



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

European Economic Review 48 (2004) 617–643

EUROPEAN
ECONOMIC
REVIEW

www.elsevier.com/locate/econbase

Mergers, brand competition, and the price of a pint

Joris Pinkse^a, Margaret E. Slade^{b,*}

^a*Department of Economics, The Pennsylvania State University, University Park, PA, USA*

^b*Department of Economics, The University of Warwick, Coventry CV4 7AL, UK*

Received 31 August 2001; accepted 9 September 2002

Abstract

Mergers in the UK brewing industry have reduced the number of national brewers from six to four. The number of brands, in contrast, has remained relatively constant. We analyze the effects of mergers on brand competition and pricing. Brand-level demand equations are estimated from a panel of draft beers. To model brand-substitution possibilities, we estimate the matrix of cross-price elasticities semiparametrically. Our structural model is used to assess the strength of brand competition along various dimensions and to evaluate the mergers. In particular, we compute equilibria of pricing games with different numbers of players.

© 2003 Elsevier B.V. All rights reserved.

JEL classification: L13; L41; L66; L81

Keywords: Beer; Mergers; Brand competition; Differentiated products; Semiparametric instrumental-variable estimation

1. Introduction

Historically, the UK brewing industry was relatively unconcentrated. The last decade, however, has witnessed a succession of successful mergers that have increased concentration in the industry, as well as proposed mergers that, if successful, would have added to that trend. It is thus natural to ask how those mergers have changed both product pricing and product offerings. In particular, the mergers could have resulted in higher prices, a reduction in the number of brands, an increase in brand uniformity, and a move towards competition through national advertising. In this paper, we attempt

* Corresponding author. Fax: 2479-523-032.

E-mail address: m.slade@warwick.ac.uk (M.E. Slade).

to assess the effects of actual mergers and to predict how unsuccessful mergers would have affected the industry. Our formal analysis is limited to price changes, but we analyze other consequences informally.

Our merger evaluations are based on a structural model of demand, cost, and market equilibrium. In particular, we estimate the demand for brands of a differentiated product and we use our estimated demands, together with engineering data on costs, to predict equilibrium prices and margins, which we compare to those that were observed. We then assess the effects of the mergers by solving for equilibria of games with different numbers of players. In other words, changes in market structure – mergers and divestitures – are modeled as changes in the number of decision makers, where each decision maker controls the prices of some set of brands. This means that when two firms merge, some pricing externalities are internalized. Moreover, if brands of the differentiated product are demand substitutes and prices are strategic complements, prices rise after a merger. The question is: by how much do they rise?

The use of econometric and simulation techniques for merger analysis is becoming increasingly common. Indeed, competition authorities on both sides of the Atlantic recognize the advantages of supplementing traditional merger analysis with quantitative assessments. In this paper, we provide such an assessment for the UK brewing industry. We use panel data on all brands of beer that constitute at least one half of one percent of a regional market to estimate the demand for brands of draft beers sold in two regions of the country (Greater London and Anglia), two bimonthly time periods (August/September and October/November 1995), and two types of establishments (multiples and independents).

A number of classes of demand models have been used in merger-assessment exercises, including the logit, nested logit, and random-coefficients discrete-choice specifications.¹ In contrast to those studies, our demand model is a continuous-choice specification. In particular, we extend the spatial model of Pinkse et al. (2002) to encompass substitution among brands, and we estimate the matrix of cross-price elasticities semiparametrically as a function of a number of measures of the distance between brands in product-characteristic space. We believe that our specification combines the simplicity of the logit and nested logit with the flexibility of the random coefficients. Furthermore, it can encompass endogeneity and measurement error in a straightforward fashion.

In the next section, we set the stage by describing the structure of the UK brewing industry and some recent positions taken by UK competition authorities that have affected the industry's horizontal-market structure. This is followed by a review of the merger-evaluation literature and the development of our model of demand, cost, and equilibrium. Our model is then used to assess two mergers – one that was allowed and one that was prohibited. To anticipate, we find that the Scottish–Courage merger had little effect on prices, whereas the proposed merger between Bass and Carlsberg–Tetley would have resulted in more substantial price increases.

¹ For example, Werden and Froeb (1994), Ivaldi and Verboven (2000), and Nevo (2000) evaluate mergers using logit, nested-logit, and random-coefficient demand models, respectively.

2. The UK beer market

To get a feel for the UK industry, it is useful to begin with some international comparisons. 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 about 70% of the total, accounted for almost three times the comparable percentages in France and Germany and about six times the percentages in North America.²

Substantial changes in both consumption and production have occurred in the UK industry in the last few decades. 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 fifty.³ However, in spite of the reduction in the number of brewers, prior to the mergers the UK industry was substantially less concentrated than its counterparts in the US, Canada, and France, where beer tends to be mass produced. Production in Germany, in contrast, where specialty beers predominate, was much less concentrated.

Changes in tastes have also occurred. To illustrate, beers can be divided into three broad categories: ales, stouts, and lagers. Although UK consumers traditionally preferred ales, 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 of those lagers are foreign brands that are brewed under license in the UK.

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, and keg products tend to contain less alcohol than real products. Many of the more recently introduced brands, however, particularly the lagers and hybrid ales, are premium beers with relatively high alcohol contents.

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, compared to many other countries, the UK brewing sector was only moderately concentrated. Recent developments in the industry, however, have resulted in substantial changes in ownership patterns.

In 1990, there were six national brewers: Bass, Allied Lyons, Scottish & Newcastle, Grand Metropolitan (Grand Met), Courage, and Whitbread. Moreover, those six firms had dominated the market for decades. Since 1990, however, a sequence of mergers has increased concentration in brewing. First, three large mergers were approved by

² Only in Ireland was it higher, where draft sales accounted for over 80% of consumption.

³ In addition to incorporated brewers, there are over 100 microbreweries operating at very small scales.

UK competition authorities: Courage and Grand Met merged to form Courage, Allied Lyons and Carlsberg merged to form Carlsberg–Tetley, and Courage and Scottish & Newcastle merged to form Scottish Courage. After 1995, however, horizontal-merger policy became less lenient. Indeed, a proposed merger between Bass and Carlsberg–Tetley was denied, and still more recently, when the Belgian firm Interbrew acquired the brewing assets of Bass and Whitbread, it was ordered to sell its Bass breweries. We discuss the two mergers that we evaluate in greater detail.

The Courage/Scottish & Newcastle merger: The third merger occurred in 1995, when the merged firm Courage combined with Scottish & Newcastle to form Scottish Courage. This event reduced the number of national brewers from five to four and created the largest brewer in the UK with a market share of 28%. In spite of the fact that the majority of the groups that were asked to comment on the merger favored a full investigation by the Monopolies and Mergers Commission (MMC), the Office of Fair Trading did not refer the matter to the MMC. Instead, it allowed the merger to proceed subject to a number of undertakings, all of which involved the relationship between the brewer and its retail outlets.

The Bass/Carlsberg–Tetley merger: A fourth merger was proposed in 1997 but not consummated. This involved the numbers two and three brewers, Bass and Carlsberg–Tetley, and would have created a new firm, BCT, with a market share of 37%. The MMC estimated that, after the merger, the Hirshman/Herfindahl index of concentration (HHI) would rise from 1,678 to 2,332.⁴ Furthermore, it noted that the US Department of Justice’s 1992 Merger Guidelines specify that a merger should raise concerns about competition if the post-merger HHI is over 1,800 and the change in the HHI is at least 50 points. Nevertheless, the MMC recommended that the merger be allowed to go forward.⁵ In spite of the MMC’s favorable recommendation, however, the BCT merger was not consummated because the president of the Board of Trade did not accept the MMC’s advice.

UK competition authorities’ views towards horizontal concentration in brewing seem to have changed over the decade of the 1990s. In particular, early on the Commission was more concerned with vertical relationships in the industry, whereas, by the end of the decade, a concern with horizontal concentration assumed prominence. Was increased concern with horizontal-market power justified? As a first cut to answering that question one can examine the market shares of the firms before and after each merger. That exercise reveals that, with all three consummated mergers, a few years afterwards the merged firm’s market share was less than the sum of the premerger shares. This suggests that increased efficiency did not overwhelm increased market power.

3. Quantitative methods for merger evaluation

Traditional merger analysis places heavy reliance on defining a market and evaluating horizontal concentration within that market. However, many economists believe that

⁴ The HHI is the sum of the squared market shares of the firms, multiplied by 10,000.

⁵ The one economist on the Commission, David Newbery, wrote a dissenting opinion.

concentration indices such as the HHI are less reliable measures of market power in industries in which products are differentiated. In particular, markups can be high in relatively unconcentrated industries if the products that are sold are not close substitutes. The effect of a merger thus depends more on the identities of the merging firms and the brands that they produce and less on their shares of the market.

In evaluating a merger, it is therefore important to assess substitution possibilities. In other words, the matrix of own and cross-price elasticities must be estimated. When products are homogeneous, this task is relatively simple, since a single price prevails and the substitution matrix reduces to a singleton – the industry own-price elasticity of demand. Most consumer products, however, are differentiated, either by their intrinsic characteristics or by the way in which they are marketed. Moreover, it is not uncommon for several hundred brands of a product such as beer or breakfast cereal to be sold. Under those circumstances, there are several hundred own and many thousand cross-price elasticities to determine. Without imposing some structure, this task is empirically intractable. Fortunately, a number of demand models have been proposed to simplify the estimation problem.

For the purpose of merger evaluation, the ideal demand model would possess the following characteristics: it would be (i) flexible in the sense that it would impose no restrictions on the estimated own and cross-price elasticities, (ii) simple, transparent, and easy to estimate using standard computer software, and (iii) capable of handling a large number of brands or products. Unfortunately, no model is ideal, and one must consider tradeoffs among the strengths and weaknesses of each, taking into consideration the features of the market and the data. We highlight some popular demand systems that have been used for merger analysis and discuss the circumstances under which they are more likely to be (un)satisfactory.

Perhaps the simplest demand system that can incorporate a large number of brands is the multinomial logit.⁶ Its advantages are that it is simple to estimate and requires relatively little data. In fact, many applications that involve the logit are calibrated rather than estimated (see, e.g., [Werden and Froeb, 1994](#)). Its principal disadvantage is that it places severe a priori restrictions on the substitution matrix.⁷ Indeed, the logit is a single-parameter model of substitution, and all off-diagonal entries in a column of the substitution matrix are identical.

A slight generalization of the logit, the nested logit ([McFadden, 1974, 1978a](#)), provides greater flexibility without increasing either data or computational requirements substantially.⁸ With the nested logit, brands are grouped into exhaustive and mutually exclusive classes, and substitution is presumed to be easier within than across groups. For example, when the product is beer, the classes might be lagers, ales, and stouts. Although more flexible than the logit, the nested logit is still only a two-parameter model of substitution, and the cross-price elasticities in a column of the substitution matrix take on only two values – one for brands in the same group and one for brands

⁶ This discussion pertains to a logit in which individual demands have been aggregated to obtain brand market shares.

⁷ See [Berry \(1994\)](#) for a discussion of this issue.

⁸ The nested logit has been used to evaluate mergers by [Ivaldi and Verboven \(2000\)](#).

from different groups. The logit, nested or otherwise, is therefore most problematic with respect to characteristic (i) in our list of desirable attributes.

A substantial generalization of the logit, the random-coefficients or mixed multinomial logit (Berry et al., 1995; McFadden and Train, 2000), provides a much more flexible model of substitution. With this model, when the price of one brand is raised, consumers that stop purchasing that brand gravitate towards other brands that have similar characteristics. Increased flexibility is not obtained costlessly, however, since the computational and data requirements are more demanding. The random-coefficients model is therefore most problematic with respect to characteristic (ii) in our list.⁹

The demand models described thus far are based on a discrete-choice assumption under which each consumer purchases at most one unit of one brand of the differentiated product. This assumption is apt to be appropriate for large purchases such as automobiles. In other contexts, however, consumers might have a systematic taste for diversity and might purchase several brands in varying amounts. Some models are therefore based on the opposite assumption—that consumers purchase all brands.¹⁰ Those models draw on the flexible-functional-form literature that was developed to assess substitution among a small number of broad classes of products, such as food, housing, and clothing (e.g., Diewert, 1971). Unfortunately, when a merger involves a large number of brands, demand cannot be modeled in a completely flexible manner.

One method of solving the dimensionality problem is to adopt a multi-stage-budgeting assumption (Gorman, 1971). For example, one can assume that consumers first decide how much to spend on the product (beer). They then decide how to allocate that amount among broad classes of brands (lagers, ales, and stouts), and finally, they decide how much to spend on each brand. This approach is taken by Hausman et al. (1994), who use the almost ideal demand system of Deaton and Muellbauer (1980) to model substitution among brands within each group. The advantage is that within-group substitution is modeled flexibly. The disadvantage is that each group can contain only a small number of brands. The two-stage-budgeting model is therefore most problematic with respect to characteristic (iii) in our list.

4. The beer-brands model

We wish to model a market in which a large number of brands of a differentiated product are sold. Moreover, the assumption that consumers purchase several brands in varying quantities seems more appropriate for beer than the opposite assumption under which they are limited to at most one unit of one brand. We therefore draw on the flexible-functional-form literature. In particular, we modify our earlier work on the demand for differentiated products (Pinkse et al., 2002). In that paper, we consider the case where buyers are downstream firms. Here, we extend that model to encompass buyers that are individuals.

⁹ The random-coefficients model has been used by Nevo (2000) to evaluate mergers.

¹⁰ This assumption is required for consistent aggregation across consumers in order to obtain aggregate demands at the brand level. When consumers are firms or when income effects are not present, the assumption is unnecessary.

The idea behind our demand model is as follows. It is obvious that brands of a differentiated product can compete along many dimensions in product-characteristic space. For empirical tractability, however, one must limit attention to a small subset of those dimensions. Nevertheless, it is not desirable to exclude possibilities a priori. Our demand specification allows the researcher to experiment with and determine the strength of competition along many dimensions and can thus be used to construct an equation that relies on few a priori assumptions. However, although many hypotheses concerning the way in which products compete can be assessed in our framework, only the most important measures will typically be used in the final specification. Like earlier models, ours is not ideal. There are circumstances, however, under which it is the most appropriate choice, and we believe that this is the case for our application.

In what follows, the demand and cost sides of the differentiated-product market are described and equilibria of pricing games are discussed. We consider a situation in which there are n brands of a differentiated product, each of which is produced by one of K firms. The market for the differentiated product is assumed to be imperfectly competitive. All other goods are aggregated into an outside good, which is competitively supplied at a parametric price p_0 .

4.1. Demand

Our demand model is based on a normalized-quadratic indirect-utility function (Berndt et al., 1977; 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. Our individual indirect-utility functions are in Gorman polar form and can therefore be easily aggregated and differentiated to obtain brand-level demands.¹¹ In particular, aggregation does not depend on the distribution of unobserved consumer heterogeneity or of income. However, simplicity in aggregation is not obtained costlessly, since we must assume that all consumers have the same constant marginal utility of income.

Since the indirect-utility function is quadratic, the aggregate-demand equations are linear in normalized prices and income,¹² and brand sales can be written as

$$q_i = a_i + \sum_j b_{ij} p_j - \gamma_i y, \quad i = 1, \dots, n, \quad (1)$$

where $B = [b_{ij}]$ is an arbitrary $n \times n$ symmetric, negative-semidefinite matrix, and normalized prices $p = (p_1, \dots, p_n)^T$ and aggregate income y have been divided by p_0 .

Eq. (1) has more parameters than can be estimated using a single cross section or short panel. It is therefore assumed that a_i is a function of the characteristics of brand i , $a_i = a(x_i)$.¹³ The brand-level demand intercept then depends on product and market

¹¹ See Blackorby et al. (1978) for a discussion of the conditions that are required for consistent aggregation across households.

¹² By Roy's identity, demands must be divided by the marginal utility of income, which is a price index here that differs over time but not by consumer. In a short time series, this number can be set equal to 1.

¹³ One can restrict the indirect-utility functions, the individual demand equations, or the aggregate demand equation (as is done here). The three possibilities are equivalent.

characteristics, x_i and y ,¹⁴ an assumption that 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.

The diagonal elements of B , which determine the own-price elasticities, are also assumed to depend on the characteristics, $b_{ii} = b(x_i)$.¹⁵ For example, the characteristics might be the brand's alcohol content, product type (lager, ale, or stout), and brewer identity. With those characteristics, a hypothesis might be that the demand for high-alcohol beers is systematically less elastic than that for low.

Finally, the off-diagonal elements of B are assumed to be functions of a vector of measures of the distance between brands in some set of metrics, $b_{ij} = g(d_{ij})$, $i \neq j$. For example 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, and whether they are brewed by the same firm. With these measures, a hypothesis might be that brands that have similar alcohol contents are closer substitutes.

Let X be the matrix of observed brand and market variables with typical row $X_i = (x_i^T, y)$. If there are also unobserved brand and regional characteristics u , (1) can be written in matrix notation as

$$q = \alpha + X\beta + Bp + u, \quad (2)$$

where α and β are vectors of parameters that must be estimated. The random variable u , which captures the influence of unobserved product and market variables, can be heteroskedastic and correlated across observations. We assume, however, that the unobserved characteristics are mean independent of the observed characteristics, $E[u_i|X] = 0$. Whereas this assumption is problematic, it is standard in the literature.¹⁶ Moreover, it can be tested, as we do below. Finally, as we are interested in placing as little structure as possible on substitution patterns, we estimate the function $g(\cdot)$ that determines the off-diagonal elements of B by semiparametric methods.

4.2. Marginal costs

A good approximation to demand is necessary but not sufficient for our equilibrium calculations. In addition, we need estimates of marginal costs. There are two common methods of estimating marginal costs econometrically. With the first, researchers assume that a particular game is played (e.g., Bertrand–Nash) and write down the first-order conditions for that game.¹⁷ As those conditions typically include marginal-cost variables, one can estimate the first-order condition along with the demand equation and use the estimated equations to infer costs. This method is efficient if the firms are

¹⁴ A market is a regional/time-period pair with zero cross-price elasticities across markets.

¹⁵ The own and cross-price elasticities are related. However, it would be empirically intractable to make the own-price elasticities depend on all of the distance measures between brand i and all other brands. In the application, we allow own-price elasticities to depend on the number of a brand's neighbors in product-characteristic space, which is a measure of substitution possibilities.

¹⁶ For a discussion of the difficulties involved in relaxing this assumption, see Berry (1994).

¹⁷ See, for example, Berry (1994), Berry et al. (1995), and Petrin (2002).

indeed playing the assumed game. If they are playing a different game, however, the estimates of marginal cost so obtained are biased.

The second method involves estimating marginal costs from first-order conditions that contain a vector of parameters, Γ , that are often called market-conduct parameters. Those parameters summarize the outcome of the game that the firms are playing without specifying that game. In other words, the econometrician takes an agnostic position and lets the data approximate the market outcome. This literature, which is summarized in Bresnahan (1989), has recently been criticized. Indeed, the interpretation and identification of market-conduct parameters that are estimated jointly with costs has been questioned by Corts (1999), who notes that they can be biased, especially when deviations from the null ($\Gamma = 0$) are large.

We do not use either of the econometric techniques to obtain our marginal costs. Instead we make use of a detailed engineering study of beer-production, distribution, and retailing costs by product type that was performed by the UK *Monopolies and Mergers Commission* (1989). We update the MMC estimates to reflect inflationary trends, and we use the updated costs to solve the games. In so doing, we avoid the difficulties that are inherent in the first two approaches. In particular, we eliminate the possibility of contaminating the demand estimates, in which one typically has more confidence, through joint estimation with a misspecified first-order condition. However, we must assume that marginal costs are constant.

4.3. Pricing games

The estimated demands and costs can be used to evaluate the brewers' game. We begin by considering a static pricing game and then discuss how one can test the Bertrand assumption.

Brewers either transfer beer internally to establishments that they operate, in which case the brewer sets the retail price, or they sell beer at wholesale prices to independent or affiliated retailers, in which case the retailer sets the retail price. In the former situation of vertical integration, the joint surplus, brewing plus retailing, is maximized. In the latter situation, the transaction between brewer and retailer is usually not arms-length. Indeed, fixed fees are involved that can be used to distribute the surplus. We assume that nonintegrated brewers and retailers bargain efficiently to maximize the total surplus, given rival prices.¹⁸ The division of that surplus, however, which determines the wholesale price, will depend on the relative bargaining strengths of the two parties. Furthermore, those strengths can change over time.¹⁹ Our assumption is equivalent to having a single party choose the retail price optimally.

Formally, suppose that player k , $k = 1, \dots, K$, controls a set of prices p_i with $i \in \tilde{k}$, where $\kappa = [\tilde{1}, \tilde{2}, \dots, \tilde{K}]$ is a partition of the integers $1, \dots, n$. Let $p_{\tilde{k}}$ be the set of prices

¹⁸ In other words, the vertical game between retailer and brewer is cooperative with side payments, whereas the horizontal retail game is noncooperative.

¹⁹ For example, on average, the retail price of beer in the UK has increased faster than the wholesale price, implying that retailers are now receiving a larger fraction of the total. This fact, however, is consistent with our assumption.

that k controls. For a given partition, κ , player k seeks to

$$\max_{p_{\tilde{k}}} \pi_k(p, \kappa) = \sum_{j \in \tilde{k}} \left\{ (p_j - c_j) \left(A_j + \sum_{m=1}^n b_{jm} p_m \right) \right\} - f_k, \quad (3)$$

where $A_j = \alpha_j + X_j \beta$, c_j is the marginal cost of producing brand j , and f_k is firm k 's fixed cost.

The first-order conditions for this game are linear in prices, and solution of the game involves only matrix inversion. A merger (divestiture) can be simulated by changing the number of players to K' and the partition of the brand space to κ' . One can then evaluate a merger by solving for the Nash equilibrium, $p_{\kappa'}$, of the new game and comparing before and after prices.

The approach just described is valid as long as the firms in the market are engaged in a static pricing game. There are a number of reasons, however, why this might not be the case. Complexities can arise for at least three reasons: the game is repeated, it can be played by agents (retailers), not principals (brewers), and, if there are rigidities such as costs to adjusting prices, Markov-Perfect equilibria can differ from static equilibria.

If the equilibrium assumption is incorrect, the model's predictions will be biased. It is therefore important to assess our assumption. Since we have cost data, we can evaluate the Bertrand assumption by testing for equality between observed margins, $L_{oi} = (p_i - c_i)/p_i$, and the margins that are predicted by the game κ , $L_{ki} = (p_{ki} - c_i)/p_{ki}$. We call the difference between the vectors of observed and equilibrium margins excess margins, and we use θ_{κ} to denote those differences. Finally since the predicted excess margins are functions of the estimated parameters, we can test if they are on average zero.

5. Estimation and testing

Our semiparametric estimator is described in detail in Pinkse et al. (2002) and is therefore discussed only briefly here. Our estimating equation is the demand function (2). This equation contains a vector, d , of measures of distance between brands in different metrics. Specifically, the off-diagonal elements of the matrix B , b_{ij} , $i \neq j$, are a common function $g(\cdot)$ of the distance measures, d_{ij} . The elements of d must be specified by the econometrician; the functional form of g , however, is determined by the data.

We use a series expansion to approximate g and, as is standard, allow the number of expansion terms that are estimated to increase with the sample size. There are three concerns that must be dealt with in deriving our estimator. First, the right-hand-side variables contain prices that are apt to be correlated with u , second, in addition to the error term u , there is an approximation error that is due to neglected expansion terms, and third, the number of instruments must grow as the number of expansion terms increases.

We deal with endogeneity by taking an instrumental-variables (IV) approach. Our concern here is with the choice of instruments. In particular, we need instruments that vary by brand. The exogenous demand and cost variables, X and c , are obvious

choices, and some of them vary by brand. A number of other possible choices have been discussed in the differentiated-products literature. For example, Hausman et al. (1994) assume that systematic cost factors are common across regions so that prices in one region are correlated with those in another but not with shocks to demand. This allows them to use prices in one city as instruments for prices in another. In addition, Berry et al. (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 that we make involve a combination of the two suggestions. First, we use prices in the other region as instruments. The brands in our sample are not brewed locally and thus have a common cost component. Moreover, it is likely that the error in our demand equation principally reflects local promotional activity that is uncorrelated across regions. Unfortunately, however, common demand shocks such as national advertising can render price instruments invalid.²⁰ Furthermore, unobserved product characteristics that are correlated with prices can also cause problems. For this reason, it is advisable to experiment with other sets of instruments and to develop tests of instrument validity, as we do below.

Second, we use rival characteristics to form instruments by multiplying the vectors of characteristics by weighting matrices W , where each W is created on the basis of one or more of the distance measures. 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 and zero otherwise) and that x^1 is the vector of alcohol contents of the brands. The product, $W^1 x^1$, has as i th element the average alcohol content of rival brands that are of the same type as i .²¹ We are thus able to create additional instruments, and our model is overidentified.

Finally, we create additional instruments when the number of expansion terms grows by interacting the exogenous regressors with each term in the series expansion. For example, if the expansion involves polynomials, we interact exogenous regressors with powers of the continuous distance measures.

In Pinkse et al. (2002) we suggest a semiparametric estimator for g and β that is based on the traditional parametric IV estimator. Furthermore, we demonstrate that g and β are identified and that our estimator is consistent, and we derive the limiting distributions of $\hat{\beta}$ and \hat{g} . Finally, we show how their covariance matrix can be estimated. Our covariance-matrix estimator is similar to the one that is proposed in Newey and West (1987) in a time-series context. In particular, observations that are close to one another are assumed to have nonzero covariances, where closeness is measured by one of the distance metrics. 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.²²

²⁰ Although national advertising presents a problem, it is not nearly as prevalent in the UK as in North America. For example, the MMC (1989) presents data in which advertising and marketing expenditures are about 1% of takings.

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

²² 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.

We have assumed that our instruments are uncorrelated with the errors in our estimating equation. The exogeneity of some of them, however, in particular price in the other region, is questionable. Furthermore, other instruments are created from that variable and might also be suspect. We therefore develop a test of exogeneity that is valid in the presence of heteroskedasticity and spatial correlation of an unknown form.

Suppose that the estimating equation is $y = R\gamma + u$ and that $\{(z_i, u_i, Q_i, R_i)\}$ 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 u_i is the error for observation i . For z to be a valid instrument, u and z must be element-wise uncorrelated, i.e. $E(z_i u_i) = 0$. Let $P_Q = Q(Q^T Q)^{-1} Q^T$, $\Omega = \text{Var}(u|R, z, Q)$, $M = I - R(R^T P_Q R)^{-1} R^T P_Q$, $\tilde{V} = z^T M \tilde{\Omega} M^T z$, where $\tilde{\Omega}$ is our estimate of Ω , and \hat{u} be the residuals from an IV estimation using Q (but not z) as instruments. Then, under mild regularity conditions on $\tilde{\Omega}$,

$$\tilde{V}^{-1/2} z^T \hat{u} = \tilde{V}^{-1/2} z^T M u \quad (4)$$

has a limiting $N(0, 1)$ distribution (see Pinkse et al., 2002).

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.

6. Data and preliminary data analysis

6.1. Demand data

Most of the data were collected by StatsMR, a subsidiary of A.C. Nielsen Company. An observation is a brand of draft beer sold in a type of establishment, region of the country, and 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 bi-monthly time periods are considered, August/September and October/November 1995, two regions of the country, London and Anglia, and two types of establishments, multiples and independents. There are therefore potentially 504 observations. Observations were dropped in both regions of the country if one of the numbers (price, quantity, or coverage) was missing for that observation. 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. Independents, in contrast, can be public houses that are not owned by a brewer or chain, or they can be clubs or bars in hotels, theaters, cinemas, or restaurants.

For each observation, we have price, sales volume, and coverage. All are averages for a particular brand sold in a particular region, time period, and type of establishment.

Price, which is measured in pence per pint, is denoted PRICE. Volume, which is total sales measured in 100 barrels, is denoted VOL. Finally, coverage, which is the percentage of outlets that stock the brand, is denoted COV.

In addition, we have data that vary by brand but not by region, establishment type, or time period. Those variables are alcohol content, product type, and brewer identity. 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. We therefore created a dichotomous alcohol-content variable PREM that equals one for premium brands and zero otherwise.

Brands are classified into four product types, lagers, stouts, keg ales, and real ales. Unfortunately, three brands – Tetley, Boddingtons, and John Smiths – have both cask and keg-delivered variants. Since it is not possible to obtain separate data on the two variants of these brands, we adopt the classification that is used by StatsMR. Dummy variables that distinguish the four product types are denoted $PROD_i$, $i = 1, \dots, 4$.

There are ten brewers in the sample, the four nationals, Bass, Carlsberg–Tetley, Scottish Courage, and Whitbread, two brewers without tied estate,²³ Guinness and Anheuser Busch, and four regional brewers, Charles Wells, Greene King, Ruddles, and Youngs. Brewers are distinguished by dummy variables, $BREW_i$, $i = 1, \dots, 10$.

In addition, we created dummy variables that distinguish the establishment types: PUBM for multiples, regions of the country; REGL for London, and time periods and PER1 for the first period.

We also created a number of interaction variables, which are denoted PRXXX, where XXX is a characteristic. For example, $PRALC_i$ is $PRICE_i \times ALC_i$.

All variables that were constructed from prices or volumes are considered endogenous. Coverage, in contrast, is considered to be weakly exogenous. Although coverage would be endogenous in a longer-run model, there is considerable inertia in brand offerings. This is mostly due to the existence of long-term contracts between wholesalers and retailers.

Table 1 shows summary statistics by product type. Table 1A divides observations into the three major product groups: lagers, stouts, and ales, whereas Table 1B gives statistics for the two types of ales. In those tables, total volume is the sum of sales for that product type, whereas average volume is average sales per establishment. 1A shows that, on average, 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 Guinness is an outlier that is carried by a very large fraction of establishments. Finally, cask-conditioned ales have higher prices and sell larger volumes than keg ales. However, the volume statistics must be viewed with caution, since some of the most popular brands have keg variants.

²³ Brewers without tied estate are not vertically integrated into retailing.

Table 1
Summary statistics by product type^a

<i>A: Three major groups</i>				
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

<i>B: Ales</i>				
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

London and Anglia draft beer brands in sample.

^aAverages taken over brands, regions, and time periods.

6.2. Marginal cost data

The UK Monopolies and Mergers Commission performed a detailed study of brewing and wholesaling costs by brand. In addition, they assessed retailing costs in managed public houses.²⁴ A summary of the results of that study is published in MMC (1989). Although costs were assessed on a brand basis, only aggregate costs by product type are publicly available.

Brewing and wholesaling costs include material, delivery, excise, and advertising and marketing expenses per unit sold. Retailing costs include labor and wastage, and combined costs include VAT.²⁵ If average-variable costs in brewing are constant, these are marginal costs. If not, there will be a bias in the marginal-cost measures.²⁶ Unfortunately, we have no quantitative information that can be used to assess this issue and therefore use the MMC unit-cost figures.

²⁴ Managed public houses are owned and operated by the brewer.

²⁵ See Slade (2002) for a more detailed breakdown of brewing and retailing costs.

²⁶ If a bias exists, we cannot determine the direction. For example, even under increasing returns, average-variable costs can rise.

We updated the MMC cost figures to reflect inflation. To do this, we collected a price index for each category of expense. We then multiplied each cost category by the ratio of the appropriate price index in 1995 to the same index in 1985.

6.3. The metrics

We experimented with a number of notions of distance or its inverse, closeness: beers that are of the same product type, are brewed by the same brewer, have similar coverages, and have similar alcohol contents. Furthermore, we considered beers that are nearest neighbors or share a market boundary in alcohol/coverage space.

It is reasonable to assume that many customers who drink stout, for example, and do not find their favorite brand are apt to choose another brand of stout as a substitute. Our first measure of closeness, WPROD, therefore has i, j element equal to one if beers i and j are the same type of product and zero otherwise.

Normally, one would not expect brewer identity to play a large role in determining substitution patterns. The UK system of tied houses, however, that involves brewer exclusivity agreements, could cause beers that are brewed by the same firm to substitute for one another. Our second measure of closeness, WBREW, therefore has i, j element equal to one if beers i and j are brewed by the same firm.

The above measures are discrete. We also consider two continuous measures. The first of these captures closeness in coverage space. Specifically, $WCOV_{ij} = 1/[1 + |\log(COV_i) - \log(COV_j)|]$.²⁷ We use this measure to test if, for example, popular national brands are substitutes for other popular national brands, whereas specialty brands are substitutes for other specialty brands. Our second continuous measure captures closeness in alcohol-content space, $WALC_{ij} = 1/(1 + 2|ALC_i - ALC_j|)$. We use this measure to test if, for example, light beers substitute for other light beers.

Our two continuous characteristics are also used to construct two-dimensional market areas. These can be defined either exogenously, as a function of Euclidean distance, or endogenously, as a function of ‘delivered prices’. There are four market configurations for each measure, one for each region and time period. To construct these configurations, we averaged over multiple and independent establishments in each market using volume weights.

First, consider the nearest-neighbor measures. The elements of the first matrix, WNNX where X stands for exogenous, are dummy variables that equal one if i is j 's nearest neighbor and vice versa, $\frac{1}{2}$ if i is j 's or j is i 's nearest neighbor but not both, and zero otherwise. In performing this calculation, i 's nearest neighbor is the beer that is the shortest Euclidean distance from i in alcohol/coverage space.

The second nearest-neighbor matrix, WNNN, is determined endogenously, and the letter N is used to indicate this fact. With this measure, brand i 's nearest neighbor has the lowest utility loss or ‘delivered price’ at i 's location. To find losses, we used a quadratic utility-loss function. Specifically, a consumer located at a point $x = (x_a, x_c)^T$ who purchases a brand located at a point $y = (y_a, y_c)^T$, with subscripts a and c denoting

²⁷ The functional form of this measure was chosen somewhat arbitrarily. However, with the nonparametric estimations, functional form is irrelevant.

positions on the alcohol and coverage axes, receives a utility loss equal to $\text{PRICE}_y + b_a(x_a - y_a)^2 + b_c(x_c - y_c)^2$.

The utility-loss coefficients, b_a and b_c , were found by maximizing the fit between observed market shares and those predicted using our utility-loss function, where brand i 's market area is the set of customers for whom the utility loss associated with i 's product is less than or equal to the loss associated with any other brand. We assumed that consumers are uniformly distributed in alcohol/coverage space, and normalized so that both variables range between 0 and 1. Normalization implies that a brand's market area and its market share are equal. Finally, we performed a grid search to find the coefficients, \hat{b}_a and \hat{b}_c that yield the best fit.

There are also two common-boundary measures, which we denote WCBX and WCBN. The elements of the exogenous common-boundary matrix, WCBX, are dummy variables that equal one if i and j share an exogenous-market boundary but are not nearest neighbors, and zero otherwise, where i 's exogenous market consists of the set of consumers who are at least as close in Euclidean distance to i as to any other brand. The boundary between markets i and j thus consists of customers who are equidistant from the two. The endogenous common-boundary measure, WCBN, is similar except that with this measure, boundaries are determined by relationships of equality of utility losses.²⁸ Fig. 1 depicts endogenous market areas for London in the first time period.

We also constructed variables NCBX (NCBN) to equal the number of exogenous (endogenous) common-boundary neighbors for each brand. On average, brands have 7 common-boundary neighbors.

Matrices corresponding to each discrete metric were normalized so that the elements of each row sum to one. This normalization was performed so that when the price vector is multiplied by a matrix, the i th element of the resulting vector is the average price of, for example, rival beers that are of the same type of product as i .

7. The econometric estimates

7.1. Parametric estimates

Table 2 summarizes the IV estimates of demand. Table 2A is divided into three sections: the first is the intercept term, a_i , whereas the second is the own-price term, b_{ii} . Both of these are functions of the characteristic vector, X_i . The characteristics in b_{ii} , however, have been interacted with price. The third section is the rival-price term b_{ij} , which is a function of the distance measures, d_{ij} . Variables in this section are weighted averages of rival prices and are denoted RPW, where W is one of the distance matrices.

To conserve space, the estimated coefficients of the characteristic variables are not shown. Instead, the first row in Table 2A indicates their signs, whereas subsequent rows indicate their significance. Furthermore, each equation in this table contains at most one distance measure. Finally, since brewer and product-type fixed effects were never significant, the equations that are shown do not include fixed effects.

²⁸ This measure of closeness is very similar to the one used by Feenstra and Levinsohn (1995).

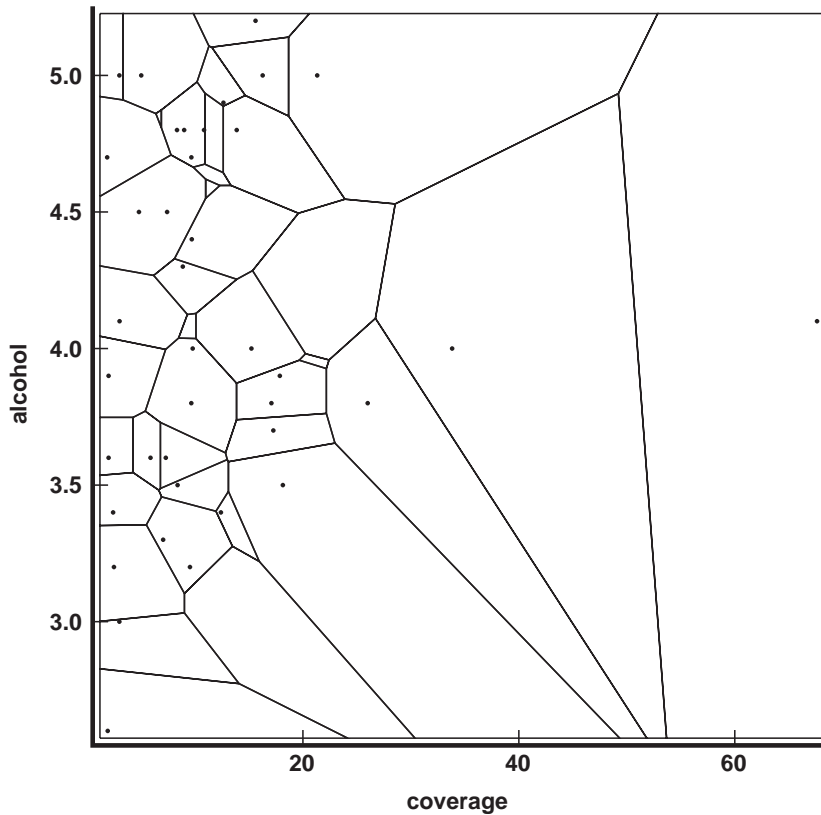


Fig. 1. Endogenous market areas for London.

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. We have tried to allocate the variables in what we think is a sensible fashion. Nevertheless, the allocation is somewhat arbitrary.²⁹ In addition, since coverage was found to be an important determinant of a brand's market size and own-price elasticity, we have included coverage in both parts of the table. Different functional forms were used to avoid collinearity, with $LCOV = \log(COV)$ and $COVR = 1/COV$.

First consider the intercepts, a_i . In all specifications, high coverage is associated with high sales. In addition, sales are higher in independent establishments and in London. Finally, a high alcohol content has a positive but weak effect on sales. None of these results is surprising.

²⁹ We experimented with other specifications and found that our principal conclusions were not affected.

Table 2
IV demand equations

A: Equations with a single distance measure

Intercept (a_i)					Own price (b_{ii})					Rival price (b_{ij})			
CONS	LCOV	ALC	PUBM	PER1	REG1	PRICE	PRCOVR	PRPREM	PRNCBX	VAR.	COEF.	t	STAT
—	+	+	—	+	+	—	+	—	—				
***	***	*	***		***	*	***	**	*	None			
***	***	*	**		***	***	***	*	*	RPPROD	0.82	2.6	
	***	*	***		***		***	**	*	RPBREW	−0.40	−1.1	
***	***		***		***	*	***		*	RPALC	0.16	0.7	
			***		***					RPCOV	−1.38	−1.7	
***	***		***		***		***	**	*	RPCBN	0.04	0.7	
***	***	**	***		***	**	***	**	*	RPCBX	−0.08	−1.7	
***	***		***		***		***	**	*	RPNNN	0.06	1.2	
***	***	*	***		***	*	***	**	*	RPNNX	0.008	0.2	
<i>Semiparametric estimates with RPPROD</i>													
***	***		**		***		***		**		3 expansion terms		
***	***		*		***		***		***		5 expansion terms		

B: IV Equation with multiple distance measures

RPPROD	RPBREW	RPALC	RPCBN	RPNNN
0.82	−0.19	0.16	0.02	0.05
(2.5)	(−0.5)	(1.2)	(0.3)	(0.9)

C: IV Equation used in evaluation of games

Intercept (a_i)					Own price (b_{ii})					Rival price (b_{ij})	
CONS	LCOV	ALC	PUBM	PER1	REGL	PRICE	PRCOVR	PRPREM	PRNCBX	RPPROD	RPALC
−159.1	60.3	8.80	−11.0	3.81	31.5	−1.13	0.17	−0.03	−0.12	0.71	0.22
(−4.0)	(11.7)	(0.7)	(−1.9)	(0.8)	(6.4)	(−2.9)	(7.8)	(−0.1)	(−2.7)	(2.6)	(1.6)

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

*Denotes significance at 10%.

**Denotes significance at 5%.

***Denotes significance at 1%.

Next consider the own-price effects, b_{ii} . Premium and popular brands have steeper (i.e., more negative) slopes (recall that COVR is an inverse measure of coverage). In addition, when a brand has a large number of neighbors, its sales are more price sensitive.

More important and the focus of the paper are the determinants of brand substitutability. The table shows that the most significant measure of rivalry is RPPROD,

which implies that competition is strongest among brands that are of the same product type. Indeed, none of the other measures of substitution is significantly different from 0 at the 5% level.

Table 2B shows a specification with multiple distance measures. As before, the discrete measure same-product-type has the highest explanatory power. In addition beers with similar alcohol contents tend to compete, but this effect is weaker.

Table 2C shows the final parametric specification. Only the same-product-type and similar-alcohol-content measures, RPPROD and RPALC, are included in this specification. This 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 compete, regardless of type.

7.2. *Semiparametric estimates*

We have carried out experiments in which g is an unspecified function of the distance between brands in alcohol and/or coverage space and several discrete brand-similarity measures. The results are not qualitatively different from the fully parametric case, and hence we present only one specification. This one is identical to the second equation in Table 2A, except that the same-product-type distance measure, WPROD, is interacted with terms of a Fourier-series expansion of difference in alcohol contents, WALC. We report specifications with three and five expansion terms, which can be found in the bottom portion of Table 2A. The table shows that the semiparametric estimates are identical in sign and similar in significance to the parametric ones. However, since the number of regressors is greater in a semiparametric specification, standard errors tend to be larger.

The estimate of the function g for the specification with five expansion terms is shown in Fig. 2. The graph shows how competition between brands that are of the same product type varies as the difference in their alcohol contents increases. In addition we show 5% and 2.5% asymptotic one-sided pointwise (Bonferroni) confidence bands. Those bands have been corrected for spatial correlation, which makes them wider than without such a correction. At a 5% level of significance, we conclude that the cross-price elasticities are nonzero. However, we cannot reject the hypothesis that g is constant.

Given that g is an approximately linear function of the distance measures, which means that our IV estimates are consistent, we use the equation that appears in Table 2C for validity checks and merger evaluation.

8. Model assessment and use

We assess our model of demand, cost, and market equilibrium in three ways: we examine the implied own and cross-price elasticities and compare them to previous elasticity estimates for beer, we test if our equilibrium assumption is consistent with observed margins, and we compare observed and predicted prices under the ownership structure that prevailed when the data were collected.

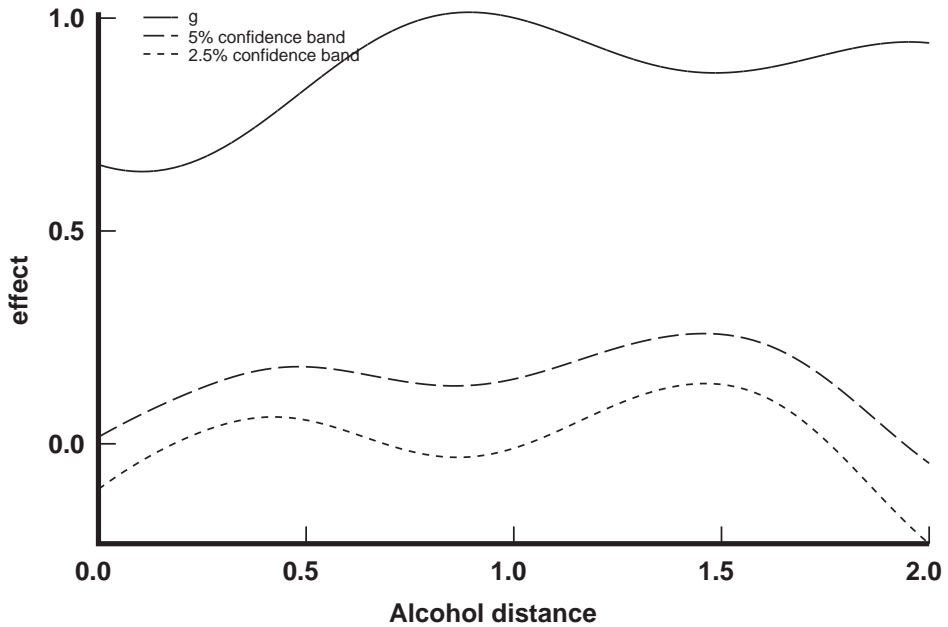


Fig. 2. g , Same product type rivals (5 expansion terms).

First, however, we assessed identification and checked regularity conditions. With respect to identification, we used our test of correlation between the residuals and various groups of instruments. This process revealed no evidence of endogeneity. For example, when we examined price in the other region by itself, the p -value for the test was 0.20, and when we examined the price instruments as a group, the p -value was 0.38. We also estimated the demand equation with each set of instruments separately, as well as with both sets together, and found that our estimated elasticities were not very sensitive to that choice.³⁰

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

8.1. Own and cross-price elasticities

There are many previous estimates of the industry own-price elasticity of demand for beer, which is the percentage change in total beer consumption due to a 1% increase

³⁰ The equations that we report were estimated with both sets of instruments.

Table 3

Own and cross-price elasticities for selected brands (London) evaluated at observed prices and quantities

Brand	Tennants Pilsner	Stella Artois	Lowenbrau	Toby Bitter	Websters Yorks Bitter	Courage Best	Greene King IPA	Guinness Stout
Alcohol content	3.2%	5.2%	5.0%	3.3%	3.5%	4.0%	3.6%	4.1%
product type	Reg. Lager	Prem. Lager	Prem. Lager	Keg Ale	Keg Ale	Real Ale	Real Ale	Stout
No. of neighbors	12	8	8	12	8	15	9	2
Tennants 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
Guinness	0.002	0.002	0.002	0.001	0.002	0.004	0.002	-0.93

in the prices of all brands. Our estimated industry own-price elasticity is -0.5 , which is in line with previous estimates.³¹

Estimated own-price elasticities for individual brands, which are calculated holding the prices of rival brands constant, are much rarer. Our average brand own-price elasticity is -4.6 , which is in line with estimates obtained by Hausman et al. (1994). Their brand own-price elasticities average -5.0 .

One can 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. We find that this elasticity averages 4.1. Partial cross-price elasticities, which are percentage changes in one brand’s sales due to a 1% increase in the price of a single rival, vary by brand pair. Moreover, as there are many brands, partial cross-price elasticities must be small. Indeed, since our cross-price elasticities are nonnegative (i.e., the brands are substitutes), stability requires that, on average, their sum be less than the absolute value of the own-price elasticities.

It is not practical to examine 63 own and approximately 4,000 cross-price elasticities. Table 3 therefore contains elasticities for a selected subsample of brands sold in London. 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, Guinness, is an outlier with a coverage that is substantially higher than that of any other brand. In addition to identifying the type of each brand, the first row of the table shows the brand’s alcohol content and the number of its exogenous common-boundary neighbors.

³¹ For example, Clements and Johnson (1983), Johnson et al. (1992), Lee and Tremblay (1992), Hogarty and Elzinga (1972) and Hausman et al. (1994) estimate industry own-price elasticities for beer to be -0.1 , -0.3 , -0.6 , -0.9 , and -1.6 , respectively.

Table 4
Nash-equilibrium prices (London)

	Actual	Status quo	Before	%Δ	After	%Δ
Mean	167.8	168.4	167.4	−0.6	173.5	+3.0
Standard deviation	20.2	29.5	22.1		30.2	

Table 3 shows that there is substantial variation in brand own-price elasticities. Furthermore, most of the magnitudes are plausible. In particular, if one ranks the reciprocals of the absolute values of the own-price elasticities and ranks the price/cost margins, the rankings are very similar. The table also shows that, as expected, cross-price elasticities are greater when brands are of the same type and have similar alcohol contents.

All brand own-price elasticities are significant at 1%. Cross-price elasticities for brands of the same type are also significant at 1%. When brands are of different types, however, their cross-price elasticities are not significant at 5% but are at 10%.

8.2. Comparing status quo margins and prices

We can use our estimates of demand and cost to test our equilibrium assumption. In particular, we wish to determine if a static game in prices is a reasonable description of interactions in the market. To do this, we construct excess margins, $\hat{\theta}_{\kappa i}$, for the *status quo* game κ and test if they are on average zero, where the *status quo* ownership configuration is the situation that prevailed when the data were collected. In other words, with the *status quo*, there are ten brewers, four nationals, Bass, Carlsberg–Tetley, Scottish Courage, and Whitbread, two brewers without tied estate, and four regionals.

The mean excess margin is -0.009 , the median is -0.024 , and the range is -0.30 to 0.56 . Furthermore, the t statistic for the hypothesis that $1/n \sum_i \hat{\theta}_{\kappa i} = 0$ is -0.66 , which means that, on average, predicted and observed margins are equal and Bertrand behavior cannot be rejected.³²

On average, excess margins are zero. Nevertheless, there are systematic deviations from zero in the estimates. In particular, more popular and higher-strength brands, as well as those that are sold in multiple establishments, are less competitively priced.

Next consider prices. The first columns of Table 4 show observed and Nash-equilibrium (NE) prices calculated under the *status quo* assumption for London. The average predicted price is 168.4, slightly higher than the observed average of 167.8, and the standard deviation is 29.5, compared to 20.2. *Status quo* prices are thus approximately centered around the true value but are more variable. However, as with

³² The test is based on the fact that any sequence of i.i.d. variates with uniformly bounded moments greater than two, whether they are estimators or not, have a limiting normal distribution; this follows from the Lindeberg theorem, e.g. Doob (1953, Theorem 4.2). The notion that the estimators $\hat{\theta}_{\kappa i}$ are independent, even in the limit, is debatable. If they are dependent however, the t -statistics would generally be smaller.

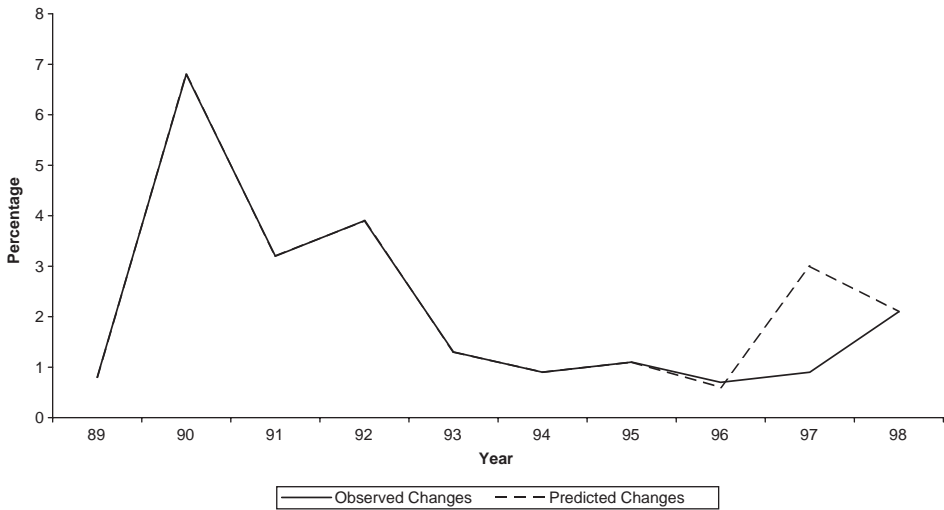


Fig. 3. Observed and predicted real retail price changes.

excess margins, there are systematic deviations between observed and predicted prices. Nevertheless, we feel that, on average, the model predicts well.

8.3. The price effects of the mergers

Given that our estimated model of demand is well behaved and that our equilibrium assumption is not rejected, we can use our model for a formal analysis of the impacts of mergers and divestitures. This analysis involves solution of market games under different assumptions about ownership.

Static NE prices are calculated under three scenarios: *before*, *status quo*, and *after*. The *before* scenario is created by undoing the Scottish Courage merger. With this scenario, there are five national brewers, and the prices of the brands that belonged to Courage before the merger are chosen by one player, whereas the prices of the brands that belonged to S&N are chosen by another. The *after* scenario is created by allowing the Bass/Carlsberg–Tetley merger to occur. With this scenario, there are three national brewers, and the prices of all Bass and Carlsberg–Tetley brands are chosen by a single player. With all three scenarios, the structure of the rest of the industry is unchanged.

Equilibrium prices under the various scenarios are summarized in Table 4. For consistency, *before* and *after* prices are compared to *status quo*, rather than to observed prices. The table shows that undoing the Scottish Courage merger would have little effect on prices. Indeed, the price decline is less than one percent. A merger between Bass and Carlsberg–Tetley, in contrast, is predicted to result in an overall price increase of 3%. To us, this number seems substantial.

Fig. 3 shows observed and predicted price changes. One can see that the observed and predicted changes in the year of the Scottish Courage merger are close to one

another. In addition, the change that is predicted to accompany the BCT merger is substantially larger than the change that occurred absent the merger.³³

Why might a merger between S&N and Courage not cause substantial price increases? The answer lies in both the geographic and the brand fit. Geographically, S&N had most of its market in the north, whereas Courage was a southern beer. In London, therefore, Courage dominated S&N, and the local difference in their market shares was greater than the national difference. Moreover, Courage had a strong position in lagers, with best selling Fosters and Kronenbourg. S&N, in contrast, had only Becks, which was not a heavy seller. Comparing ales, both Courage Best and John Smiths were big brands for Courage, whereas S&N's biggest seller, Theakstons, had a much smaller share of the London market.

The brand overlap between Bass and Carlsberg–Tetley, in contrast, was more substantial, particularly in lagers. Indeed, both had best-selling lagers, Carling for Bass and Carlsberg for Carlsberg–Tetley, as well as several other very popular brands. Geographically, the firms originated in the north (Tetley) or north central (Bass) part of the country. Their geographic strengths were thus more similar than those of Courage and S&N, which meant that their local and national market shares were closer.

The geographic and brand overlap (or lack thereof) is undoubtedly responsible for the different impacts that the mergers are predicted to have on prices. Indeed, the difference in market shares, 37% for Scottish Courage versus 34% for BCT in the geographic and brand market studied, is small. Moreover, if one were to consider local-market shares alone, one would expect the Scottish Courage merger to be more anticompetitive, which is the opposite of what we find.

9. Other consequences of the mergers

Mergers can affect all aspects of a firm's business, not just prices. We therefore examine the possible ramifications for brand selection and productive efficiency informally.

Although, in general, there is considerable brand churn, very little of it concerns products that constitute as much as $\frac{1}{2}$ % of local markets. Moreover, it is clear from interviews with industry managers and marketers that firms are principally interested in the fate of their best-selling brands. None of these disappeared during the ten-year period. Furthermore, according to people in the industry, the brands that did disappear would have done so absent the merger, albeit at a possibly different rate.

In contrast to brand disappearance, new best-selling brands were introduced. These fall into two classes: foreign-owned lagers and hybrid ales. The lagers that are in the sample but were not sold in the UK in 1985 are Grolsch, a Dutch beer, and Coors, an American beer. It is highly unlikely, however, that absent the mergers (or in the case of BCT, had the merger occurred) those brands would not have entered the market.

³³ The large price increases that occurred in the beginning of the decade are perhaps due to forced sales of retail outlets (see Slade, 1998).

Hybrid ales are a relatively recent addition to the UK market. A hybrid is a keg ale that uses a nitrogen and carbon-dioxide mix in dispensing that causes it to be smoother and to more closely resemble a cask ale. After Bass introduced Caffreys, which is the most popular hybrid ale, other brewers followed suit. As with the lagers, it is highly unlikely that the introduction of hybrid ales would have been retarded or aborted by mergers or divestitures.

The two mergers under consideration are thus unlikely to have had much of an impact on product selection. However, they could have caused promotional efforts to be concentrated on a smaller number of best-selling brands. If true, there might have been less variety in consumption if not in offerings.

When a merger occurs it is usually accompanied by restructuring and cost reduction. Indeed, increased efficiency is one reason why firms undertake mergers. It is therefore possible that cost reductions could have offset market-power increases, and the mergers could have caused prices to fall.

The BCT merger, like other mergers in the brewing industry, was forecast to result in cost savings of two types: brewery closings and savings in distribution and wholesaling costs. Retail costs were not supposed to be affected. In the absence of the merger, however, the breweries that were scheduled to close have in fact closed. In contrast, there has been little reduction in distribution and sales forces. If there are merger-specific offsetting cost savings, we must therefore look for them in a reduction in wholesaling costs, and, according to the MMC's cost estimates, wholesaling costs constitute only 10% of the retail price of beer.

It is possible to assess whether cost reductions could have offset increased market power formally. To do this, we solved for the cost reduction that would be required to insure that, in equilibrium and on average, no price increase would occur. In performing this exercise, we recognized that only the costs of the merging firms would be reduced.³⁴ We found that the costs of the merging firms would have to fall by about 20% to just offset the increase in market power. Given that wholesaling costs are only 10% of price, a reduction of the required magnitude would not have been possible.

Finally, entry is unlikely to mitigate the effects of mergers. The number of brewers has been in decline for decades, and no firms have entered the national segment of the industry. Indeed, there were six national brewers in 1960, each of which is roughly identifiable with one of the six that existed in 1990, just prior to the Courage/Grand Met merger.

10. Conclusions

Our analysis of brand substitution in UK markets for draft beer indicates that competition is relatively local. In particular, we find that brands that are the of the same product type compete most vigorously. Our model of demand is thus similar to a nested logit. Competition among groups of products, however, is not a maintained assumption,

³⁴ Specifically, we substituted $\tilde{c}_i = (1 - \alpha) \times c_i$ if i was a BCT brand, and $\tilde{c}_i = c_i$ if not, into the *after* first-order conditions, and we solved for the value of α that would leave prices unchanged on average.

since this possibility is encompassed in a broader model of substitution. We also find that beers with similar alcohol contents compete regardless of type, but the strength of this competition is considerably weaker.

When we use our estimated demands and costs to analyze the game that the brewers are playing, we find that a static Nash equilibrium in prices cannot be rejected. We therefore use the Bertrand assumption to predict the effects of mergers and divestitures, both actual and counterfactual. It is of course possible that changes in ownership structure could cause changes in collusiveness. If higher concentration is associated with more collusive outcomes, however, our estimates of price changes due to mergers and divestitures are conservative.

Our analysis of the mergers indicates that, whereas the (consummated) merger between Courage and Scottish & Newcastle had little effect on prices, the proposed merger between Bass and Carlsberg–Tetley would have raised prices by a more substantial amount. This conclusion relies heavily on our findings about the structure of demand. Indeed, the local market shares of the post-merger firms would have been similar, so that, if competition were symmetric, the effects of the mergers would also be similar. With localized competition, in contrast, the identity and product mix of each merging partner is key in determining whether the merger will be anticompetitive.

Acknowledgements

This research was supported by grants from the Social Sciences and Humanities Research Council of Canada. We would like to thank the following people for thoughtful comments: Aviv Nevo, Otto Toivanen, two anonymous referees, and seminar participants at the ES meetings in Boston, the Canadian Bureau of Competition Policy, the conferences on Industrial Organization and the Food Processing Industry in Toulouse and Antitrust Issues in International Markets in Berlin, and the Universities of Calgary, Southern California, Carnegie Mellon, Vanderbilt, Victoria, Carlos III, Oxford, Warwick, and Paris I.

References

- Berndt, E.R., Fuss, M.A., Waverman, L., 1977. Dynamic models of the industrial demand for energy. Electric Power Institute, EA-580, Palo Alto.
- Berry, S., 1994. Estimating discrete-choice models of product differentiation. *RAND Journal of Economics* 25, 242–262.
- Berry, S., Levinsohn, J., Pakes, A., 1995. Automobile prices in market equilibrium. *Econometrica* 63, 841–890.
- Blackorby, C., Primont, D., Russell, R., 1978. Duality, Separability, and Functional Structure: Theory and Economic Applications. North-Holland, Amsterdam.
- Bresnahan, T.F., 1989. Empirical studies of industries with market power. In: Schmalensee, R., Willig, R. (Eds.), *The Handbook of Industrial Organization*, Vol. 1. Elsevier, Amsterdam.
- Clements, K.W., Johnson, L.W., 1983. The demand for beer, wine, and spirits: A systemwide analysis. *Journal of Business* 56, 273–304.
- Corts, K.S., 1999. Conduct parameters and the measurement of market power. *Journal of Econometrics* 88, 227–250.

- Deaton, A., Muellbauer, J., 1980. An almost ideal demand system. *American Economic Review* 70, 312–326.
- 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., Levinsohn, J.A., 1995. Estimating markups and market conduct with multidimensional product attributes. *Review of Economic Studies* 62, 19–52.
- Gorman, W.M., 1971. Two-stage budgeting. Mimeo.
- Hausman, J., Leonard, G., Zona, D., 1994. Competitive analysis with differentiated products. *Annales d'Econometrie et de Statistique* 34, 159–180.
- Hogarty, T.F., Elzinga, K.G., 1972. The demand for beer. *Review of Economics and Statistics* 54, 196–198.
- Ivaldi, M., Verboven, F., 2000. Quantifying the effects from horizontal mergers: The European heavy trucks market. Mimeo.
- Johnson, J.A., Oksanen, E.H., Veall, M.R., Fretz, D., 1992. Short-run and long-run elasticities for Canadian consumption of alcoholic beverages: An error-correction mechanism/cointegration approach. *Review of Economics and Statistics* 74, 64–74.
- Lee, B., Tremblay, V.J., 1992. Advertising and the market demand for beer. *Applied Economics* 24, 69–76.
- McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), *Frontiers in Econometrics*. Academic Press, New York, pp. 105–142.
- McFadden, D., 1978a. Modeling the choice of residential location. In: Karlqvist, A., et al. (Ed.), *Spatial Interaction Theory and Planning Models*. North-Holland, Amsterdam.
- McFadden, D., 1978b. The general linear profit function. In: Fuss, M., McFadden, D. (Eds.), *Production Economics: A Dual Approach to Theory and Applications*, Vol. 1. North-Holland, Amsterdam, pp. 269–286.
- McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15, 447–470.
- Monopolies and Mergers Commission (MMC), 1989. *The Supply of Beer*. Her Majesty's Stationary Office, London.
- 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., West, K.D., 1987. A simple positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 55, 703–708.
- Petrin, E., 2002. Quantifying the benefits of new products: The case of the minivan. *Journal of Political Economy* 110, 705–729.
- Pinkse, J., Slade, M.E., Brett, C., 2002. Spatial price competition: A semiparametric approach. *Econometrica* 70, 1111–1153.
- Slade, M.E., 1998. Beer and the tie: Did divestiture of Brewer-owned public houses lead to higher beer prices? *Economic Journal* 108, 565–602.
- Slade, M.E., 2002. Market power and joint dominance in the UK brewing industry. Mimeo., University of Warwick.
- Werden, G.J., 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.