Diagnosing Foreclosure due to Exclusive Dealing

John Asker[1]

October 14, 2005

Leonard N. Stern School of Business, NYU

Abstract

Exclusive dealing arrangements, in which a distributor contracts to work exclusively with a single manufacturer, can be efficiency enhancing or they can be an anticompetitive means to foreclose markets. This paper evaluates the effect of exclusive distribution arrangements on competition in the Chicago beer market in 1994. A diagnostic test is provided to judge whether exclusive arrangements between brewers and their distributors lead to foreclosure. To implement this test I estimate a model of consumer demand and firm behavior that incorporates industry details and allows for distribution through exclusive and shared channels. The test indicates that foreclosure effects are not present in this market, suggesting that the most likely effect of intervention would be to reduce social welfare.

JEL Nos: L12, L22, L42, K21

Keywords: Exclusive dealing, foreclosure, vertical restraints

1 Introduction

This paper considers exclusive dealing agreements between manufacturers and distributors. These agreements stop the distributor selling the products of the manufacturer's rivals. Such agreements raise antitrust concerns, since they may restrict the access of upstream rivals to distributors. Rivals may be foreclosed from the market altogether or, more commonly, forced to use higher cost, or less effective, methods to bring their products to market. In either case, competition can be reduced through either reducing the number of manufacturers serving the market or by artificially raising the costs of some manufacturers.

In contrast to these anti-competitive effects, it is frequently argued that exclusive dealing agreements help a firm organize their distribution more efficiently. In such cases, where these agreements result in cost reduction or some other efficiency dividend, antitrust concerns may be unfounded.

This paper formulates and implements a test for foreclosure using non-proprietary data on prices and shares. Foreclosure is considered in both the sense of forcing competitors to use higher cost distribution channels, and also in the sense of forcing competitors to use distributors who are less effective at promotion. The results suggest that, at least in the beer market, concerns about foreclosure are unfounded.

The American beer industry is dominated by a few large brewers. During 1994, the period of analysis, Anheuser Busch was the largest brewer nationally, followed by Miller, Coors, Stroh, Heileman and Pabst. Anheuser Busch produced 45% of the beer shipped in the USA, Miller produced 22%, Coors, 10%, Stroh 6%, Heileman 5% and Pabst 4%. Imported beers, treated together, accounted for about 5% of the market.[2]

In 1997, the US Department of Justice began an investigation into distribution practices in the industry.[3] During this time Anheuser Busch was engaged in a campaign to increase their number of exclusive distributors. This campaign, dubbed "100% Share of Mind" by the company, offered distributors extended credit, increased marketing support and other incentives for going exclusive.[4] This drive for more exclusive distributors raised antitrust concerns about exclusive dealing. Paraphrasing, Section 3 of the Clayton Act provides that an exclusive contract for the provision of goods that substantially lessens competition between upstream firms is in breach of the law.

---

[2] Perhaps unsurprisingly, the beer industry has had a long history of involvement with antitrust authorities. See Tremblay (1993) and Klein and Murphy (1988) for a discussion.

[3] This investigation was reported widely in the business press. See, for example, Canedy (1997) and Chicago Sun-Times (1997)

[4] See Lenderman (1997), Butler (1997) and Stokes (1997) for more details. Sherwood (2000) provides an interesting account of the industry at this time. The Anheuser Busch Annual Reports for 1998 and 1999 reveal that in 1997 40% of A-B beer was sold through exclusive distributors; by 1999 this had risen to 60%.

This paper asks whether the use of exclusive distributors lessens competition by foreclosing distribution channels that might otherwise be available to competing brewers.

The new data set used in this study is a detailed description of the Greater Chicago distribution networks of 12 beer brewers. This is merged with data on the sales of beer from a major Chicago supermarket chain. These demand data include information about the price paid by the chain and the price paid by consumers for the units of beer sold. Combining these sources results in a rich data set that allows the sale of beer to be tracked right up through the vertical chain.

The analysis proceeds in several steps. First, demand estimates are obtained. Then an industry model is used with the demand estimates to compute markups. These markups are combined with the price to retail to back out the costs to the brewer and distributor. These estimated costs are then compared to see whether exclusivity agreements lead to cost-based foreclosure. The demand estimates are used to examine the incidence of promotion-based foreclosure.

The bulk of the existing literature on exclusive dealing is theoretical. Theoretical work on exclusive dealing has investigated justifications for, and the consequences of, exclusionary contracts. The literature is best organised in terms of the offered justifications, which can be grouped into foreclosure, spillover and incentive stories. Salop and Scheffman (1983), Comanor and Frech (1985), Schwartz (1987), Mathewson and Winter (1987), Rasmusen et al (1991) and Bernheim and Whinston (1998) all provide models that describe the incentives for a manufacturer to engage in exclusive dealing in order to foreclose rivals from the market. The essence of all these papers is that a firm can gain from an exclusive dealing contract by forcing competitors to use higher cost distributors or by completely foreclosing them from the market. Besanko and Perry (1993) and Segal and Whinston (2000) provide models where investment spillovers can encourage firms to engage in exclusive dealing. They show that exclusive dealing arrangements can provide an incentive for firms to invest more in their distributors. Marvel (1982) and Ornstein (1989) make similar, informal, arguments. Bernheim and Whinston also provide a model that shows that exclusive dealing can solve incentive problems between manufacturers and shared distributors. This gives a reason why we might expect to see exclusive distributors providing better service. Prat and Rustichini (2003), Klein and Murphy (1988) and Martimort (1996) provide related models.[5]

A number of empirical papers exist which are related to this study. Sass (2002) uses a detailed survey of the cost structure of 391 beer distributors to present evidence that exclusive dealing increases the productivity of the sales and promotional effort of distributors. Sass argues that these productivity increases are most likely due to incentives put in place by brewers. Chipty (2001) examines foreclosure in the context of integration between programming and distribution in the cable television industry. She finds evidence in support of foreclosure, in the sense that integrated operators

---

[5]In the Klein and Murphy model incentives are provided by the threat of contractual termination rather than via incentive contracts.

have less externally produced programming. The study here differs from Chipty's in that it focuses on promotional and cost side impacts of any possible foreclosure effect.

Using a related approach to this paper, Brenkers and Verboven (2002) evaluate the welfare impacts of enhanced competition between car dealers due to the removal of exclusive territories and exclusive distribution agreements in the European car market. They proceed by estimating demand and then running experiments based on different scenarios describing the price setting behaviour of dealers and manufacturers. These experiments benefit from sanity checks afforded by limited data on average dealer markups at the country level. Due to much finer data on distribution arrangements, the study in this paper is able to delve deeper into the specific effects of particular vertical arrangements.

Villas-Boas (2005a,b) and Bonnet et al (2004) use inferred costs (derived using similar methods to those used here) to test different hypotheses about the interaction of supermarkets and manufacturers. Whereas in their markets the method of interaction is unknown, in this study information from interviewing beer distributors and the regulator is used to model this interaction. In this sense, this paper is different in that the mode of conduct in the vertical market is observed. A similar approach, using some of the same data as in this paper, is also used by Hellerstein (2006) to invesigate exchange rate pass through for beer.

A feature of the beer industry in many states, including Illinois, is the regulatory requirement that exclusive sales territories be assigned to distributors. This is used in this paper to track products up the vertical chain and determine the nature of interaction between distributors. Sass and Saurman (1996), Culbertson and Bradford (1991) and Jordan and Jaffe (1987), all examine the effects of exclusive territories on beer prices. They find that the use of territories have a positive effect on price, while Sass and Saurman also conclude that the territories enhance the provision of dealer services.

The rest of this paper is organized as follows: Section 2 describes the structure of the Chicago beer market. Section 3 provides an overview of the model used to provide structure to the data, then describes the empirical implementation of this model. Section 4 describes the data set. Section 5 discusses the empirical results, including the demand estimates and the testing of both the promotional and cost-side promotional hypotheses. Finally, Section 6 offers some concluding remarks.

2 The Chicago Beer Market

Two Illinois state statutes provide the basis for the regulatory structure which governs the sale of beer in Illinois. [6] This legislation provides the basis for a three-tiered vertical market structure, with the tiers being brewer, distributor and retailer. A member of one tier is prohibited from operating in another tier. Under the legislation

---

[6] The Liquor Control Act of 1934 (235 ILCS 5/.) and The Beer Industry Fair Trading Act (815 ILCS 720/.)

the brewer has to supply their beer to one or more distributors who are in turn responsible for distributing the beer to retailers.

A distribtor must be given a distribution territory by the brewer. Within this territory the distrbutor has the exclusive distribution rights for the brands asigned to them. These territories must be registered with the Liquor Control Commission. The standard practice is that territories be contiguous geographic areas commonly bordered by major roads or county boundaries, although brewers are free to configure them as they wish.

Within this three-tier structure the market works as follows. Brewers produce the product and deliver it to the distributor. The distributor buys the beer at a flat rate from the brewer and then sells it on to retailers within their territory. Figure 1 shows a stylized example of the way the three-tier system works.

Insert figure 1 here

The brewer has varying degrees of input into the wholesale price charged by the distributor to retailer. When dealing with a large supermarket chain, a sales representative from the brewer will arrive at a wholesale price for the chain with the chain's buyer. Distributors are then expected to supply at that price.[7] Discussions with distributors indicated that they had little or no freedom to negotiate another price; as far as they were concerned they had to take the price as given. The demand and price data used in this study is taken from the Dominick's Finer Foods (DFF) supermarket chain. Since this is a large chain, this data allows us to test the claim that the price to retail is fixed by the brewer. This institutional structure is confirmed in the data - less than 2% of observations at the UPC level have the price to retail varying across distributors serving Dominick's Finer Foods.[8]

Distributors also have a promotional role. While brewers take care of national and regional advertising campaigns, distributors are responsible for point of sale (POS) promotion and advertising. Distributors supply POS materials to retailers and coordinate other promotional activity at the retail outlet.[9]

The laws regulating the Illinois beer industry make it easy to identify the vertical structure of this market. Elsewhere, the vertical structure is an artifact of individual distribution contracts which are very hard to observe. Here, the registration of exclusive territories makes the vertical structure observable and the laws give a clear basis for the interpretation of the nature of competition at the distributor level. The territory records provide a way to map from a beer sale in a particular store, to

---

[7]While resale price maintenance is prohibited explicitly by the Beer Industry Fair Trading Act (sections 5(3) and 2(A)(i)), this practice does not appear to invite legal sanction.

[8]Universal Product Codes (UPCs) are the bar codes found on most products.

[9]The provision of distributor promotional service is seen as very important by all brewers; so important that Anheuser-Busch and Miller initiated legal proceedings against each other in 1998 over contractual clauses they alleged interfered with each others' products (see Stamborski 1998 and Bloomberg News 1998).

the distributor who provided the beer and then to the brewer. This feature of the regulatory environment enables much of the following analysis to proceed.

3 The Foreclosure Tests, the Industry Model and its Empirical Implementation

To investigate the potential impact of exclusive dealing arrangements on the beer industry in Chicago I propose an industry model and then use this model to estimate demand and then cost parameters. The model provides the environment in which the tests for foreclosure can be implemented. Rather than jumping straight into this model, I first describe how the foreclosure tests work. I then describe the model, which is used to take the analysis to the point where these foreclosure tests can be implemented.

Although the data are described fully in section 4, it will be useful to have in mind the following facts. The demand data contain UPC level observations for 73 beer brands drawn from 71 Dominick's Finer Foods (DFF) stores spread through Greater Chicago, in 24 weeks. Pricing by DFF is done at a chain-wide level. Moving up the vertical chain, 12 brewers use 42 distributors to serve DFF, averaging 6 distributors each. I observe the price to the consumer and the price to retail for each product.

3.1 The Foreclosure Tests

3.1.1 Concept

This paper examines foreclosure arising from competitors being unable to gain access to low cost distributors and also from competitors being unable to gain access to those distributors that provide better promotional services. In both cases we are concerned about distributors with desirable characteristics being foreclosed. In what follows this characteristic is assumed to be low cost distribution, however the logic remains the same when promotional aptitude is considered.

The test considers two types of market; those with a distributor to whom access is restricted by an exclusive arrangement with a brewer: and markets without a distributor restricted by an exclusive arrangement. If the distribution of potential distribution costs in these markets were the same then, if the foreclosure hypothesis were correct, the presence of an exclusive arrangement would mean that, for those brewers that never use exclusive distributors, this distribution of costs would be truncated from the left. That is, an exclusive arrangement would prohibit competitors from getting access to the distributors in the low cost tail of the distribution. The effect of this would be that, in expectation, distribution costs for the set of potentially foreclosed brewers should be higher in markets that have access to a distributor restricted by an exclusive arrangement.

Of course, if exclusivity arrangements are being used to protect investments then this pattern would not necessarily be expected. It is easy to imagine a setting in

which the initial state of distributors is the same and exclusive distributors, via some investments, are made more cost effective. In this setting there would be no difference between the two types of markets in terms of the distribution costs for competitors.

The test is designed to pick up whether foreclosure effects are present. This is distinct from diagnosing whether they were intended. It may be that exclusivity arrangements are being used to protect spillovers and the distributors selected to be exclusive are also from the low cost tail of the distribution. An implication of this would likely be foreclosure, in the sense of competitors having higher costs in exclusive markets. In such instances the social benefit of exclusive practices would be ambiguous since the added incentive for investment may or may not compensate for the distortion in competitors costs.

<center>**Insert figure 2 here**</center>

The central problem of formulating a test for foreclosure is in the lack of understanding of the structure on equilibrium in distribution chains with and without and exclusive dealer. Because of this the test offered here is not "necessary". This is illustrated in Figure 2. Figure 2 shows two markets in equilibrium. In both markets the brewer marked by the grey circle is matched to the low cost distributor (distrbutor 1). In market 2 this relationship is exclusive. The equilibirum allocation of brewers to distributors in market 1 results in the average cost of distribution being c^* for brewers other than the grey brewer. In market 2, the configuration of brewers and distributors changes so that all other brewers use distributor 2. In market 2, despite on brewer being forclosed, the average distribution cost has decreased from c^* to c_2. Despite forclosure occuring in this market, the change in equilibrium configurations will result in this not being detected by the proposed test. The cases that will prove problematic for the test, as made clear by the example in Figure 2, are where one brewer is excluded but equilibrium reallocations of distributors actually benefit other brewers in a substantial way. In these cases it may even be that the social welfare effect is positive overall. Thus the cases where the test may falsely reject the hypothesis of foreclosure are likely to those that are less important from the perspective of policy.

Having laid out the basis for the test I now turn to questions of implementation.

3.1.2 Implementation

Implementation of the concept underlying the test is aided by the structure of the data. In the Chicago beer market all Anheuser Busch distributors are exclusive, half the Miller distributors are exclusive and all other brewers are using shared distributors. This means that the treatment effect of exclusivity of Anheuser Busch is constant across the region, while there is variation in Miller.

In running the test each store is taken to be a distinct market. The test compares the cost to brewers, other than Anheuser Busch and Miller, of serving a market with a Miller exclusive to the cost of serving markets without a Miller exclusive. If the

<center>6</center>

cost of serving a market with a Miller exclusive is higher then this can be taken to be indicative of either foreclosure or some other market specific effect correlated with the exclusivity of Miller distributors. To untangle these effects the Anheuser Busch distributors are used as a control group to pick up market specific effects. That is, the cost effect for Anheuser Busch of being in a market with a Miller exclusive is measured and this is taken to be the market effect. This inferred market effect is then used to normalize the measured effect on the potentially foreclosed brewers.

This implementation requires two important conditions: First, that the second and higher-order moments of the distribution of potential distribution costs are invariant across markets; And second, that the access of Anheuser Busch to distribution is not affected by the presence of a Miller exclusive. This second condition is satisfied if Anheuser Busch were to get first pick of the distributors. If Anheuser Busch is affected by the presence of a Miller exclusive this would mean that the market specific effect may have a positive bias, presuming that the more likely effect would be that the presence of a Miller exclusive would make access harder. However, given Anheuser Busch always chooses to be exclusive it is hard to imagine how the exclusivity or otherwise of another distributor would affect them.[10]

Implementation of this foreclosure test requires cost information. Cost information is highly sensitive commercial data and is unavailable here. To combat this issue what I do here is use the industry model outlined below to infer cost data from price and quantity data.

When foreclosure of promotionally adept distributors is examined the test is run in a similar way. Instead of comparing costs in different markets, the demand effect of facing a Miller exclusive is the variable of interest. If facing a Miller exclusive has a negative impact on the performance of products from potentially foreclosed brewers then we have some basis for being concerned about foreclosure.

To begin with the test outlined above is implemented using the demand data. I examine the effect of facing a Miller exclusive on demand, using Anheuser Busch as a control group to manage any market specific effects.

An identification problem emerges when considering running the test in this way. There are at least three different reasons why product performance may be lower in markets with a Miller exclusive: first, there may be some market specific effect, such a strong dislike for beer in these markets; second, the Miller exclusives may be better at promotion, causing people to substitute away from other brands; third, the promotional aptitude of other brewers distributors may be lower due to foreclosure effects. The third effect is of interest here. In the basic form of the test the Anheuser Busch control group is being relied on to control for the first two effects. Since cross-product substitution effects are likely to be different for different products this is especially unsatisfactory for dealing with the second effect. With this in mind, a reformulated version of the test is also run.

---

[10]One possible way it might affect Anheuser Busch is if it changed the "order" of who picked distributors when.

Distributors who do business for the potentially foreclosed brewers are divided into three classes. The first class comprises those distributors who never serve a market with a Miller exclusive. The second class comprises those distributors who serve some markets with a Miller exclusive and some markets without. The third class comprises those distributors who only serve a markets with a Miller exclusive. In implementing the test I compare distributors in classes two and three in markets where Miller is exclusive. Distributors in class two are used in markets where Miller is not exclusive. This means that they are selected even when the market is not potentially foreclosed. Thus, under a foreclosure hypothesis, distributors in class two should be more promotionally adept than distributors in class 3. By comparing these two groups of distributors in markets with Miller exclusives the first two reasons (outlined above) why product performance may be lower in markets with a Miller exclusive are controlled and the third reason, that the promotional aptitude of other brewers distributors may be lower, can be evaluated.

One condition that is required for this implementation of the test is that class two distributors do not substitute promotional activity away from markets with a Miller exclusive toward those markets without a Miller exclusive. This would lead us to be biased toward rejecting foreclosure to often. A sense of the importance of this condition can be gleaned from the data, and will be discussed with the results.

3.2   The Industry Model

The purpose of the industry model is to provide a framework for analysis of the data, so that the tests outlined above can be implemented. When examining cost-based foreclosure the cost data needs to be inferred from pricing decisions and shares. This is due to the fact that actual cost data is unavailable due to its commercial sensitivity. The model outlined below allows this inference to be made. The investigation of promotion-based foreclosure utilizes a demand system which is part of the empirical industry model.

The proposed model divides the industry into its three tiers: retailers, distributors and brewers. Products are taken to be horizontally differentiated, with consumers having different preferences for different products. The retailer (Dominick's Finer Foods) sets the price paid by consumers, taking into account the cross-price elasticities with all the products they stock, the price they pay at wholesale and their other costs. Further, when they set a price for a product they set a chain-wide price. That is, for a given product, price across different stores does not vary.

The brewer sets the price paid by the retailer and the price paid by the distributor. In this sense, the brewers' pricing decision is modeled in the same way as resale price maintenance is modeled in standard double marginalization models. When the brewer sets prices they take into account their costs, a participation constraint for the distributor, the cross-price elasticities with other products they produce and the impact of the price paid by the retailer on the ultimate price paid by the consumer. In this model the retailer is assumed to be a monopolist while the brewer is taken to be an oligopolist.

The timing of the model is set out in Figure 3. Brewers move first, taking into account the impact of the prices they set on the prices subsequently set by the retailer. The retailer sets the price paid by consumers after the brewers have set wholesale prices. The model is a dynamic game of complete information, solved by application of the equilibrium concept of subgame perfection. As is standard in these games I use backward induction to find equilibria.

The retailer solves the following problem when choosing the retail price[11]

$$\max_{p_j^c} \sum_k \sum_j \left(p_j^c - p_j^r - c^r\right) \left[M_k s_{jk}\left(p^c, x|\theta\right)\right] \tag{1}$$

Where p_j^c (p_j^r) is the price to consumers (retailers) of product j at store k, c^r captures store level marginal costs, M_k is the market size for store k and $s_{jk}\left(p^c, x|\theta\right)$ is the share of product j at store k which is a function of consumer prices, other product characteristics and model parameters. Solving this problem gives the following first-order condition for each product l

$$\sum_k M_k \left(s_{lk}\left(p^c, x|\theta\right) + \sum_m \frac{\partial s_{mk}\left(p^c, x|\theta\right)}{\partial p_l^c}\left(p_m^c - p_m^r - c^r\right)\right) = 0 \ \forall l \tag{2}$$

The brewer solves the following price setting problem:

$$\Pi^b = \max_{p^r, p^d} \sum_{j \in B^b} \sum_k \left(p_{jd(k)}^d - c_j^b\right) M_k s_{jk} \tag{3}$$

$$\text{subject to } p_j^r - \left(c_{jd(k)}^d + r_{jd(k)}^d\right) - p_{jd(k)}^d = 0 \quad \forall d\left(k\right)$$

Where $p_{jd(k)}^d$ is the price paid by the distributor who serves store k for product j, B^b is the set of products made by the brewer and c_j^b represents the costs of of production of the brewer. The term $\left(c_{jd(k)}^d + r_{jd(k)}^d\right)$ catches the per unit cost of distribution internalized by the brewer. A component of this is the cost of the physical distribution process, $c_{jd(k)}^d$, while the rest are any quasi-rents that may have to be transferred to the distributor per unit distributed, $r_{jd(k)}^d$. This $r_{jd(k)}^d$ may be a result of some bargaining process or transfers used to induce particular actions. In any case, $\left(c_{jd(k)}^d + r_{jd(k)}^d\right)$ summarizes the markup on $p_{jd(k)}^d$ received by the distributor, which is treated as a cost term when the brewer is setting the optimal level of p_j^r.

The first-order condition for expression (3) is

$$\frac{\partial \Pi^b}{\partial p_j} = \sum_k M_k s_{jk} + \sum_{l \in B^b} \sum_k \left(p_l^r - \left[c_{jd(k)}^d + r_{jd(k)}^d\right] - c_l^d\right) M_k \sum_m \frac{\partial s_{lk}}{\partial p_m^c} \frac{\partial p_m^c}{\partial p_j^r} = 0 \ \forall j \in B^b \tag{4}$$

---

[11]The retailer pricing problem is used in two ways. First, the theoretical structure is important for computing the markup on the price to retail (and the weights on distributor cost variables). Second, it is estimated as a robustness check on the demand system.

It is important to note the essentially passive role of the distributor in this model. While it seems reasonable to assume distributors choose a service level to direct toward each brand they service, taking as given the margin they earn on the brand, I assume that the potential range in service chosen by an individual distributor will not have sufficient impact on chain-wide brand performance to affect brewers' pricing decisions. That is, I am assuming the brewer sets price to maximize profits without considering the response of the distributors' effort and its effect on demand.[12] This assumption is made to ease the computational burden in estimation.[13]

3.3 Econometric Implementation

Estimation of the model proceeds in two steps. First, a demand system is estimated using the scanner data from Dominick's Finer Foods (DFF). These demand estimates are then combined with the industry model to impute the markups at the retail and the wholesale levels (recall that we observe the price paid by the retailer). The second stage takes the observed prices paid by the retailer and backs out the costs using the estimated markups from the first stage. Then, the impact of various cost drivers are investigated. In what follows I discuss elements of each step in turn.

3.3.1 Demand

Beer is a differentiated product. Hence, any demand model used must be able to handle product differentiation. Two demand models are estimated. The first, the logit model, is nested in the second, the random coefficients logit model as implemented by Berry, Levinsohn and Pakes (1995) (BLP hereafter). Nevo (2000) and BLP provide a detailed overview of the mechanics of these models.

The convenience of the logit model for estimation is offset by the restrictions the specification places on the resulting demand function. The most unattractive aspect of the model is that cross-price effects become solely a function of product shares. The resulting substitution matrix is thus flawed, to the extent that consumers in the market substitute more readily between products that are 'closer' in the characteristic space. The random coefficients logit model (BLP) relaxes these restrictions.

---

[12]This is distinct from ignoring distributors' impacts on brand performance altogether. When demand is estimated the large size of the data set allows brand dummy variables to be interacted with distributor dummies. This incorporates the mean level of service provision by each distributor on each brand into the estimated product characteristics. When markups are calculated the precise level of service provision for each store and week is incorporated into the estimated utility of the consumer (in the 'deltas', or the non-interacted component of utility). Hence, the only substantive assumption is on the brewers' pricing decision.

[13]The assumption is on the derivative of effort w.r.t. distributors' markup times the derivative of demand w.r.t. effort. Since both these effects are likely to small at the margin, their product is likely to be very small. The extra layer of complexity required to relax this assumption would increase the computational complexity by several orders of magnitude, making estimation impractical.

The basic specification of the utility function giving rise to demand is given by

$$U_{ijkt} = x_j\beta_i - \alpha_i p^c_{jt} + \xi_{jkt} + \varepsilon_{ijkt}$$

where ε_{ij}'s capture idiosyncratic tastes that are assumed to be distributed according to a gumbel (extreme value type 1 (maximum)) distribution. Following BLP and Nevo (1998) the random coefficients, $[\alpha_i, \beta_i]$ are functions of demographic variables taken from census information (denoted D_i) and unknown consumer characteristics v_i. Thus $[\alpha_i, \beta_i]$ are given by

$$[\alpha_i, \beta_i] = [\alpha, \beta] + \Gamma D_i + \Upsilon v_i$$

D_i takes an empirical distribution implied by market level demographic data. v_i are assumed to be distributed $N(0, I)$. The matrices Γ and Υ contain the parameter values to be estimated. Γ contains the parameters on interactions between demographic variables and product characteristics. Υ contains a lower triangular matrix that represents a cholesky decomposition of the variance-covariance matrix of the v_i's.[14]

The model is completed by the inclusion of a good, denoted good zero, which captures all the outside goods. As is standard in the discrete choice demand literature I assume that consumers purchase only one unit during an observed week.[15]

If $[\alpha_i, \beta_i] = [\alpha, \beta]$ then we have the familiar logit model, which can be transformed to give the following estimating equation

$$\ln s_j - \ln s_o = x_j\beta - \alpha p^c_j + \xi_j \qquad (5)$$

where s_o is the share of the outside good. In the more general model a contraction mapping theorem is used to extract the unobserved terms, ξ_j allowing moments to be formed for estimation (see BLP for details).

In these data, price is likely to be correlated with unobservable characteristics known to consumers and price setters, but not the econometrician. As a first step to solve this problem distributor-brand dummies are included in the demand specification. These dummies pick up any time invariant unobservable common at the distributor-brand level. Together with variables for packaging, this takes care of time invariant unobservables at the product level. Variables that capture in-store promotional activity and seasonal effects will capture some time variant effects. However, in these data there is still concern about short run unobservables like promotional activity at the national or regional level, changes in shelf placement, unobserved packaging changes, community events and so on. This means that a set of instruments in place

---

[14]If the $[\alpha_i, \beta_i]$ coefficients are not dependent on the individual factors (that is $[\alpha_i, \beta_i] = [\alpha, \beta]$) then the model is just the GMM implementation of the familiar logit model (Mcfadden (1974)).

[15]Here, and throughout the paper, I use heteroskedasticity robust standard errors. The weight matrix used in these regressions is the two stage least squares weight. These weights are efficient weights if the error is spherical, if this is not true then these weights are not efficient. Given the large sample size this is unlikely to be a significant problem.

of price is desirable for consistent parameter estimates. In this study lagged and lead mean prices for the brand are used as instruments for each product. The lag and lead is four and eight weeks. These instruments are similar in spirit to those used by Hausman et al (1994) and Nevo (2000a) & (2001), the difference here being that cost shocks are assumed to be shared across time rather than across geographic markets.[16] The instruments are valid to the extent that time variant demand unobservables at the retail level have a short duration of effect, while underlying cost changes are more persistent.

In the BLP (as opposed to logit) demand model, price needs to be interacted with demographic data (so as to form the α_{ik}). The demographic data used here are household income and age taken from the 1990 census. The characteristics were matched to the zip code of each store to create the appropriate set of demographics for each observation.[17]

3.3.2   Distributor and Brewer Costs

An objective of this paper is to measure any cost differences between markets that may be foreclosed and those that are not. To do this I take the markup, computed from earlier demand estimates, away from the price to retail, leaving costs, and then investigate the determinants of cost variation between different brewer-distributor combinations. The two econometric issues are to correctly instrument for an endogeniety problem that arises from the fact each brewer has to take into account the costs of several distributors in making each pricing decision and, secondly, to construct the standard errors of the second stage estimates, taking into account variance introduced from the use of first stage coefficient estimates.

The first-order condition in equation (4) summarizes the economic problem of the brewer. To use this first-order condition I need $\frac{\partial s_{lk}}{\partial p_m^c}$ which is available from the demand system. I also need $\frac{\partial p_m^c}{\partial p_j^r}$. This is derived from the retailer's first-order condition using the implicit function rule.

To set one price for the chain the brewer must take into account the different costs of servicing each store of the chain. Since each distributor is serving only a subset of stores in the DFF chain, when the brewer sets the price for the chain, the costs of several different distributors must be taken into account, according to the demand conditions in the individual markets they serve. Pricing equations for upstream firms dealing with chains will usually have this form, since the chains tend to buy at the same price across stores. Thus the measurement issues confronted here will be found in industries other than the beer industry.

---

[16]Recall that, since products are priced on a chainwide basis, we have only one market (in the sense of Hausman et al, or Nevo) at any point in time.

[17]Where a store stood on the boundary of a zip code it was assumed customers where drawn from each zip code with equal probability. To aid in sampling in the simulation estimator the income data were fitted to a lognormal distribution and the age data to a gamma distribution, following the methodology used in BLP.

To recover costs as a linear function of the markup and price we need to invert the system of equations (4). A difficulty is introduced in this context by the introduction of distributor costs.[18] Since $\left(c_{jd(k)}^d + r_{jd(k)}^d\right)$ in equation (4) is a market specific variable we cannot "pop" it out cleanly in the same way as a chain wide variable can be recovered. This can be seen if we rewrite equation (4) as follows:

$$\sum_k M_k s_{jk} + \sum_{l \in B^b} \left(p_l^r - c_l^b\right) \left[\sum_k M_k \sum_m \frac{\partial s_{lk}}{\partial p_m^c} \frac{\partial p_m^c}{\partial p_j^r}\right]$$
$$-\underbrace{\sum_{l \in B^b} \left[\sum_k \left(c_{ld(k)}^d + r_{ld(k)}^d\right) M_k \sum_m \frac{\partial s_{lk}}{\partial p_m^c} \frac{\partial p_m^c}{\partial p_j^r}\right]}_{part\ 3} = 0 \ \forall j \in B^b$$

which, letting the ij^{th} element of $\mathbf{G}_k = \sum_m \frac{\partial s_{jk}}{\partial p_m^c} \frac{\partial p_m^c}{\partial p_i^r}$, can be rendered in matrix notation as

$$\mathbf{B} + \left[\sum_k M_k \left(\mathbf{G}_k \bullet \Pi\right)\right] \left(\mathbf{p} - \mathbf{c}^b\right) - \underbrace{\sum_k \left[M_k \left(\mathbf{G}_k \bullet \Pi\right) \left[\mathbf{c}_{d(k)}^d + \mathbf{r}_{d(k)}^d\right]\right]}_{part\ 3} = 0 \qquad (6)$$

where $-\left[\left(\sum_k M_k \mathbf{G}_k\right) \bullet \Pi\right]^{-1} \mathbf{B}$ gives the markups on the prices to retail. The Π term is a J-by-J accounting matrix that makes sure that only the set $l \in B^b$ is considered in the first-order condition ($\bullet$ is used to denote element-by-element multiplication of matrices).[19] The inclusion of the $c_{ld(k)}^d + r_{ld(k)}^d$ terms in the summation over k in part 3 of equation (6) is the problem. Were this $\mathbf{c}_{ld(k)}^d + \mathbf{r}_{ld(k)}^d$ able to be factored out of the summation we would be able to include it in the $(\mathbf{p} - \mathbf{c})$ vector in the second part of the equation.[20]

Rearranging equation (6) gives the estimating equation for a specific product j

$$p_j - \mu_j(x, \theta, \xi) = \alpha c_j^b + \beta \sum_k W_{kj}(x, \theta, \xi) \left[c_{ld(k)}^d + r_{ld(k)}^d\right] + \varepsilon_j \qquad (7)$$

where

$$W_k(x, \theta, \xi) = \left[\sum_k M_k \left(\mathbf{G}_k \bullet \Pi\right)\right]^{-1} M_k \left(\mathbf{G}_k \bullet \Pi\right) \mathbf{1}_{[J \times 1]}$$

These weights capture the marginal importance to each product of the stores served by each distributor. $\mu_j(x, \theta, \xi)$ is the markup on the price to retail.

The concern is that the inclusion of the $\frac{\partial s_{lk}}{\partial p_m^c} \frac{\partial p_m^c}{\partial p_j^r}$ terms in the regressor, $\sum_k W_{kj}(x, \theta, \xi)(c_{jk}^d + r_{ld(k)}^d)$, will give rise to an endogeniety problem (prices should be correlated with cost

---

[18] There is also a minor issue arsing from the fact that the price in the demand equation is price per 12oz. I have to trivially adjust equation (4) to model the brewer as choosing a per 12oz price, if the demand estimates are to inform the estimation. Thus estimated costs etc are per 12oz.

[19] All vectors are column vectors, unless otherwise noted.

[20] Note that $[\mathbf{A} \bullet \Pi] = \sum_k M_k \left(\mathbf{G}_k \bullet \Pi\right)$

unobservables and so functions of prices are likely to be as well). Since $\frac{\partial s_{lk}}{\partial p_m^c} \frac{\partial p_m^c}{\partial p_j^r}$ is only a function of demand level parameters[21], we are worried about endogeniety here if a cost unobservable is feeding through to the price to the consumer (as would be expected if firms are profit maximizing) or, more generally, if the cost unobservable is correlated with the level of the utility derived from the product. An example would be the hypothetical 'Green Beer' which differs from ordinary beer only in its colour (which is unobserved). If it costs more to produce green beer, but people like to drink it (say on special occasions like St. Patrick's Day), then we would expect $\frac{\partial s_{lk}}{\partial p_m^c} \frac{\partial p_m^c}{\partial p_j^r}$ to be correlated with the cost unobservable associated with green beer, since the s_{lk} would be correlated with the cost unobservable.

Consider a scenario where the $\frac{\partial s_{lk}}{\partial p_m^c} \frac{\partial p_m^c}{\partial p_j^r}$ terms are such that those stores served by distributors with high $\left[c_{ld(k)}^d + r_{ld(k)}^d \right]$ terms are also ones measured as having higher marginal importance (higher weights) for the costly, but popular, 'green beer'.[22] Then this story would lead to a positive bias in the estimates of β.[23,24]

Happily, I can deal with this problem using appropriate instruments. In this instance the $\left[c_{d(k)}^d + r_{d(k)}^d \right]$ terms give us potential instrumental variables, which, by construction, are correlated with the endogenous regressor but orthogonal to the error.

The last issue to confront is making the adjustments for the two step estimation procedure being adopted here. The empirical strategy requires preliminary estimates of demand as an input into the estimation of costs. The difficulty in implementing the cost estimator in this context is that it is being done as the second step in a two step process. This has no impact on the consistency of results but, since the estimator turns out not to be adaptive, will affect the standard error estimates. Thus the fact that estimated parameters are being used in this second step needs to be directly incorporated into the estimates of the variance-covariance matrix. The method for making this adjustment is discussed in Wooldridge (2002).

4 Data Description

The Chicago Distribution System   The new data set introduced in this paper documents the exclusive sales territories of 12 brewers' distributors in Illinois in 1994.

---

[21] To see this note that $\frac{\partial s_{lk}}{\partial p_m^c}$ is already available from the work on retailer costs and $\frac{\partial p_m^c}{\partial p_j^r}$ is recoverable without using data on p^r.

[22] Say the $\left[c_{ld(k)}^d + r_{ld(k)}^d \right]$ term was an exclusive dummy. This might occur due to exclusive distributors providing relatively better support for more expensive brands, or due to some territory specific component of peoples preferences.

[23] This is distinct from the $\left[c_{ld(k)}^d + r_{ld(k)}^d \right]$ terms being correlated with unobserved costs. Consider a story where all products go through all the distributors. Unobserved product cost is uncorrelated with distributor characteristics, $\left[c_{ld(k)}^d + r_{ld(k)}^d \right]$, but the weights will still be correlated.

[24] Exploration of this issue with simulated data found that it was easy to find parameter settings under which the bias was large enough to be economically significant.

14

These 12 brewers comprises all of the major domestic brewers and most of the major imported brands. It also includes some smaller brewers. The selection of brewers provides a sample covering of over 90% of the beer sold through the Dominick's Finer Foods supermarket chain, the source of the demand data used in this study.

Data on exclusive sales territories is collected by the Illinois Liquor Control Commission. Brewers are required to keep an up-to-date register of the sales territories of their distributors with the Commission. These registration documents must specify the relevant brands, the precise geographic area covered by the territory and the time period of the distribution agreement.[25]

Insert figure 4 here

Figure 4 shows the result of the data collection process. For each brewer we can draw a map of the territories of their distributors. The maps in Figure 4 are stylized for ease of illustration. Each square corresponds to a page in the Rand-McNally street directory (3.5 x 4.5 miles). Shaded squares indicate the presence of at least one Dominick's Finer Foods supermarket. Comparing the territories of the three brewers illustrates the variation in territory configuration across brewers (note that none of the distributors are shared between these brewers).

An examination of Figure 4 and the underlying data is quite informative. The following stylized facts quickly emerge:

- Different brewers do not divide Chicago into the same territories;

- The sales territories for a single distributor serving two brewers tend not to overlap perfectly;

- Different brewers use different numbers of distributors; and

- Territories tend to remain the same over the course of several years.

It is worth noting some caveats on the territory data. The files from which the data were drawn occasionally had missing documents. In most cases this problem was able to be overcome since, given territories are contiguous, the territory for a distributor will be defined by the boundaries of the neighboring distributors (recall territories may not overlap). This feature of the data, together with the complete lists of distributors serving a brewer contained in the files, meant that the territory of the problematic

---

[25]The availability of these data was made possible by the generosity of the Commission in allowing access to their archives and their patience with questions about how the data were organized. The data exist in hard copy form in Illinois.

The data supplied by the brewers is often far more detailed than is required, at times even including the complete text of the distribution agreements. To an economist, these contracts tend to be remarkably vague documents. However, it is interesting to note that where they are most precise is in describing the data collection and data sharing responsibilities of the distributor. Actual performance benchmarks and targets tend to be qualitative at best.

distributor could be described. However, in two instances the files had too many missing documents to enable the recovery of all the territories for a brewer's brands. In these instances the brewer was dropped (this resulted in Heineken and Guinness, both major imported brands, being omitted from the sample).

Information on the presence of exclusive dealing arrangements was collected by talking to industry insiders and comparing various industry sources. It is accepted in the Chicago industry that there are three different distribution "systems" in the city: the Anheuser Busch system, which is completely exclusive; the Miller system, which is somewhat exclusive; and the "all others system" which is a catch-all for distributors who are not in the Anheuser Busch or Miller systems. Table 1 contains information on the relationships between the 42 distributors in the data and the 12 brewers. Table 1 shows which distributors deal for which brewers and also whether the relationship is exclusive (indicated by a shaded rectangle). The Barton and Heilman distributors who are indicated as being exclusive are marked this way as, according to the data, the are only serving one brewer. This is distinct from having an exclusive arrangement with the brewer. Since there is no corroborative evidence to suggest that there was an exclusive arrangement between these distributors and the brewers, they are not considered as being party to such an arrangement.[26]

Insert table 1 here

Demand Data from Dominick's

Demand data is drawn from the Dominick's Finer Foods data set collected by the Chicago GSB Marketing Department.[27] This is a data set that provides weekly scanner data on consumer purchases at stores in the Dominick's chain across a wide range of products. This study uses data on beer sales during 1994, drawn from this, much larger, data set.

Dominick's Finer Foods is the second biggest supermarket chain in greater Chicago (Jewel being the biggest), having around a 20% market share, depending on the year and source (Hoch et al (1994)).[28]

For each beer, the data provide a considerable amount of information. The brand, store and week are recorded, together with the price at which the beer was sold to consumers. When promotional activity occurs this is also recorded.[29] Lastly, the

---

[26]Indeed, examination of the available data on distributor characteristics suggest that several of these distributors as highly likely to be lemons, in the sense of being very poor quality.

[27]This data set is avaliable online at http://gsbwww.uchicago.edu/research/mkt/Databases/DFF/DFF.html

[28]The importance of Dominick's to individual distributors varies. Using reported revenues from the 1994 Beverage Marketing Directory, the Dominicks stores make up around 3% of the business of the average distributor of Anheuser Busch or Miller products. This number varies however, the maximum is around 10% and the minimum less than 1%. Since revenue is self-reported and recorded in ranges the numbers are only indicative at best.

[29]The documentation supporting the DFF data comment that this variable may fail to report some promotions. There is no way of telling whether this is a problem in the data used here.

data records the average prices paid by the retailer for their inventory of each beer.[30] Thus, the this aquisition price and the true economic cost may be slightly different.

In the data this average aquisition cost (or price to retail) varies across stores by less than 1¢ for 98% of UPCs. 76% of UPCs show no variation. The remaining 22% is accounted for by the way DFF computes the cost of inventory (Dominick's reports the average acquisition cost of inventory. Hence, where stores are running through inventory at slightly different rates we would expect slightly different average acquisition costs). Given the legal requirements that beer distributors sell only to stores in their territory, this pattern of wholesale price indicates that the brewer is coordinating the price to retail across distributors. As noted earlier this confirms the anecdotal evidence collected from distributors. This institutional feature is an important aspect of the model used in this study.

The same analysis on the price to consumers suggests that DFF sets prices centrally, although with more variation in prices across stores at the UPC level as compared to the price to retail. Once again this is incorporated into the model.

The data on beer are provided at the UPC level, by store, by week. This results in a very large data set. Organising the data requires drawing a balance between the size of the data set and the desire to use the actual transaction prices paid by consumers rather than resorting to some weighted average measure of price. To do this a product is defined as a set of observations where the brand, serving size, price to consumer, week, store and promotional status are the same.

The price to retail is not required to be the same. Across stores, this can vary slightly due to stores going through inventory at slightly different rates (refer back to the earlier discussion). For each product the price to retail is a quantity weighted average. However, as the earlier discussion suggests, these averages are over prices that very similar.[31] Due to the cost of inventory issue, this averaging procedure will always be required, regardless of how the data is organised.

Table 2 provides some descriptive statistics of the demand data. The data contain 138213 observations, where an observation is the weekly purchases of a product (as defined above), in a store, in a week. The data are drawn from 71 stores spread across greater Chicago (see Figure 4 for some indication of the geographical distribution). 73 brands of beer are used, which together account for over 91% of the beer sold at Dominick's stores in 1994. Overall Dominick's sells in excess of 3 million gallons of beer each year. These 73 brands are drawn from 12 brewers, listed (together with their distributor allocations) in Table 1.

Insert table 2 here

---

[30]Beer distributors vistit major accounts like Dominicks several times per week, so it is likely that the impact of this averaging of the price paid by the retailer is minimal. In particular, containimation from aquistions made in previous weeks should be small.

[31]Differences, where they occur, are usually at the 3rd or 4th decimal place. That is, tenths or hundredths of a cent, when price is quoted as per 12oz.

Table 2 shows descriptive statistics of the price paid by the consumer for a 12oz serve of beer (equivalent to a bottle). The mean price is 60¢ with a large standard deviation (17¢). The minimum price per bottle in the data is 19.8¢ and the maximum price is $2.97. Table 3 also shows corresponding statistics for the price paid by the retailer. Mean price to retail per 12oz is 50¢ with a standard deviation of 13.5¢. The minimum price to retail is 15.3¢ and the maximum is $1.11. These statistics suggest an average retail markup (ignoring shelf costs) of 10c per bottle, a conclusion bourne out by closer examination of the data. The mean difference between price to retail and the price paid by consumers per bottle is 9.7¢, with a maximum of $2.49 and a minimum of -34¢, reflecting a sale. Promotional activity is quite common in this data set with just under $\frac{1}{3}$ of all observations being recorded as a sale item.

The data provides information on the total number of customers visiting the store for each week, in about $\frac{2}{3}$ of cases. This is used as the market size where available. Where this is not available the number of customers visiting the store is forecast from regressions done on those stores where these data is available. The regressions are run using an OLS specification, with population, median age, median household income and a constant as regressors. The demographic data are for the surrounding zip code taken from the 1990 census. These regressions are done for each week, so week specific fixed effects (due to holidays, for instance) are taken into account by the constant. The standard approach in demand estimation has been to make an *a priori* assumption about the geographical market size and then infer the population size using the assumption and census data.[32] Given the richness of the information in this data set, this paper departs from the traditional approach and lets the data guide the market size.[33]

Once the size of the market has been set, shares can be calculated from the data on quantity. Here, an assumption is made; that consumers buy only one unit of beer, where a unit is defined as at the UPC level. That is, I assume that if I see a six pack of Bud sold then that was the only beer the consumer bought. This assumption allows shares for each product to be calculated as the number of units sold divided by the market size. To make everything add up to one, an outside option is introduced with collects all those consumers who chose not to buy any beer. Table 2 shows the average share of a product to be 0.05% with a standard deviation of 0.09%. The smallest share is 0.003% and the largest is 3.7%.

Data on beers' alcohol by volume (ABV) and calories per 12oz, was collected from industry publications from around the time of the study (Brewing Matters (1995) and (1996)) or by contacting brewers directly. Table 2 shows that the average ABV among the brands in this sample is 4.4%, while the average amount of calories per 12oz is 130.

---

[32]See BLP (1996), Nevo (1997) and (2000), Mortimer (2003) and Villas Boas (2002), for a few examples of this approach.

[33]That said, there is still an implicit assumption that if the price of a brand goes down this will not cause people to go to the supermarket, who otherwise would not have. Chintagunta et al (2002), provide evidence in support of this assumption in the DFF data

5 Empirical Results

5.1 Demand

Several different demand specifications are considered in this study, all nested in the BLP framework considered above. These estimated demand systems are identified from variation in products' shares due to variation in the composition of characteristics available across the different choice sets. In the simple logit specification (equation 5) the impact of a products characteristics on its own and other products' shares is captured by the $\ln s_j - \ln s_o$ term that makes up the left-hand-side of the estimating equation. In the full BLP specification the relationship is algebraically more complex, but similar in flavor. As described in section 3.3, identification of the price coefficients is aided by the use of instruments and distributor brand dummies. Hence, the component of price variation used for identification here is within-brand and within-distributor-territory variation.

As a starting point, the logit specification, which lacks interactions, was run. This specification has several attractions. First, it is simple. Second, it provides a useful comparison for the full model with interactions. The problem with the logit model is that it generates potentially unrealistic substitution patterns. Thus, running a full BLP style specification with interactions between individual characteristics and product characteristics is likely to be informative, if for no other reason than as a robustness check on the appropriateness of the logit specification.

Insert table 3 here

Table 3 contains the results from the logit demand models. In the OLS specification (specification A) the estimated price coefficient is -4.64, whereas when instruments are used for price (specification B) the estimated coefficient on price is -6.51. This estimate corresponds to a price elasticity for beer of around 3.4, depending on the beer, store and week. The results indicate that instrumenting for price has a significant impact on the coefficient estimate, correcting for the expected positive bias in the price coefficient. The effect of a promotion, independent of the price effect is found to be small in the instrumented logit case, being equivalent to a five cent drop in the price of a six pack of beer. The coefficients on serving size and serving size squared suggest that the smaller size units are preferred in the relevant region (six packs are preferred to cases).[34][35]

Table 3 also provides demand estimates for BLP demand models. These models allow for consumer heterogeneity through random coefficients. The first specification, C,

---

[34]Here, and throughout the paper, I use heteroskedasticity robust standard errors. The weight matrix used in these regressions is the two stage least squares weight. These weights are efficient weights if the error is spherical, if this is not true then these weights are not efficient. Given the large sample size this is unlikely to be a significant problem.

[35]Joint estimation of the demand system with the retailer's pricing equation was also examined. For the logit specification this did not change the demand estimates in any significant fashion. Estimation with BLP was only feasible after over 80% of the data was discarded, due to the excessive time needed to compute markups at each iteration. As a results the estimates were not particularly informative.

allows the coefficient on price and a constant (interpreted as the value of the outside option) to vary with the inverse of income. The inverse of income was used since as income gets larger we would expect the importance of price to decrease; this functional form seems a natural way to incorporate this. Thus the price coefficient is $\left(\alpha + \frac{\beta}{y_i}\right)$ where $\alpha = -6.59$, $\beta = 0.29$ and y_i is household income divided by 2500. For a consumer with annual household income of \$10,000 this corresponds to a price coefficient of -6.51. As income increases, the coefficient converges to -6.59. This implies that income does not change a consumer's sensitivity to price very much. At the same time the value of the outside option is decreasing with income, although the magnitude of this effect is very small. In a joint test, the null hypothesis that the coefficients on the interacted terms are zero is unable to be rejected at any reasonable significance level, implying that the logit model cannot be rejected. This is also the case for specification D, which uses randomness generated from a normal distribution, capturing unknown sources of consumer heterogeneity.[36]

Since specifications C and D are unable to reject the logit specification of demand, and the magnitude of the impact of consumer heterogeneity implied by these specifications is small, I proceed with the logit model of demand with instruments for price (specification B) in the rest of this paper.

5.2 Testing for Foreclosure

5.2.1 Foreclosure in Promotional Aptitude

To begin to investigate the possibility of foreclosure in promotional aptitude the standard test was run, as described in Section 3.1. This test compares the demand effect of facing a Miller exclusive to the demand effect of not facing a Miller exclusive, using Anheuser Busch as a control group to manage any market specific effects. The results of this test are contained in specification A of Table 4. The basic demand specification used above is also used here, except that instead of using brand-distributor dummies only brand dummies are used. There are two variables of interest, '"Foreclosed" Stores - Potentially Foreclosed', which is a dummy for when a potentially foreclosed brewer's distributor faces a Miller exclusive, and '"Foreclosed" Stores - AB' which is a dummy for an Anheuser Busch distributor in the same position (the control group). The estimates indicate that both Anheuser Busch and potential foreclosed brewers seem to get better product performance in markets where Miller is exclusive. Further, there appears to be no significant difference between the magnitudes of this effect. On this basis, there appears to be no evidence of foreclosure.

As discussed in Section 3.1, an identification problem arises when running the test in this way. In the basic form of the test the Anheuser Busch control group is being relied on to control for the market specific demand shifters and the effect of a (potentially) more active Miller distributor. Since cross-product substitution effects are likely to

---

[36]Chi-Squared statistics for the joint test of the null that the coeffiecients are zero are 0.0050715 for specificatiom D and 1.9542 for specification C.

be different for different products this may be unsatisfactory for dealing with this second effect. With this in mind, a reformulated version of the test is also run.

This reformulated test compares product performance when serviced by a distributor who is used in markets where Miller is exclusive and also where Miller is not exclusive (class two distributors) to product performance when serviced by a distributor who only services markets where Miller is exclusive (class three distributors). Performance is compared across the two types of distributors in markets where Miller is exclusive. This holds constant market specific effects that may be correlated with the Miller distributor being exclusive and the impact of a more active (exclusive) Miller distributor.

Insert table 4 here

Table 4 shows the results of this form of the test. The variable of interest is denoted 'All Exclusive Markets'. This variable is a dummy for a product serviced by a class three distributor. If the coefficient on this dummy is negative then evidence consistent with foreclosure exists. In specification B, which is the test as described above, the coefficient on 'All Exclusive Markets' is strongly positive, rejecting the conclusion that foreclosure in promotional aptitude is present in this market. Indeed, the magnitude of the coefficient suggests that the class three distributors are providing extra promotional service equivalent to a 3.5¢ price reduction on a bottle of beer.

A possible concern in this analysis is that class two distributors may be switching promotional effort away from their exclusive markets toward their non-exclusive ones. To see whether this might be an issue specification C was run. Specification C includes all the markets that class two distributors service, not just those in which Miller is exclusive. If the this substitution in effort were a problem it might be expected that the coefficient on 'All Exclusive Markets' would be smaller. Instead it gets larger, both in absolute size and in terms of the equivalent price reduction of the extra promotional benefit provided by the class three distributors. This result is interpreted as suggesting that a substitution effect is not likely to be distorting the conclusion.

5.2.2 Foreclosure in Costs

Before delving straight into the test for foreclosure in terms of distribution costs it is useful to examine data that might allow us to reject the hypothesis more easily.

One way to cast doubt on the hypothesis is to compare the production technologies of exclusive and non-exclusive distributors. If the foreclosure hypothesis, or an investment type theory, were correct we would expect to be able to see differences in the production technologies used by each type of distributor. An investment story (e.g. Besanko and Perry 1993) would suggest that these differences were a function of better investment by brewers, while a foreclosure story would suggest that these differences are a function of innate distributor characteristics.

While ideal data is not available for a full econometric treatment of this issue, some data does exist which is informative about the characteristics of distributors in Greater Chicago. Figures 5 and 6 show the number of employees and the number of trucks that Anheuser Busch and Miller distributors have, compared to other distributors. The data are drawn from the 1994 Beverage Marketing Directory, which contains survey data collected in the middle and later part of 1993. The data are collected as ranges, with respondents asked to tick the bin into which they fall. Thus, the data contain a noisy view of the industry.[37]

Insert figures 5 and 6 here

Despite this, the data paint an interesting picture. 50% of Anheuser Busch distributors have between 30 and 39 employees, this being the modal response for this group. Other distributors are less homogenous in terms of their numbers of employees. This same pattern, of greater homogeneity in Anheuser Busch distributors as compared to other distributors is also present in the trucks data. This pattern is maintained when the sample is extended to include all the distributors in the state of Illinois.

This relative homogeneity in the labor and capital used by Anheuser Busch (and to a lesser extent Miller) distributors seems consistent with both foreclosure and investment type theories of why lower cost distributors are more likely to be exclusive.

Having examined these data, it remains to test for foreclosure in costs. To do this I use the industry model and the demand system to infer costs, as described in section 3. Then I project these inferred costs down onto cost drivers using equation (7). Recall that to implement the test, I define each store as a market. If the store is served by an exclusive Miller distributor then it is classed as a market with an exclusive distributor. These markets are allocated a dummy which is applied to product observations from brewers other than Anheuser Busch or Miller. These dummies form part of the $\left[c_{ld(k)}^d + r_{ld(k)}^d \right]$ terms in equation (7). In the specification, costs, being price to retail minus markup, are projected onto product and brewer characteristics and a dummy for when there is an exclusive distributor in the market.[38] This last variable is called "'Foreclosed" Stores'. If the coefficient on this variable is negative then the costs of distributors who serve brewers who never use an exclusive distributor are lower in markets where there is an exclusive distributor, rejecting the foreclosure argument.

---

[37] The distributor details collected from the ILCC registration documents were matched to distributors in the directory. For some distributors this was difficult due to the existence of different business names. It is also possible that missing distributors did not respond to the Beverage Marketing Directories survey. Unfortunately 5 of the 42 distributors where unmatched, 2 of which were Anheuser-Busch distributors.

[38] This dummy is only applied to brewers that never use an exclusive distributor - the estimation is done on only those brewers who might be foreclosed (thus estimation is done within the set of potentially foreclosed brewers). The "dummies" are really sums of dummies, as required by the aggregation process discussed in section 5.3 and equation (7). The instruments used for the resulting weighted sum are is the proportion of stores stocking the product with an exclusive Miller distributor.

Identification of the coefficient on the dummies that comprise the $\left[c_{ld(k)}^{d} + r_{ld(k)}^{d} \right]$ terms is a challenge. The problem is due to the fact that a brewer, when they make their pricing decision, is taking into account the costs of several different distributors. Some of these distributors are in exclusive markets, while others are in non-exclusive markets. There are two potential sources of identification of the impact of different distributor types. The first takes advantage of the fact that different brewers and brands have different exposure to exclusive and non-exclusive markets. This is due to variation in stocking decisions and also variation in the relative importance of the different types of markets due to, for example, consumer preferences. This type of variation relies heavily on the instrument on the $\sum_k W_{kj}(x, \theta, \xi) \left[c_{ld(k)}^{d} + r_{ld(k)}^{d} \right]$ term in equation (7) being orthogonal to any unobserved product characteristic. Thus, controlling for product characteristics is especially important.

The second source of identification comes from within-product variation across time in the importance of the two types of distributors (those in exclusive markets versus those in non-exclusive markets). This source of identification seeks to exploit demand shock hitting different parts of the distribution chain differentially. A simple example might be that if a football game is occurring in a region dominated by exclusive markets, we might expect to see the relative importance of distributors servicing exclusive markets increase. This would mean that the costs of distribution through these distributors would be more strongly reflected in the pricing decisions of the brewer for that week.

Thus identification can be thought of as coming from some combination of across- and within-brand sources.

The remaining issue in identification is the nature of the instrument used on the $\sum_k W_{kj}(x, \theta, \xi) \left[c_{ld(k)}^{d} + r_{ld(k)}^{d} \right]$ term in equation (7). The instrument is the proportion of stores that a product is stocked at that are served by a Miller exclusive. A sufficient condition for the validity of the instrument is that the stocking decision be independent of whether the store is exclusive or not.

Insert table 5 here

The results in Table 5 show the estimates from projecting costs down onto various cost drivers. The coefficients on the "'Foreclosed" Stores' terms are of primary interest as they capture differences in distribution costs in the two types of markets. The rest of the specification captures product-specific cost drivers. The constant (per unit/SKU) term is controlling for differences in serving size (e.g. 6 packs vs 12 packs). Light, Ice, NA (for non-alcoholic) and Import are dummies for different types of beer. ABV (alcohol by volume) and Calories are two important beer characteristics that reflect different production processes and ingredient mixes. Lastly, specifications 1 and 2 include brewer dummies, while specifications 3 and 4 include brand dummies. These variables are intended to capture any returns to scale or other brewer (or brand) specific cost characteristics that might be present.

23

In specification 1 (with no instruments) the estimated coefficient on the "'Foreclosed" Stores - Potentially Foreclosed' term is 0.011. This suggests that in markets with a Miller exclusive the potentially foreclosed brewers have distribution costs that are on average 1¢ higher than in markets where Miller is not exclusive. Specification 2 shows the impact of instruments on the coefficient estimate on the "'Foreclosed" Stores - Potentially Foreclosed' term. The estimate changes to 0.0007, which is not significantly different from zero. To evaluate the effect of market specific cost variables, this is compared to the coefficient on "'Foreclosed" Stores - AB' which captures the effect on Anheuser Busch's costs of serving an exclusive market. The results do not indicate that there is evidence of foreclosure, instead it appears that the potentially foreclosed brewers have relatively lower costs in exclusive markets than their Anheuser Busch counterparts.

The results in specifications 1 and 2 use both the across- and within-brand sources of identification. This may be problematic as we might expect higher cost brands to be higher quality and better able to compete in markets where Miller is exclusive, hence more likely to be placed in these markets. If this were the case then the instrument used here may not be orthogonal to the unobserved cost terms, leading to a positive bias in the estimates. To control for these effects the specification was re-run replacing all the brand specific variables with a full set of brand dummies. Thus only the second, within-brand source of identification is exploited.

Specifications 3 and 4 shows the results. In specification 4 (with instruments) the "'Foreclosed" Stores' coefficient for the potentially foreclosed brewers changes to -0.009, while the coefficient for Anheuser Busch changes to -.006. Taking the difference between these two estimates and testing the hypothesis that this is not different from zero, suggests that, for the potentially foreclosed brewers, the cost advantage from serving markets with a Miller exclusive is attributable to differences in market specific effects and not reflecting any underlying cost difference at the distributor level. Thus, the hypothesis of foreclosure in costs remains rejected by the test.

6   Concluding Remarks

This paper has provided evidence that the use of exclusive distributors in the beer industry should not lead to concerns about foreclosure, at least in metropolitan markets.

An issue left in the background by this paper is the influence of scale on the structure of distribution. It would be expected that distributors have considerable fixed costs to cover, in addition to their marginal costs. In discussions with distributors, rent of premises and insurance were stated as significant costs that might be considered fixed. If fixed costs are considerable, then having an exclusive distributor would be more attractive to brewers that can provide a large market share.

Such a story has interesting implications. If any cost advantage from exclusivity is due to foreclosure effects, then fixed costs and accompanying returns to scale would

help preserve foreclosure by discouraging new distributors from entering the market.[39] New entrants would have to cover potentially high fixed costs, and to do this may need to persuade several of the smaller brewers to break their existing distribution agreements and switch. Thus, entry by talented, low cost, distributors would be impeded.

In the more likely event that exclusive distributors attract higher investment levels, then scale is likely to be even more important in restricting access to any cost advantages. If a smaller brewer cannot provide the scale to support a distributor by themselves then they will be forced to share, removing incentives to invest in lowering the marginal costs of distribution.

Both these stories suggest that the next step in understanding the organizational structure of distribution in this market is to get estimates of fixed costs and better data on the contractual durations and obligations of distributors. This would be a first step to unpacking questions of how distributors and brewers match up. To properly understand the impact of specific interventions by a regulatory authority, modelling this matching process would be important.

Lastly, more work on the methods of identifying different reasons for using exclusive distributors would be extremely useful. Here I am able to reject foreclosure effects in a simple test, however it would be interesting to confront the investment hypothesis with a similar test. This would likely require data over a longer time period and over several distinct markets, so as to observe cross-market and -time variation in the structure of distribution. More detail on distributor inputs to production would also be helpful. It is also likely that some theoretical work would be required to nest different models within the same broad framework, so as to facilitate empirical evaluation.

While many interesting issues remain to be investigated in this industry, this paper makes a preliminary step in establishing that exclusive distributors are not foreclosing smaller brewers in this market, implying antitrust intervention is not warranted. Developing identification strategies for the exact motivation for using exclusive distributors and studying the matching process behind the choice of distributor is left for future research.

---

[39]See Rasmusen et al (1991) for a related theoretical discussion.

References

[1] Areeda, Phillip and Herbert Hovenkamp (2002),*Fundamentals of Antitrust Law Volume 2*, Aspen Law and Business, New York.

[2] Bernheim, B. Douglas and Michael Whinston (1998), Exclusive Dealing, *Journal of Political Economy*, 106(1), 64-103.

[3] Berry, Steve, James Levinsohn and Ariel Pakes (1995), Automobile Prices in Equilibrium, *Econometrica*, 63(4), 841-890.

[4] Besanko, David and Martin Perry (1993), Equilibrium Incentives for exclusive dealing in a differentiated products oligopoly, *RAND Journal of Economics*, 24(4), 646-667

[5] Bloomberg News (1998), Anheuser Busch, Miller Settle Lawsuits, printed in *Los Angeles Times*, May 20, Section D, p3.

[6] Bonnet, Celine, Pierre Dubois and Michel Simioni (2005), Two-Part Tarrifs versus Linear Pricing between Manufacturers and Retailers: Empirical Tests on Differentiated Products Markets, IDEI Working Paper 370.

[7] Bork, Robert (1978), *The Antitrust Paradox: A Policy at War with Itself*, Basic Books, New York.

[8] Brenkers, Randy and Frank Verboven (2002), Liberalizing a Distribution System: the European Car Market, *mimeo*, K.U. Leuven.

[9] Brewing Matters (1995), *1995 Great American Beer Festival Program*, Association of Brewers, Colorado.

[10] Brewing Matters (1996), *1996 Great American Beer Festival Program*, Association of Brewers, Colorado.

[11] Butler, Rachel (1999), Harness the Power, *Beverage World*, September 1999, p40.

[12] Canedy, Dana (1997), Anheuser Busch Queried on Distribution Practices, *New York Times*, October 3, Section D, p17, column 1.

[13] Chicago Sun-Times (1997), Antitrust Probe Targets Busch, *Chicago Sun-Times*, October 2, Financial p2.

[14] Chintagunta, Pradeep, Jean-Pierre Dube and Vishal Singh (2002), Balancing Profitability and Consumer Value: An application to zone-pricing by a supermarket chain, working paper, University of Chicago.

[15] Chipty, Tasneem (2001), Vertical Integration, Market Foreclosure, and Consumer Welfare, *American Economic Review*, 91(3), 428-453

[16] Comanor, W.S. and H.E. Frech (1985), The Competitive Effects of Vertical Agreements, *American Economic Review*, 75, 539-546.

[17] Culbertson, W. Patton and David Bradford (1991), The Price of Beer: Some Evidence from Interstate Comparison, *International Journal of Industrial Organization*, 9(2), 275-89.

[18] Hausman, Jerry, Gregory Leonard and J. Douglas Zona (1994), Competitive Analysis with Differentiated Products, *Annales D'Economie et Statistique*, 34, 159-180.

[19] Hellerstein, Rebecca (2006), A Decomposition of the Sources of Incomplete Cross-Border Transmission: The Case of Beer, working paper, FRBNY.

[20] Hoch, Stephen, Xavier Dreze and Mary Pruk (1994), EDLP, Hi-Lo and Margin Arithmetic, *Journal of Marketing*, 58, 16-29.

[21] Jordan, John and Bruce Jaffe (1987), The Use of Exclusive Territories in the Distribution of Beer: Theoretical and Empirical Observations, *The Antitrust Bulletin*, 32(1), 137-64.

[22] Klein, Benjamin and Kevin Murphy (1988), Vertical Restraints as Contract Enforcement Mechanisms, *Journal of Law and Economics*, 31(4), 265-297.

[23] Lenderman, Max (1997), All Stressed Out ... and Nowhere to Go, *Beverage World*, August 1997, p53.

[24] Nevo, Aviv (2000a), Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry, *RAND Journal of Economics*, 31(3), 395-421.

[25] Nevo, Aviv (2000b), A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand, *Journal of Economics and Management Strategy*, 9(4), 513-548.

[26] Nevo, Aviv (2001), Measuring Market Power in the Ready-to-Eat Cereal Industry, *Econometrica*, 69(2), 307-342.

[27] Mathewson G.F. and R.A. Winter (1987), The Competitive Effects of Vertical Agreements: Comment, *American Economic Review*, 77, 1057-1062.

[28] McFadden, Daniel (1974) Conditional Logit Analysis of Qualitative Choice Behavior, in P. Zarembka (ed.), *Frontiers in Econometrics*, 105-142, Academic Press: New York, 1974.

[29] Mortimer, Julie (2002), The Effects of Revenue-Sharing Contracts on Welfare in Vertically-Separated Markets: Evidence from the Video Rental Industry, mimeo, Harvard University.

[30] Ornstein, Stanley (1989), Exclusive Dealing and Antitrust, *The Antitrust Bulletin*, vol 34, 65-98

[31] Pakes, Ariel and David Pollard (1989) Simulation and the Asymptotics of Optimization Estimators, *Econometrica*, 57(5), 1027-1057.

[32] Pakes, Ariel (1997), Lecture Notes for Advanced Applied Econometrics, mimeo, Harvard University.

[33] Prat, Andrea and Aldo Rustichini (2003), Games Played through Agents, *Econometrica*, 71(4), 989-1026.

[34] Rasmusen, Eric, J. Mark Ramseyer and John Wiley Jr. (1991), Naked Exclusion, *American Economic Review*, 81(5), 1137-1145.

[35] Salop, Steven and David Scheffman (1983), Rising Rivals Costs, *American Economic Review*, 73(2), 267-271.

[36] Sass, Tim (2002), The Competitive Effects of Exclusive Dealing: Evidence from the US Beer Industry, Florida State University Working Paper.

[37] Sass, Tim and David Saurman (1996), Efficiency Effects of Exclusive Territories: Evidence from the Indiana Beer Market, *Economic Inquiry*, 34(3), 597-615.

[38] Segal, Ilya and Michael Whinston (2000), Exclusive Contracts and Protection of Investments, *RAND Journal of Economics*, 31(4), 603-633

[39] Sherwood, Sonja (2000), A Full House, *Beverage World*, September 2000, p22.

[40] Stokes, Patrick (1997), Anheuser Busch: Taking the Initiative, *Beverage World*, July 1997, p78.

[41] Tremblay, Victor (1993), Consistency Between the Law and Its Enforcement, *The Antitrust Bulletin*, 38(2), 327-348

[42] Schwartz, M. (1987), The Competitive Effects of Vertical Agreements: Comment, *American Economic Review*, 77, 1063-1068.

[43] Stamborksi, Al (1998), A-B and Miller Settle Distribution Suits, *St Louis Post Dispatch*, May 20, p C1.

[44] Villas-Boas, Sofia (2005a) "Vertical Contracts between Manufacturers and Retailers: An Empirical Analysis", Department of Agricultural & Resource Economics, UCB. CUDARE Working Paper 943.

[45] Villas-Boas, Sofia (2005b) "Wholesale Price Discrimination: Inference and Simulation", Mimeo, Department of Agricultural & Resource Economics, UCB.

[46] Wooldridge, Jeffrey (2002) *Econometric Analysis of Cross Section and Time Series Data*, MIT Press, Cambridge.

Figure 1

Brewer B1 **Brewer B2** **Brewer B3**

Y

X

Distributor D3

Distributor D1 Distributor D2

Z

W

Territories:

D1: YZW   D1 can distribute B1's products to retailers marked with a ■

D2: WXY   D2 can distribute B1's products to retailers marked with a □

    WXYZ   D3 can distribute B2's products to all retailers [□ & ■]

D3: WXYZ   D3 can distribute B3's products to all retailers [□ & ■]
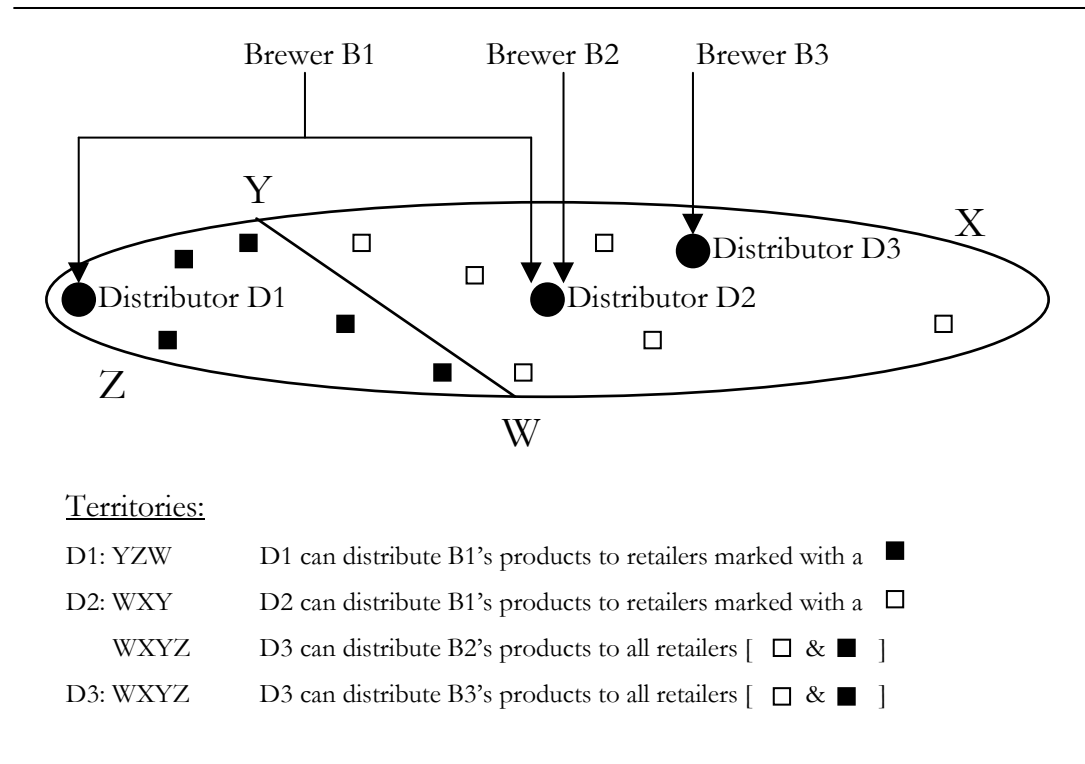
Figure 1: Exclusive Territories and the Three Tier System
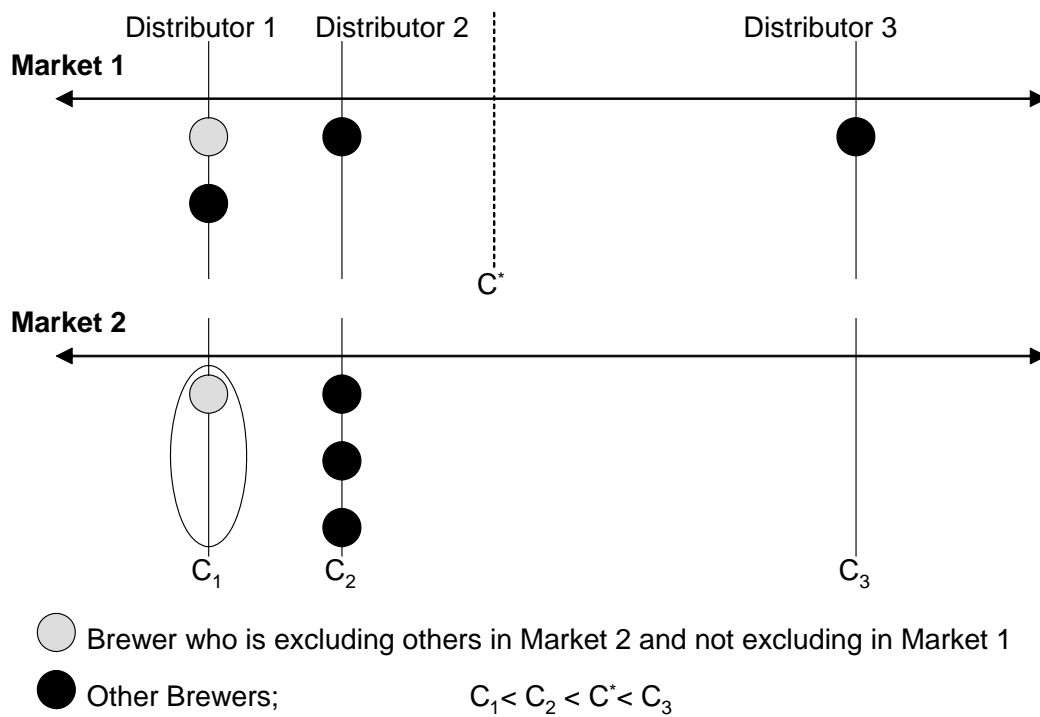
: Figure 2: Equilibrium in networks in two markets

Figure 2

Figure 3

	1	2	3

Brewer sets prices paid
by the distributor and
the retailer

Retailer sets the price
paid by consumers

Consumers make
purchases; distributors
and retailers supply
the quantity demanded
at given prices
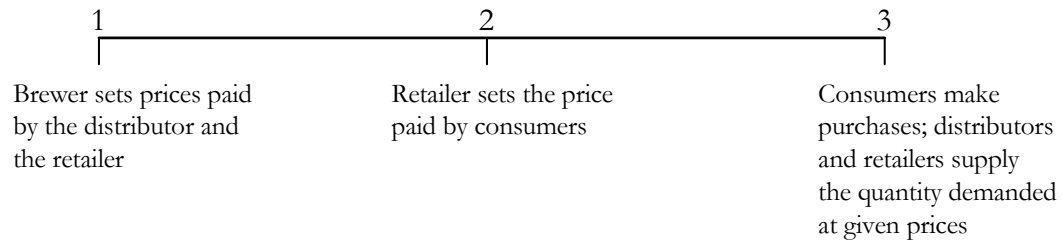
Figure 3: Order of Moves in the Industry Model
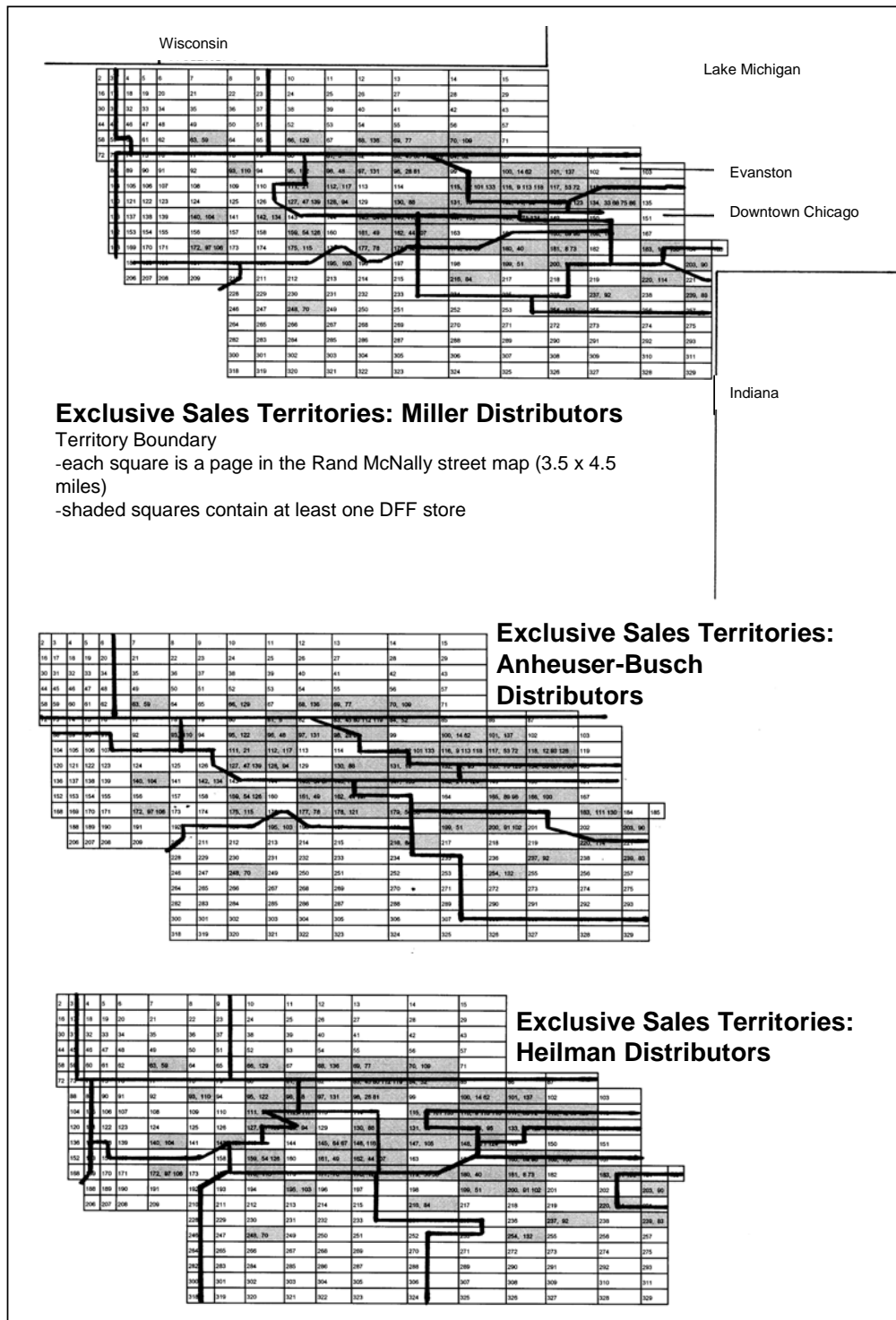
Figure 4: Exclusive Sales Territories for Selected Brewers
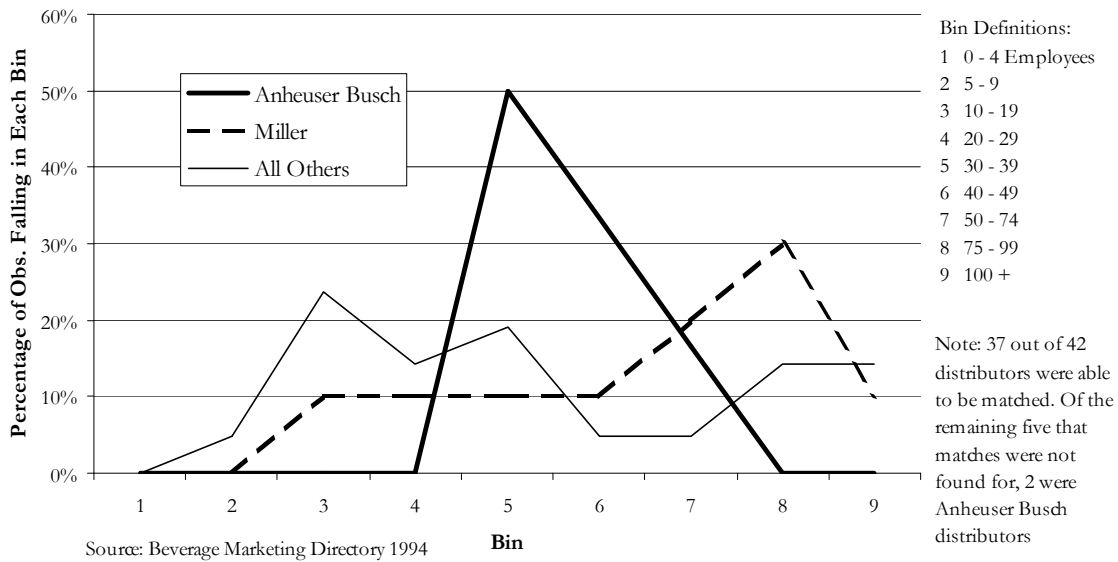
Figure 2.5: Distributors' Employees by Brewer

Bin Definitions:
1 0 - 4 Employees
2 5 - 9
3 10 - 19
4 20 - 29
5 30 - 39
6 40 - 49
7 50 - 74
8 75 - 99
9 100 +

Note: 37 out of 42 distributors were able to be matched. Of the remaining five that matches were not found for, 2 were Anheuser Busch distributors

Source: Beverage Marketing Directory 1994
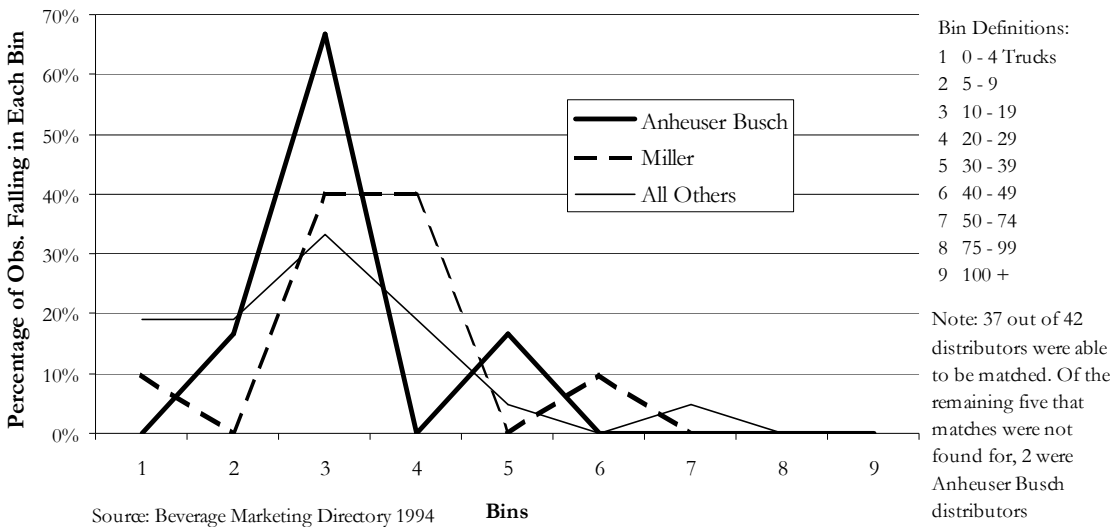
Figure 2.6: Distributors' Trucks by Brewer

Bin Definitions:
1 0 - 4 Trucks
2 5 - 9
3 10 - 19
4 20 - 29
5 30 - 39
6 40 - 49
7 50 - 74
8 75 - 99
9 100 +

Note: 37 out of 42 distributors were able to be matched. Of the remaining five that matches were not found for, 2 were Anheuser Busch distributors

Source: Beverage Marketing Directory 1994

:

35

Table 2: Descriptive Statistics drawn from DFF Data

Number of Obs	138213
Number of Stores	71
Number of Beers	73
Number of Brewers / Importers	12

A. Prices [\$ per 12oz][1]	p^c	p^r		p^c - p^r
mean	0.6001	0.5033		0.0968
stdev	0.1777	0.1345		0.0906
min	0.1983	0.1533		2.4866
max	2.9683	1.1073		-0.3375

B. Other Characteristic	Shares[2]	Serving Size[3]	ABV[4]	Calories[5]
mean	0.054	132.0	4.41	130.8
stdev	0.094	88.8	1.02	25.0
min	0.003	21.0	0.25	58.0
max	3.710	661.0	6.94	200.0

C. Promotional Activity

% of Obs with Promotional Activity Tagged	31.3%
% of Obs with p^c - p^r Negative	17.6%
% of Obs overlapping[6]	15.0%

Notes:
[1] p^c referes to the price to consumers; p^r refers to the price to retailers

[2] expressed as a percentage, taking into account the share of the outside good

[3] expressed in in oz

[4] ABV is Alcohol By Volume

[5] Calories per per 12 oz

[6] This is the % of obs. that have a promotion tagged and p^c - p^r negative

:

Table 1: Brewer - Distributor Matchings

Brewer	**Distributor Allocations (shaded entries are exclusive)**									
Anheuser Busch	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**		
Barton	10	15	18	21	22	23	24	**25**	26	**31**
Boston Beer	9									
Becks	10	11	12	13	14	15	16			
Petes*	9	32								
Sierra Nevada	32									
Sapporo	9									
Miller	21	23	24	**33**	**34**	**35**	**36**	37	38	**39**
Hielman	14	15	17	18	19	20	27	**28**	29	**30**
Pabst	12	13	26	27	29	37	40	41		
Stroh 1	11	12	13	38	40	41	42			
Stroh 2	10	11	12	13	16	40	42			
Coors	11	14	15	17	18	19	20			

* switches form 9 to 32 in May

Table 3: Logit and Random Coefficient (BLP) Demand Models

Specification:	A		B		C		D	
Instruments:	None		Lagged and Lead Mean Prices		Lagged and Lead Mean Prices		Lagged and Lead Mean Prices	
	Coeff.	Std. Err.	**Coeff.**	Std. Err.	**Coeff.**	Std. Err.	**Coeff.**	Std. Err.
Interacted with (1/Income)*2500								
Constant	-	-	-	-	0.029	0.129	-	-
Price (per 12oz)	-	-	-	-	0.293	0.283	-	-
Interacted with draws from N(0,1)								
Constant	-	-	-	-	-	-	-0.177	3.134
Price (per 12oz)	-	-	-	-	-	-	0.063	3.093
Not interacted								
Price (per 12oz)	-4.648	0.106	-6.509	0.112	-6.590	1.341	-5.966	0.744
Promo	0.260	0.013	0.059	0.013	0.104	0.081	0.118	0.062
Serving Size[1] (div by 10)	-0.052	0.002	-0.071	0.002	-0.066	0.010	-0.064	0.009
Serving Size^2 (div by 10000)	0.097	0.005	0.119	0.005	0.011	0.001	0.011	0.001
Holiday[2]	0.364	0.006	0.360	0.006	0.364	0.009	0.361	0.016
Temperature (^{o}C, div by 10)	0.084	0.002	0.083	0.002	0.085	0.003	0.085	0.008
Distributor-Brand Dummies	+	+	+	+	+	+	+	+
r^2	0.436		0.431		0.438		0.433	
SSR	98527		99501		105800		99031	
SER	0.846		0.850		0.877		0.848	
N	138213		138213		138213		138213	

Notes: Note that in specifications B, C and D the r-squared statistics are merely indicative of goodness of fit due to the presence of instrumental variables

[1] Serving size is always in oz.

[2] The Holiday variable is a dummy for weeks in the fortnight leading up to Superbowl Sunday, the 4th of July, Labor Day and Christmas

Table 4: Testing for Promotional Foreclosure

Specification:	A		B		C	
Instruments:	Lagged and Lead Mean Prices		Lagged and Lead Mean Prices		Lagged and Lead Mean Prices	
	Coeff.	Std. Err.	**Coeff.**	Std. Err.	**Coeff.**	Std. Err.
Price (per 12oz)	-6.323	0.115	-7.130	0.246	-6.730	0.193
Promo	0.064	0.014	-0.069	0.030	-0.061	0.023
Serving Size[1] (div by 10)	-0.067	0.002	-0.048	0.005	-0.046	0.004
Serving Size^2 (div by 1000)	0.011	0.001	0.009	0.001	0.009	0.001
Holiday	0.359	0.006	0.337	0.013	0.335	0.010
Temperature	0.008	0.000	0.007	0.000	0.007	0.000
Brand Dummies	+	+	+	+	+	+
"Foreclosed" Stores - Potentially Foreclosed	0.129	0.010	-	-	-	-
"Foreclosed" Stores - AB	0.117	0.008	-	-	-	-
All Exclusive Markets	**-**	-	0.255	0.012	0.300	0.011
r^2	0.377		0.399		0.388	
SSR	108836		22678		35040	
SER	0.888		0.840		0.841	
N	138213		32174		49555	

Table 5: Testing for Cost-Based Foreclosure

Dependant Variable:	Price to Retail - Estimated Markup							
Specification:	**1**		**2**		**3**		**4**	
Estimation Proceedure	OLS, as second step		IV, as second step		OLS, as second step		IV, as second step	
	Coeff.	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**	**Coeff.**	**Std. Err.**
Constant (per unit/SKU)	0.1461	0.0096	0.1451	0.0096	0.0958	0.0078	0.0948	0.0078
Light	0.0183	0.0031	0.0189	0.0031	-	-	-	-
Ice	0.0435	0.0058	0.0435	0.0058	-	-	-	-
NA	-0.4757	0.0785	-0.4764	0.0786	-	-	-	-
Import	0.2757	0.0063	0.2757	0.0063	-	-	-	-
ABV	-0.1785	0.0410	-0.1793	0.0410	-	-	-	-
ABV^2	0.0128	0.0045	0.0128	0.0045	-	-	-	-
Calories	-0.0086	0.0006	-0.0086	0.0006	-	-	-	-
Calories^2 (div by 1000)	0.0425	0.0022	0.0425	0.0022	-	-	-	-
A "Foreclosed" Stores - Potentially Foreclosed	0.0113	0.0029	0.0007	0.0030	0.0021	0.0020	-0.0094	0.0020
B "Foreclosed" Stores - AB	0.0273	0.0035	0.0133	0.0035	0.0090	0.0017	-0.0059	0.0018
Brand Dummies:	-		-		Full Set		Full Set	
Brewer Dummies:	Full Set		Full Set		-		-	
A - B	-0.0160	0.0045	-0.0127	0.0046	-0.0070	0.0026	-0.0035	0.0027
r^2	0.7883		0.7878		0.9480		0.9475	
SSR	43.49		43.59		10.68		10.78	
SER	0.0662		0.0663		0.0329		0.0331	
N	9935		9935		9935		9935	

Notes: The instrument is the proportion of stores that a product is going to that are also served by an exclusive Miller distributor