is that the instruments that are used to identify the demand equation are often correlated with conduct. When they are, parameter estimates are biased, especially when deviations from static equilibrium are large (see Corts 1999). Corts also shows that this specification confounds average and marginal collusiveness. Econometric estimates of marginal costs are denoted $\hat{c}_i,\ i = 1, \ldots, n$.

It is also possible to estimate the marginal–cost function econometrically after having set $\theta = 0$ (or to any other value).[23] With this alternative, the above iden-

[22] This literature is summarized in Bresnahan (1989).

[23] Examples include Berry, Levinsohn, and Pakes, (1995) and Petrin (1998).

tification problem disappears. However, the procedure is subject to the criticisms of method 2. Furthermore, since demand and cost are estimated jointly, if the equilibrium assumption is incorrect, the misspecificaton contaminates the estimates of demand.

In the application, I assess whether the method that is chosen to estimate marginal cost is an important determinant of predicted prices and markups.

5 Assessing Market Conduct

The term market conduct has been used in a variety of ways. However, it generally refers to the degree of interaction among firms and the extent to which they recognize their interdependence. In this paper, the term is used in a precise way to denote deviations from Nash equilibria of static pricing games. With this in mind, a number of possible pricing games are described.

Suppose that there are K sellers of the differentiated product and that player $k,\ k=1,\ldots,K$, controls a set of prices p_i with $i \in \tilde{k}$, where $\mathcal{K} = [\tilde{1}, \tilde{2}, \ldots, \tilde{K}]$ is a partition of the integers $1,\ldots,n$. Let $p_{\tilde{k}}$ be the set of prices that k controls. Assume also that sellers of the differentiated product play a game, whereas the outside good is competitively supplied. For given $\mathcal{K}$ and prices p_j with $j \notin \tilde{k}$, player k chooses $p_{\tilde{k}}$ to

$$max_{p_{\tilde{k}}}\ \pi_k = \Sigma_{i\in\tilde{k}}[p_i q_i - C_i(q_i)] - F_k, \tag{10}$$

where $C_i(\cdot)$ is the variable-cost function for brand i and F_k is the fixed cost for firm k.[24]

The n first-order conditions for this maximization, $\partial\pi_k/\partial p_i = 0,\ k=1,\ldots,K,\ i\in\tilde{k}$, can be used in a number of ways. In order to nest various models, I employ a fairly general specification for those conditions, one that includes a vector of market–conduct parameters, $\theta_i,\ i=1,\ldots,n$.[25] In other words there is one θ for each brand. The ith first-order condition is

$$q_i + \Sigma_{j\in\tilde{k}_i}\Big\{(p_j - c_j)\Big[\frac{\partial q_j}{\partial p_i} + \theta_i \Sigma_{m\notin\tilde{k}_i}\frac{\partial q_j}{\partial p_m}\Big]\Big\} = 0, \tag{11}$$

where $\tilde{k}_i$ is the element of the partition to which p_i belongs, and $c_j = C_j'(q_j)$, is marginal cost.

Equation (11) can be estimated jointly with the demand equation. When $\theta_i = 0$ for all i, it nests the following well-defined models:

---

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

[25] The term 'conjecture' is often used because the parameters are often interpreted as conjectured responses, $\Theta_{ji} = E(\partial p_j/\partial p_i),\ j \notin \tilde{k}_i$. This interpretation, however, is often not useful, and it might be better to think of the Θs as misspecification parameters that measure the extent of the deviation from the null hypothesis of static Nash–equilibrium behavior. Nevertheless, the first–order conditions are obtained by allowing these partial derivatives to be nonzero and then setting $\Theta_{ji} = \theta_i$ for all j.

i) *Bertrand behavior with single-product firms:* $K = n$ *.*
ii) *Bertrand behavior with multiproduct firms:* $K < n$.[26]
iii) *Joint-profit-maximizing behavior:* $K = 1$.

Each of the above ownership structures results in a set of parameter restrictions that can either be imposed on the estimation *a priori* or tested. For example, one can estimate the three models and use non-nested hypothesis tests to choose the one that best fits the data.[27]

It is also possible to use equation (11) and independent data on marginal costs to obtain implicit market–conduct parameters. Specifically, given a partition, $\mathcal{K}$, and observed price and exogenous cost vectors, p and $\check{c}$, one can solve the first–order conditions for the vector of implied market–conduct parameters, $\tilde{\theta}$, that rationalize the observed prices. One can then use the implicit market–conduct parameters to test the Bertrand (or any other) assumption. Indeed, since the implicit estimates are random variables, it is possible to test if they are, on average, zero (or some other value).[28]

In addition, given any partition, $\mathcal{K}, \theta$ can be estimated econometrically as a set of free parameters in the first–order conditions. If one has a sufficiently long time series, it is possible to exploit the temporal variation in brand elasticities to identify the conduct of each brand. With a single cross section or short panel, however, one cannot estimate n different θs. Nevertheless, one can structure the estimation so that θ_i is a function of a small set of variables that determine conduct, such as the product attributes.[29] When θ is estimated, models can result whose interpretation is not straight forward. However, if the hypothesis $\theta \leq 0$ is rejected, the models are interpreted as collusive, where collusion can be either tacit or overt.[30] Econometrically estimated market-conduct parameters are denoted $\hat{\theta}_i,\ i = 1, \ldots, n$.

Finally, one can use equation (11) to assess market conduct by solving for the marginal costs that are implied by any of the models, i) to iii) and comparing the implicit cost vector, $\tilde{c}$, to independent data on costs, $\check{c}$.[31] Like $\tilde{\theta}_i$, the implicit estimates $\tilde{c}_i,\ i = 1, \ldots, n$, are random variables that can be used to perform hypothesis tests.

[26] There are of course many partitions with $K < n$.

[27] This method of assessing market conduct is used in Bresnahan (1987), Gasmi, Lafont, and Vuong (1992), and Feenstra and Levinsohn (1995).

[28] This procedure is used in Pinkse and Slade (2000).

[29] Conduct can be estimated jointly with marginal cost, or the estimation can make use of exogenous cost information.

[30] Tacit collusion is defined as a Nash equilibrium of a dynamic game that is preferred by the players to the Nash equilibrium of the static game. Overt collusion, in contrast, involves explicit agreement.

[31] This procedure is used in Nevo (1997), where margins rather than costs are compared.

6 Decomposing Market Power

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. Following Nevo (1997), it is possible to decompose those margins into three components: one that is due to differentiation, one that is due to size or multibrand production, and the third that is due to collusion. This procedure involves solving first–order conditions to obtain equilibrium prices of different games and calculating the associated margins. For example, given a partition $\mathcal{K}$ and a set of marginal costs, one can solve the first-order conditions (11) for equilibrium prices and margins, $\tilde{p}_{\mathcal{K}i}$ and $\tilde{L}_{\mathcal{K}i} = (\tilde{p}_{\mathcal{K}i} - \check{c}_i)/\tilde{p}_{\mathcal{K}i}$, of the corresponding Bertrand game. Moreover, with the DM demand equation, this calculation normally involves only matrix inversion.[32]

Let the vector of Lerner indices evaluated at $\check{c}$ and observed prices be $\check{L}$. The first step in the decomposition is to evaluate the market power that results from differentiation alone. One does this by solving game i) of the previous subsection. 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 size, or equivalently, fewness or multibrand ownership. To do this, one solves game ii) of the previous subsection, where the partition $\mathcal{K}$ with $K < n$ corresponds to the observed brand–ownership pattern. The margins that correspond to 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). This step is particularly important for evaluating mergers, which primarily involve increases in firm size rather than changes in the degree of differentiation of their products.[33]

In the final step, differences between $\check{L}$ and the margins of the second game, if positive, are 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. Finally, in evaluating mergers, in contrast to direct changes in market power that are due to increased size, so–called unilateral effects, it is difficult to predict indirect changes that could result from changed collusiveness.

[32] This is true as long as there is an interior solution with $p_i \geq 0$ and $q_i \geq 0$ for all i. When these constraints are violated, a more complex procedure must be used to solve the constrained game. See Pinkse and Slade (2000) who use a solution that is developed in Slade (1995).

[33] One can use this procedure to compare any two ownership structures and thus to evaluate mergers that involve firms that already produce many brands (see Pinkse and Slade 2000).

7 Estimation

7.1 Demand

The demand equations (2) and (8) contain endogenous right–hand–side variables and are therefore estimated by instrumental-variables (IV) techniques. Estimation of the nested logit is entirely straight forward.

The DM equation (8) can be estimated by either parametric or semiparametric methods. With the parametric estimator, $g(\cdot)$ is a parametric function of the distance measures d_{ij}. The semiparametric estimator is described in Pinkse, Slade, and Brett (1998) and Pinkse and Slade (2000) and is therefore discussed only very briefly here. The principal concern is with the estimation of the matrix B, where $b_{ii} = \lambda^T x_i$ and $b_{ij} = g(d_{ij}),\ j \neq i$. The semiparametric estimator for β_0, β, λ, and g, which is a series expansion, is based on the traditional parametric IV estimator. In earlier papers, we demonstrate that β_0, β, λ, and g are identified and that our estimator is consistent. We also derive the limiting distributions of $\hat{\beta}_0, \hat{\beta}, \hat{\lambda}$, and $\hat{g}$ and show how their covariance matrix can be estimated.

Our covariance–matrix estimator, which can be used with either the parametric or the semiparametric version of the model, 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.[34]

The issue of identification is complicated by the fact that the X variables can enter both the linear part of the model, $\beta_0 + X\beta$, 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 X, 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.[35]

7.2 First-Order Conditions

The first–order condition (11) contains vectors of marginal cost, c, and market–conduct parameters, θ. Each of those vectors can be estimated, either separately or jointly, as functions of exogenous variables. First, consider market conduct. If one has exogenous estimates of marginal costs, one can substitute those estimates, $\check{c}$, into

[34] 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.

[35] 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.

(11) and estimate market conduct as a function of brand and market characteristics. In the absence of information on functional form, a simple linear relationship is used in the application,

$$\theta_i = \gamma^T x_i + \phi_i, \tag{12}$$

where x_i is the vector of observed characteristics, and ϕ_i is an unobserved variable that affects conduct. With the DM specification, the equation that is estimated is $Y_{1i}(\check{c}) = \gamma^T x_i + \phi_i$, where

$$Y_{1i}(c) = \frac{-q_i - \Sigma_{j\in\tilde{k}_i}(p_j - c_j)b_{ji}}{\Sigma_{j\in\tilde{k}_i}\{(p_j - c_j)[\Sigma_{m\notin\tilde{k}_i} b_{jm}]\}}. \tag{13}$$

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

In an analogous fashion, if one has exogenous information about market conduct, one can use that information to estimate a marginal–cost function. Unfortunately, unlike cost information, it is not clear where information on conduct might be obtained. Nevertheless, since it is common in the literature to assume that $\check{\theta} = 0$, I make that assumption to see how it affects the estimates of marginal cost. For this purpose, I use the variable-cost function

$$C(q_i, w_i) = \Gamma_i \vartheta_i (q_i)^\delta, \tag{14}$$

where $\Gamma_i = \exp(\nu^T w_i)$ depends on observed cost factors, w_i, and ϑ_i is an unobserved (by the econometrician) cost factor.[36] The logarithm of marginal cost is then

$$ln(c_i) = \nu_0 + \nu^T w_i + (\delta - 1)q_i + \varphi_i, \tag{15}$$

where $\nu_0 = ln(\delta)$ and $\varphi_i = ln(\vartheta_i)$. Clearly, when δ is one, marginal cost is independent of q. Finally, φ is assumed to be mean independent of x and w.

Estimation of equation (15) proceeds as follows. Define the n vector e and $n \times n$ matrix E by

$$e_i(\theta) = q_i + \Sigma_{j\in\tilde{k}_i}(b_{ji} + \theta_i \Sigma_{m\notin\tilde{k}_i} b_{jm})p_j \tag{16}$$

and

$$E_{ij}(\theta) = \begin{cases} b_{ji} + \theta_i \Sigma_{m\notin\tilde{k}_i} b_{jm} & if \quad j \in \tilde{k}_i \\ 0 & if \quad j \notin \tilde{k}_i. \end{cases}$$

Then (15) can be estimated with c replaced by $Y_2(\check{\theta})$, where

$$Y_2(\theta) = E^{-1}(\theta)e(\theta). \tag{17}$$

In principle, it is possible to estimate market conduct and marginal cost jointly as functions of suitable exogenous variables, one set that shifts c but not θ and one that

[36] In practice, x and w might be the same.

shifts θ but not c. With a short time series, however, there is little temporal variation in commonly used exogenous variables, such as factor prices or brand–ownership patterns, and they are therefore not useful instruments. There is, in contrast, substantial cross–sectional variation in brand characteristics, but brand attributes affect both cost and conduct. To illustrate, premium beers, which are more expensive to produce, might have systematically higher margins. The two functions, one for cost and one for conduct, must therefore be identified by some other means. One could rely on an untested functional–form assumption to identify the model, but this practice is dubious. Fortunately, there are other ways to achieve identification. For example, if one is willing to assume that average–variable and marginal costs are equal (i.e., that $\delta = 1$ in (15)), an assumption that can be tested, then coefficients of characteristics that are interacted with prices and demand parameters can be interpreted as part of market conduct, whereas coefficients of characteristics that are interacted only with demand parameters can be interpreted as part of marginal cost.

An iterative procedure is used to estimate c and θ jointly. First equation (12) is estimated with θ replaced by $Y_1(\check{c})$. This yields estimates of θ that are denoted $\hat{\theta}^1$. Next, (15) is estimated with c replaced by $Y_2(\hat{\theta}^1)$ and δ set equal to one, which yields estimates of c that are denoted $\hat{c}^1$. $\hat{c}^1$ is then used in (12) to obtain estimates of θ that are denoted $\hat{\theta}^2$, and so forth. The iterations continue until convergence is achieved.

This procedure is not guaranteed to converge. Furthermore, when it converges to a fixed point, the fixed point need not be unique. In practice, however, it has always converged. Furthermore, the algorithm reached the same fixed point when the order of the estimation was reversed and an initial value of $\check{\theta} = 0$ was used.

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

7.3 Hypothesis Tests

Hypotheses concerning $\hat{\theta}$ and $\hat{c}$ can be tested using standard techniques. In addition, the implicit variables, $\tilde{\theta}$ and $\tilde{c}$, 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.[37] Unfortunately, the notion that the estimates, $\tilde{\theta}_i$ and $\tilde{c}_i,\ i = 1, \ldots, n$, 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 the parameters are performed, the desired quantity is calculated, and a bootstrap distribution is generated.

7.4 The Choice of Instruments

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, X and $\check{c}$, 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, Baker and Bresnahan (1985) suggest using individual–brand cost shifters. Unfortunately, brand–specific cost variables are hard to find. Hausman, Leonard, and Zona (1994), in contrast, 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. Finally, Berry, Levinsohn, and Pakes (1995) 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 second and third suggestions. First, 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.[38] Furthermore, brands that are sold in one region are not substitutes for those that are sold in another. Profit-maximizing decision makers will therefore not coordinate their price choices across regions. Finally, in the majority of establishments (more than 85%), prices are chosen by the retailer (the publican) and not by the manufacturer (the brewer).[39]

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

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

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

[39] This fact suggests modeling the vertical relationships between brewers and retailers. It is standard in the literature, however, to ignore vertical issues. For example, papers that deal with the US auto industry (e.g., Bresnahan 1981 and 1987, Berry, Levinsohn, and Pakes 1995, Feenstra and Levinsohn 1995, Goldberg 1995, and Petrin 1998), where all prices are chosen by retailers, do not consider the implications of this fact. Vertical relationships in the UK brewing industry are analyzed in Slade (1998).

as i.[40]

Unfortunately, there are circumstances under which price instruments, p_{-r} will not be valid. For example, national advertising campaigns could cause the shocks in the two regions to be correlated. Fortunately, national advertising creates less of a problem here than with, for example, US beer, which is much more heavily promoted.[41] Nevertheless, it is advisable to experiment with other instruments.

A second set of instruments exploits the conditional–independence assumption; when $E[\xi_i|X] = 0$, rival characteristics can be used to form instruments. This is done by premultiplying the vectors of characteristics by weighting matrices W. To illustrate, suppose that x^1 is the vector of alcohol contents of the brands (a column of the matrix X). The product W^1x^1 has as ith element the average alcohol content of rival brands that are of the same type as i.[42]

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

One check on instrument validity is to determine if the results obtained are sensitive to the set chosen. In other words, since there are two sets of instruments used in the application — those constructed from prices in the other region and those constructed from characteristics of rival products — it is possible to use each set separately as well as the two together and to compare the implied elasticities and markups.[43]

It is also important to assess the validity of the instruments (i.e., that they are uncorrelated with the errors in the estimating equations) more directly. In particular, the exogeneity of price in the other region is questionable. Moreover, many other instruments are created from that variable and thus might also be suspect. A formal test of exogeneity, one that is valid in the presence of heteroskedasticity and spatial correlation of an unknown form, is derived in the appendix.[44] 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.

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

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

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

[43] This method of assessing instruments is used in Nevo (1997).

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

8 The Data

8.1 Demand Data

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

Most of the demand data were collected by StatsMR, a subsidiary of A.C. Nielsen Company. An observation is a brand of beer sold in a type of establishment, a region of the country, and a time period. Brands are included in the sample if they accounted for at least one half of one percent of one of the markets. There are 63 brands. Two types of establishments are considered, multiples and independents, two regions of the country, London and Anglia, and two bimonthly time periods, Aug/Sept and Oct/Nov 1995. There are therefore potentially 504 observations. Some brands, however, were not sold in a particular region, time period, and type of establishment. When this occurred the corresponding observation was dropped in both regions of the country.[45] 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.[46] Independents, in contrast, can be public houses, clubs,[47] or bars in hotels, theaters, cinemas, or restaurants. Independent establishments are usually not tied to a particular brewer.

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

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.[48]

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

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

[46] 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.

[47] A club is an organization where consumption of liquor is restricted to members and their guests.

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

content.

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

There are ten brewers in the sample, the four nationals, Bass, Carlsberg–Tetley, Scottish Courage, and Whitbread, two brewers without tied estate,[49] Guiness and Anheuser Busch, and four regional brewers, Charles Wells, Greene King, Ruddles, and Youngs. Brewers are distinguished by dummy variables, $\text{BREW}_i,\ i = 1, \ldots, 10$.

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

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

Finally, a variable, NCB, was created as follows. First, each brand was assigned a spatial market, where brand i's market consists of the set of consumers whose most preferred brand is closer to i in taste space than to any other brand.[50] 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).[51]

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, $\text{PRICE}_i \times \text{ALC}_i$.

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.[52] Whereas coverage would be endogenous in a longer-run model,

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

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

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

[52] This assumption is tested below.

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.[53]

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 more expensive than lagers, which are more expensive than ales, and that lagers have the highest alcohol contents, followed by stouts and then ales. In addition, average coverage is highest for stouts. This statistic, however, is somewhat misleading, since it is due to the fact that Guiness is an outlier that is carried by a very large fraction of establishments. Finally, cask–conditioned ales have higher prices and sell larger volumes than keg ales. The volume statistics must be viewed with caution, however, since some of the most popular brands have keg variants.

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

8.2 The Metrics

Using the same data, Pinkse and Slade (2000) experiment with a number of metrics or measures of similarity of beer brands. These include several discrete measures: same product type, same brewer, and various measures of being nearest neighbors or sharing a market boundary in product–characteristic space. These discrete measures are local in the sense that they set most cross–price elasticities to zero *a priori*. Two continuous measures of closeness, one in alcohol–content and the other in coverage space, are also used. These measures are global (but not symmetric) in the sense that they imply that all cross–price elasticities are positive (but not equal).

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

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

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

8.3 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.[54] 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.[55] The MMC used volume weights to calculate average unit costs, where the volumes were based on the sales of each brand in managed houses.

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

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

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

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

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

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

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

9 Empirical Results

9.1 Demand

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

Logit Demand

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

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

Brand own–price elasticities are percentage changes in the shares of single brands when the brand's price increases by 1%, holding the prices of all other brands constant. At the mean of the data, the logit own–price elasticity is -0.115, which is not a reasonable value. Indeed, demand for individual brands should be highly elastic, since there are many close substitutes.

Brand cross–price elasticities are percentage changes in the shares of single brands when the price of a single rival brand increases by 1%, holding own price and the prices of all other rivals constant. With the logit, these elasticities vary only by brand, since off–diagonal entries in a column of the logit–elasticity matrix are equal by assumption. At the mean of the data, this elasticity is 0.0001, which is also very low. The logit–demand specification is therefore not very satisfactory for this application.

Nested–Logit Demand

Table 5 summarizes the estimated nested–logit–demand equations. In the first half of 5A, α, the coefficient of price, is constant. In contrast to the logit, however, brands are partitioned into four groups according to product type — lager, stout, keg ale, and real ale. As with the logit, most of the estimated coefficients of PRICE are positive (4 out of 6). The magnitudes of the negative estimates, however, are greater, and their significance is somewhat higher. Nevertheless, when they are negative, the estimated price coefficients are still not significant at conventional levels. The estimated coefficients of LGRSHARE — the log of a brand's share of the group to which it belongs — in contrast, are positive, less than one ($0 < \sigma < 1$), and significant at conventional levels.

The second half of table 5A shows specifications in which prices are interacted with brand characteristics. The table shows estimates of α, the constant coefficient of price, as well as $\alpha_i = \alpha(x_i)$, the slope of the demand equation evaluated at the mean of the product characteristics.[56] This section of the table shows that when prices are interacted with characteristics, slopes are neither larger in magnitude nor statistically more significant.

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

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

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

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

Distance–Metric Demand

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

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

The third equation in table 6 is the DM specification.[57] This equation is divided into three sections: the intercept terms, $A_i = \beta_0 + \beta^T x_i$, and the own-price terms, b_{ii}, are functions of the characteristics, x_i. The characteristics in b_{ii} however, have been interacted with price to allow the own-price elasticities to vary with those characteristics. The rival-price term $b_{ij},\ j \neq i$, in contrast, is a function of the distance measures, d_{ij}, which determine the cross-price elasticities.

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

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

The second part of the equation, which assesses the determinants of brand substitutability, shows that the coefficient of the same–product–type rival–price measure, RPPROD, is both positive and significant. 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 but not significant at conventional levels. 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 this rivalry is less pronounced.

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

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

As a check on the DM demand equation, its identification was assessed. First, I used the test of correlation between the residuals in that equation and various groups

[57] Pinkse and Slade (2000) present many specifications of DM equations, both parametric and nonparametric, that include various combinations of distance measures. To avoid duplication, only their final specification is shown here, which is parametric.

of instruments, as is discussed in the appendix. This process uncovered no evidence of endogeneity. In particular, when price in the other region was investigated by itself, the p value for the test was 0.20, and when the instruments as a group were assessed, the value was 0.38.

Second, I experimented with various sets of instruments. The equations shown in table 6 were estimated using both sets of instruments — those constructed from prices in the other region and those constructed from characteristics of rival brands. When the demand equation was estimated with either set by itself, results were similar.[58]

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

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

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

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

In addition to identifying the type of each brand, the first row of the table shows the brand's alcohol content and number of neighbors, where a neighbor shares a market boundary with the given brand, and markets are delineated in characteristic space (see subsection 8b).

The table shows that there is substantial variation in own–price elasticities, and that most of the magnitudes are plausible. In particular, if one ranks the the reciprocals of the (absolute values) of the own–price elasticities and ranks the estimated price/cost margins, the rankings are very similar. Nevertheless, the own–price elasticity of the small–sales brand, Greene King IPA, seems unrealistically high. This is due to the fact that elasticity estimates are inversely related to sales. Furthermore,

[58] To illustrate, averages of estimated own–price elasticities that were obtained using just price (p_{-r}) instruments, just characteristics (X_{-i}) instruments, and both sets (p_{-r} and X_{-i} instruments) are -6.1, -5.4, and -4.6, respectively. Averages of estimated total cross–price elasticities (see below) using the same sets of instruments are 4.8, 4.2, and 3.9, respectively.

the own–price elasticity for Guiness is very low, which is due to the fact Guiness has very high sales (as well as the fact that it has few neighbors). It therefore seems likely that the model over (under) estimates magnitudes of elasticities for brands with very small (large) market shares.[59]

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

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

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

9.2 Marginal Costs

Any of the estimated demand equations can be used to evaluate marginal costs. However, since the logit–elasticity estimates are so poor, only the nested-logit and distance–metric equations are used for this purpose. For each specification, marginal cost can be estimated implicitly or econometrically. Recall that with the first method, the first–order conditions (11) are solved for a vector of implicit marginal–cost parameters, $\tilde{c}$, and with the second, the marginal costs in (11) are replaced with functions of the brand characteristics before the equation is estimated. With both methods, θ is set equal to zero.

All results reported below pertain to London. Results for Anglia are similar.

Table 9 summarizes the implicit marginal–cost estimates. A different set of estimates is obtained for each demand and equilibrium specification, and each set can be compared to the exogenous cost estimates, $\check{c}$, that are shown in the last row of the table. The mean $\check{c}$ is 129.1 and the standard deviation is 5.2. The true standard deviation of marginal cost, however, is probably substantially higher than 5.2, since $\check{c}$ varies only by product type. However, it is unlikely to be greater than 20.2, which is the standard deviation of price.

Nested–Logit Marginal Costs

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

The first row in table 9 shows the implicit estimates of marginal cost that were obtained using the MNL demand equation and the *status–quo* game, where the *status quo* is a Bertrand game with the multiproduct brand–ownership structure that existed when the data were collected. The mean of the MNL *status–quo* estimates is 79.0, which is substantially lower than the mean of the exogenous estimates. Moreover, the hypothesis that the 95% confidence interval for the mean of $\tilde{c}_{MNL}$ includes 129.1, the mean of $\check{c}$, is rejected at conventional levels.

The finding that MNL implicit costs are lower than $\check{c}$ is driven by the relatively low MNL elasticities. Indeed, with low elasticities, Bertrand decision makers have reason to choose high markups. In order to rationalize the observed prices, therefore, implicit costs have to be lower.

The last column in table 9 shows percentage differences between means of implicit and exogenous costs. The MNL implicit costs under the *status–quo* equilibrium are 39% lower than the exogenous estimates.

Given the unsatisfactory nature of the MNL elasticities and their implications for behavior, econometric estimation of an MNL marginal–cost function does not seem worthwhile. In particular, since the first–stage estimates are not significantly different from zero, it is unlikely that the second–stage estimates, which build on the first, would be more accurate.

Distance–Metric Marginal Costs

The remaining rows in table 9 summarize implicit marginal costs that were calculated using the DM demand equation. Four hypothetical equilibria are considered. The first, which is also the simplest, is marginal–cost pricing, and the others are Bertrand games with different assumptions concerning brand ownership. With the first of these games, each brand is owned by a different decision maker, whereas with the second, decision makers control the prices of several brands, as in the data. Finally, the fourth row corresponds to joint–profit maximization by a single player who chooses the prices of all brands.

The implicit–marginal–cost mean falls and the standard deviation rises as one moves down the rows of the table. The variation in mean is due to the fact that, as the market becomes less competitive, optimal markups become higher. Observed prices, however, remain unchanged. Implicit costs must therefore fall to rationalize the prices.

It is interesting to note that DM implicit costs under monopoly are higher on average than MNL implicit costs under the *status quo*. This means that, faced with the MNL elasticities, Bertrand players would choose higher markups than a monopolist facing the DM elasticities.

The table shows that the DM *status–quo* game is associated with the implicit costs that are closest to the exogenous costs. Indeed, the percentage difference between the *status–quo* and exogenous means is less than one. With the MNL demand function, in contrast, the *status–quo* game produces marginal–cost estimates that are substantially lower than the exogenous estimates. The choice of demand specification is thus crucial here.

One can test the equilibrium assumptions more formally by seeing if the mean of $\check{c}$ lies inside the 95% confidence interval for each of the DM implicit–cost means. When this was done, marginal–cost pricing and joint–profit maximization were rejected at 1%, whereas single–product pricing was rejected at 10%. The *status–quo* game, in contrast was not rejected at any reasonable level.

Turning to the econometric estimates of the marginal–cost function, table 10 shows estimates that use the DM demand equation. The first two specifications were obtained by setting $\theta_i = 0$ for all i, as is common in the literature.[60] The two specifications that are shown differ in the way that alcohol strength is measured. The qualitative results, however, do not depend on the specification. If one takes the estimates at face value, they show that costs are higher when a product is popular, when it is sold in a multiple establishment, when it is of higher strength, and when it is a lager or a stout. One must, however, interpret the coefficients with caution. To illustrate, consider the positive coefficient of PUBM. Table 2 shows that prices are higher in multiple establishments. Nevertheless, if products were priced less competitively in such establishments, it would be possible for costs to be equal across establishment types.[61] Indeed, it is difficult to disentangle the separate effects of c and θ.

9.3 Market Conduct

As with marginal costs, it is possible to use any of the demand equations to estimate market conduct implicitly ($\tilde{\theta}$) or econometrically, ($\hat{\theta}$). As before, however, only the MNL and DM equations are in fact used. The estimates of conduct that are reported in this subsection use the exogenous marginal costs, $\check{c}$.

Nested–Logit Market Conduct

Both the mean and the median of the implicit MNL market–conduct parameters, $\tilde{\theta}_{MNL}$, are -0.6. Such low estimates imply that the market is very competitive. In particular, one can reject the hypothesis that $E(\theta_{MNL}) = 0$, which means that behavior is significantly more rivalrous than Bertrand.

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

Distance–Metric Market Conduct

The mean of the implicit DM market–conduct parameters, $\tilde{\theta}_{DM}$, is 0.014, the

[60] Equations that include VOL were also estimated, but, since the hypothesis that $\delta = 1$ (see equation 15) could not be rejected at conventional levels, those equations are not shown.

[61] Unfortunately, the MMC study did not consider costs in independent establishments and there is therefore no exogenous basis for comparison.

median is -0.011, and the range is -0.8 to 2.[62] Furthermore, the p value for the hypothesis that $1/n\Sigma_i\theta_{iDM} = 0$ is 0.46, which means that Bertrand behavior cannot be rejected. Therefore, as with marginal cost, the choice between MNL and DM specifications for demand strongly influences the conclusions that can be drawn concerning firm behavior. With neither specification, however, is there any evidence of collusion.

In contrast to the MNL estimates, the DM elasticities are precisely estimated. Estimating a DM market–conduct function therefore seems reasonable. The first two columns of table 11 show two specifications for this function that differ according to the measure of brand strength that is used. The table shows that, regardless of specification, more popular, higher–strength, and multiple–establishment beers are less competitively priced,[63] but there is no evidence of conduct differences across product types.

9.4 Joint Estimates of Cost and Conduct

The last two columns of tables 10 and 11 contain specifications in which cost and conduct are estimated jointly. Comparing these equations to the ones that use exogenous information on cost or conduct shows that a number of coefficients change sign and/or significance. For example, when θ is set equal to zero (the first two columns of table 10), one is led to believe that popular brands are more costly to produce. When cost and conduct are estimated jointly, however, not only is the significance of the coefficients of LCOV reduced, the signs are also reversed. It now appears that popular brands are not more costly to produce; instead they are less competitively priced. Higher strength brands continue to be more costly as well as less competitively priced. The higher cost of selling in multiple establishments, however, is reduced, and its significance disappears. Finally, the coefficients of the product–type fixed effects show that, when cost and conduct are estimated jointly, the evidence that lagers are more costly to brew disappears. Instead, lagers now appear to be less competitively priced. The higher cost of brewing stouts, in contrast, remains significant.

It seems that, although the estimates of cost and conduct that are obtained using exogenous information are accurate on average, there are systematic differences across brands that are not captured by those specifications. Furthermore, the differences are more striking with the marginal–cost estimates. In particular, when a brand is less competitively priced but θ is set equal to zero, the higher price is attributed to higher cost.

9.5 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

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

[63] 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).

margins can be decomposed in the manner that is described in section 6. In particular, one can assess the relative contributions of differentiation, fewness, and collusion to the determination of market power.

Table 12 summarizes the equilibrium prices and margins that are associated with various models, where margins or Lerner indices are calculated using exogenous cost estimates $\check{c}$ 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. The table shows that the mean *status–quo* MNL price is 245 pence per pint, which can be compared to the observed mean of 168. MNL *status–quo* prices are thus on average about 50% higher than observed prices, which is just another indication that, with the MNL demand equation, behavior is estimated to be substantially more rivalrous than Bertrand. Furthermore, MNL *status–quo* margins at the mean of the data are nearly 90%, which can be compared to the observed margins of 30%. One must again conclude that either this market is very competitive or that the MNL model of demand underestimates price sensitivity in the beer data.

The table 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).[64] The first results in prices that are on average 40 pence per pint lower than observed prices and in margins that are everywhere zero. Single–product prices, in contrast, which average 159 pence per pint, are only 9 pence lower than observed prices. This means that differentiation by itself endows the firms in this market with substantial pricing power and results in margins of over 23%. Finally, *satatus–quo* prices and margins are extremely close to observed prices and margins, which should not come as a surprise, given that *status–quo* behavior could not be rejected using earlier tests.

Using the DM demand equation, one can decompose the observed margins of 30% into two 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 fewness 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. While the final conclusion might have been unanticipated, it is similar to results reported in Nevo (1997) 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.

[64] 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.

10 Concluding Remarks

The choice of demand model appears to be crucial in the assessment of market power in the UK brewing industry. In particular, with a logit or nested logit, behavior in the industry is estimated to be substantially more rivalrous than Bertrand. With a distance–metric demand model, in contrast, the Bertrand assumption cannot be rejected. Furthermore, these conclusions do not depend on the type of test that is used. To illustrate, within a demand class, tests that involve recovering market–conduct parameters and those that involve matching implicit costs to exogenous costs yield the same conclusions. Indeed, holding the demand model constant, the qualitative results from all tests of firm behavior are similar.

The finding that the choice of demand model is crucial here can be contrasted with results reported in Genesove and Mullin (1998), who note that their assessment of market conduct in the US sugar industry is not sensitive to the specification of demand. They, however, only consider different functional forms. The difference between the two classes of demand models that are considered here is not just a matter of functional form. Instead, the variables that appear on the right–hand side differ across classes. Specifically, with the logit and nested logit, cross–price elasticities between brands i and j are independent of i, whereas with the distance metric, they depend on the characteristics of both brands.

It is reassuring that, in contrast to demand, the methods that are used to assess market conduct and power are not crucial in the application, since all of the methods that are applied here are flawed. To illustrate, as discussed in section 4, each of the commonly used techniques of estimating marginal cost has its problems, and it is therefore difficult to choose among them. Furthermore, industrial economists differ in their views concerning the appropriateness of the methods that have been used to evaluate market conduct. Indeed, some are purists who claim that only well specified game–theoretic models should be estimated, whereas others are happy to allow market–conduct parameters to take on intermediate values. Fortunately, in the UK brewing application, this distinction turns out to be unimportant. With applications where the Bertrand assumption is inappropriate, however, the treatment of marginal costs could be crucial.

With this data, the marginal–cost estimates that are obtained under the Bertrand assumption are reasonable on average. Nevertheless, there appear to be systematic differences in costs across brands that are not captured. For example, when Bertrand behavior is assumed, popular brands appear to have higher costs. When costs and conduct are estimated jointly, in contrast, the cost difference disappears. Instead, popular brands appear to be less competitively priced.

Unfortunately, the problems that are associated with identifying marginal costs in the absence of reliable information on conduct (and vice versa) are ubiquitous in the differentiated–products literature. Nevertheless, it is common to assume that firms are playing a static pricing game and to use that assumption to estimate marginal costs without testing its validity. If the Bertrand assumption is not appropriate, the cost estimates can be very inaccurate.

Turning to the implications for the UK brewing industry, I find that most of the market power that the firms possess is due to differentiation, a factor that competition authorities cannot control directly. Nevertheless, firm size, a factor that can easily be influenced by competition policy, also contributes substantially to brewer margins. The relatively lax position that the UK Monopolies and Mergers Commission took towards brewer mergers in the 1990s was therefore questionable.

In this industry as well as many others, the vertical structure seems to play an important role in determining margins. Indeed, I find that prices are systematically higher in multiple establishments, which are public houses that are owned by brewers or retail chains. Unfortunately, an in–depth assessment of this issue requires cost data that vary by establishment type. In the absence of such data, the problem of isolating the effects of contract form (i.e., who owns and who operates the retail establishments) appears to be empirically intractable.

Finally, I began by noting that economists are frequently asked to assess market power as an aid to determining public policy towards mergers and other antitrust issues. Unfortunately, my findings are not very encouraging. In particular, decisions concerning mergers must often be reached in a matter of months. A premium is therefore placed on quantitative techniques that are simple to use and require little data. Overly simple models, however, can yield highly inaccurate predictions.

References

Baker, J.B. and Bresnahan, T.F. (1985) "The Gains from Merger or Collusion in Product-Differentiated Industries," *Journal of Industrial Economics*, 33: 427-444.

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. (1994) "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25: 242-262.

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*, Amsterda m: North Holland.

Bresnahan, T. F. (1981) "Departures from Marginal-Cost Pricing in the American Automobile Industry," *Journal of Econometrics*, 17: 201-277.

Bresnahan, T. F. (1987) "Competition and Collusion in the American Auto Industry: The 1955 Price War," *Journal of Industrial Economics*, 35: 457-482.

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

Corts, K.S. (1999) "Conduct Parameters and the Measurement of Market Power," *Journal of Econometrics*, 88: 227-250.

Diewert, W.E. (1971) "An Application of the Shephard Duality Theorem: A Generalized Leontief Production Function," *Journal of Political Economy*, 79: 481-507.

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

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

Fershtman, C. and Gandal, N. (1998) "The Effect of the Arab Boycott on Israel: The Automobile Market," *RAND Journal of Economics*, 29: 193–214.

Fershtman, C., Gandal, N., and Markovich, S. (1999) "Estimating the Effect of Tax Reform in Differentiated Product Oligopolistic Markets," *Journal of Public Economics*, 74: 151-170.

Gasmi, F., Laffont, J.J., and Vuong, Q. (1992) "Econometric Analysis of Collusive Behavior in a Soft-Drink Market," *Journal of Economics and Management Strategy*, 1: 277-311.

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

Goldberg, P.K. (1995) "Product Differentiation and Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica*, 63: 891-951.

Gorman, W.M. (1953) "Community Preference Fields," *Econometrica*, 21: 63-80.

Gorman, W.M. (1961) "On a Class of Preference Fields," *Metroeconomica*, 13: 53-56.

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.

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.

McFadden, D. and Train, K. (1996) “Mixed MNL Models for Discrete Response,” University of California, Berkeley, mimeo, forthcoming in *Journal of Applied Econometrics*.

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

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

Nevo, A. (1997) “Measuring Market Power in the Ready-to-Eat Cereal Industry, University of California, Berkeley, mimeo, forthcoming in *Econometrica*.

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

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.

Petrin, A. (1999) “Quantifying the Benefits of New Products: The Case of the Minivan,” University of Chicago mimeo.

Pinkse, J., Slade, M.E., and Brett, C. (1998) “Spatial Price Competition: A Semiparametric Approach,” http://bee.econ.ubc.ca/Research/pricecompweb.pdf.

Pinkse, J. and Slade, M.E. (2000) “Mergers, Brand Competition, and the Price of a Pint,” http://bee.econ.ubc.ca/Research/beerweb.pdf.

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

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

Trajtenberg, M. (1989) “The Welfare Analysis of Product Innovations with an Application to Computed Tomography Scanners,” *Journal of Political Economy*, 97: 444-479.

Verboven, F. (1996) “International Price Discrimination in the European Car Market,” *RAND Journal of Economics*, 27: 240–268.

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.

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

APPENDIX

Estimation and Testing

The Two-Step GMM Estimation

Stage 1: GMM Estimation of the Demand Equation

One can write the demand equation as

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

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

The GMM estimator of β is the minimum over β of

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

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

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

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

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

Stage 2: Estimation of the First-Order Condition

One can write the first-order condition as

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

where y is an endogenous variable, X_2 is an nxp_2 matrix of exogenous variables,[65] γ is a p_2 vector of parameters, and u is an n vector of errors. Since this equation

[65] When $\delta \neq 1$ in the cost function, instruments are used.

is exactly identified, the GMM estimates, $\tilde{\gamma}$, can be obtained by simply solving the moment conditions.[66] The standard errors of $\tilde{\gamma}$, however, must be corrected to reflect the fact that β was estimated in a prior stage.

Let Ω_2 be defined like Ω_1 with $\hat{v}$ replaced by $\hat{u}$ and $H_{2\beta}$ be the nxp_1 matrix $\partial f_2/\partial\beta$, evaluated at $\tilde{\beta}$. Then, if u and v are uncorrelated,[67]

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

Tests of Instrument Validity

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

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

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

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

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

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

Table 1:

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

1A: Three Major Groups

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

1B: Ales

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

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

Table 2:

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

2A: London

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

2B: Anglia

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

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

Table 3:

Brewer Costs and Margins [a)]

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

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

Source: MMC (1989)

Table 4:

Logit Demand Equations
IV Estimates

Dependent Variable: LSHARE

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

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

Table 5:

Nested Logit Demand Equations
IV Estimates

Dependent Variable: LSHARE

5A: Various Specifications

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

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

Table 5:

Nested Logit Demand Equations
IV Estimates

Dependent Variable: LSHARE

5B: Final Specification

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

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

Table 6:

Distance-Metric Demand Equations[a)]

Dependent Variable: VOL

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

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

Table 7:

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

Evaluated at observed prices and quantities

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

Table 8:

Summary of Elasticity Estimates
Averages Across Brands

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

Table 9:

Implicit Marginal-Cost Estimates

Demand Equation	*Equilibrium*	*Mean*	*Standard Deviation*	*Minimum*	*Maximum*	*% Difference*
Nested Logit	*Status Quo*	79.0	14.6	51.6	132.3	-38.8
Distance Metric	*Marginal-Cost Pricing*	167.8	20.2	117.0	204.5	30.0
	Single-Product Firms	137.9	30.7	44.4	273.8	6.8
	Status Quo	128.0	35.7	35.1	205.5	-0.9
	Joint-Profit Maximization	99.1	41.3	15.1	201.7	-23.2
Exogenous Cost Estimates $\check{c}$		129.1	5.2	124.0	147.0	

Table 10:

Marginal-Cost Equations[a) b)]

2-Step GMM Estimates Using the DM Demand Equation

Equation	1 Using Θ = 0	2 Using Θ = 0	3 Estimated Jointly With Θ	4 Estimated Jointly With Θ
LCOV	0.123 (3.8)	0.120 (3.4)	-0.055 (-1.0)	-0.056 (-1.0)
PREM	0.203 (2.7)		0.283 (3.1)	
ALC		0.118 (2.1)		0.188 (2.0)
PUBM	0.218 (3.1)	0.220 (3.2)	0.103 (1.5)	0.114 (1.5)
$PROD_1$ *(lager)*	0.204 (2.0)	0.234 (2.2)	0.110 (1.2)	0.117 (1.7)
$PROD_2$ *(stout)*	0.366 (2.4)	0.296 (1.9)	0.385 (2.9)	0.314 (2.1)
$PROD_3$ *(keg ale)*	-0.092 (-1.2)	-0.094 (-1.1)	-0.088 (-1.4)	-0.072 (-1.0)
Constant	4.385 (31.3)	3.981 (9.5)	4.705 (26.8)	4.336 (8.4)

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

Table 11:

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

Equation	1 Using č	2 Using č	3 Estimated Jointly With Marginal Cost	4 Estimated Jointly With Marginal Cost
LCOV	0.022 (2.2)	0.021 (2.2)	0.030 (3.1)	0.030 (3.2)
PREM	0.043 (2.0)		0.039 (2.1)	
ALC		0.038 (2.2)		0.047 2.6)
PUBM	0.071 (3.8)	0.072 (4.0)	0.064 (2.9)	0.064 (2.9)
$PROD_1$ *(lager)*	0.001 (0.1)	0.002 (0.1)	0.058 (2.2)	0.058 (2.2)
$PROD_2$ *(stout)*	-0.016 (-0.3)	-0.030 (-0.5)	0.027 (0.7)	0.028 (0.8)
$PROD_3$ *(keg ale)*	0.025 (0.8)	0.037 (1.2)	0.013 (0.4)	0.014 (0.4)
Constant	0.048 (1.2)	-0.090 (-0.7)	-0.132 (-1.1)	-0.312 (-1.5)

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

Table 12:

Predicted Equilibrium Prices and Margins

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

[a)] Calculated using the exogenous cost estimates, $\check{c}$.