NBER WORKING PAPER SERIES

ESTIMATING MARKET POWER
EVIDENCE FROM THE US BREWING INDUSTRY

Jan De Loecker
Paul T. Scott

Working Paper 22957
http://www.nber.org/papers/w22957

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2016

We would like to thank John Asker, Cameron Birchall, Tim Bresnahan, Chris Conlon, Aviv Nevo, Mathias Reynaert, Frank Verboven and participants at various seminars for comments and suggestions. This paper uses restricted data that were analyzed at the U.S. Census Bureau Research Data Center in New York City. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. Contact authors at jdeloeck@princeton.edu and ptscott@stern.nyu.edu. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by Jan De Loecker and Paul T. Scott. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Estimating market power
Evidence from the US Brewing Industry
Jan De Loecker and Paul T. Scott
NBER Working Paper No. 22957
December 2016
JEL No. D2,L1,L2,L40

ABSTRACT

While inferring markups from demand data is common practice, estimation relies on difficult-to-test assumptions, including a specific model of how firms compete. Alternatively, markups can be inferred from production data, again relying on a set of difficult-to-test assumptions, but a wholly different set, including the assumption that firms minimize costs using a variable input. Relying on data from the US brewing industry, we directly compare markup estimates from the two approaches. After implementing each approach for a broad set of assumptions and specifications, we find that both approaches provide similar and plausible markup estimates in most cases. The results illustrate how using the two strategies together can allow researchers to evaluate structural models and identify problematic assumptions.

Jan De Loecker
Department of Economics
307 Fisher Hall
Princeton University
Princeton, NJ 08544-1021
and KU Leuven
and also NBER
jdeloeck@princeton.edu

Paul T. Scott
NYU Stern School of Business
Kaufman Management Center, 7-77
New York University
New York, NY 10012
ptscott@gmail.com

1 Introduction

Measuring the extent to which firms exercise market power, raising price above marginal cost of production, is a central topic in industrial organization and antitrust economics. This includes two distinct types of analysis: i) *simulating* the impact of new market structures (e.g. through a merger) and ii) *evaluating* competition policy.

It is in this context that tools were developed to recover estimates of marginal cost of production and the associated markups using data on, essentially, prices and quantities for a particular market of interest.[1] The *demand approach* relies on a fully specified model of consumer choice to derive a demand system that can be estimated using data on prices, market shares and product attributes. Traditionally, this approach relies on market-level data (Berry et al., 1995) while some more recent implementations use consumer-level data (e.g., Berry et al. (2004)). Given the estimated own and cross-price elasticities across the goods considered, markups can be recovered from first-order conditions after specifying a model of competition.

More recently, markups have played a central role in studying the impact of economic integration on welfare, and in questions related to economic growth, resource allocation, international trade, and innovation.[2] However, researchers in these fields typically cannot rely on the detailed data required for the *demand approach*. In this context, the *production approach* is an attractive alternative. The production approach relies on data on input and output levels, and requires that producers choose variable input levels to minimize costs. This approach, however, does not immediately lend itself for counterfactual simulations, as opposed to the *demand approach*.

Each approach relies on a set of largely untested behavioral and practical assumptions, but the sets of assumptions are largely non-overlapping. Thus, each approach provides a potential validation test for the other. This paper is, as far as we are aware, the first to compare the two approaches. Our demand-based analysis follows in the tradition of Bresnahan (1989) and Berry et al. (1995), BLP hereafter, while our production-based analysis follows Hall (1988) and De Loecker and Warzynski (2012), hereafter DLW. De Loecker (2011) outlined the major distinctions between both approaches, but very little is known whether the assumptions in either approach are reasonable approximations in practice.

When implementing the demand approach, researchers must specify the way in which firms in the market compete (e.g., static Nash Bertrand). The notion of market equilibrium, together with estimates of demand elasticities, allows one to infer marginal costs and, therefore, markups. To estimate demand elasticities, researchers typically specify a model of consumer heterogeneity that restricts the pattern of cross-price elasticities to

[1] See Ackerberg et al. (2007) and Verboven (2012) for an excellent overview.

[2] Arkolakis et al. (2015), Peeters (2014), Hsieh and Klenow (2009) and De Loecker et al. (2016).

some extent, and instrumental variables are needed for unbiased estimation of demand elasticities, given the potential correlation between prices and demand shocks.

To implement the production approach, researchers must specify a variable input. Given estimated production function parameters, the conditions for cost-minimization with respect to the variable input allow researchers to compute the marginal cost of production and, therefore, markups. Estimating the production function requires an assumed functional form for the production technology and identifying assumptions to avoid bias from the potential correlation between productivity shocks and variable input levels.

Our comparison of markup estimates focuses on the US Brewing industry, an important and interesting industry in its own right. The US Brewing industry went through a rapid consolidation period and has served as a prototypical example of oligopolistic market structure, in the US and abroad. It has been the focus of several studies on market power, including Asker (2016), Rojas (2008), Hellerstein (2008), Goldberg and Hellerstein (2013), and Miller and Weinberg (2016).

After mapping out a broad space of assumptions and specifications for both the demand and the production approach, we find that the two approaches largely agree on the mean levels of markups in the US brewing industry. The results also suggest that using the two approaches together provides an opportunity to evaluate modeling assumptions. For both approaches, some specifications do deliver implausible estimates, and we argue that these apparent failures can plausibly be attributed to erroneous assumptions.

A related literature attempts to validate the demand approach by performing retrospective assessments of merger simulations. These studies, including Peters (2003), Weinberg and Hosken (2013), Weinberg (2011), and Björnerstedt and Verboven (2013) estimate demand systems using pre-merger data and then predict post-merger prices using an assumed model of competition. An advantage our paper has, relative to these studies, is that our comparison of approaches does not rely on out-of-sample prediction. Proposed mergers generally have purported cost savings, and failing to properly model such cost savings is a potential pitfall of merger simulations. Furthermore, mergers may coincide with other changes in the market such as changes in demand patterns or input prices. Whether such changes influenced the timing of the merger or were mere coincidence, they pose a fundamental difficulty when assessing the performance of merger simulations.

Another related literature attempts to estimate industry conduct by modeling how a known change in marginal costs passes through to prices. These studies include Genesove and Mullin (1998), Wolfram (1999), and Rojas (2008). In this paper, we are not primarily interested in identifying market conduct, but one possible advantage of combining the demand and production approaches would be to identify conduct without relying on a setting in which changes in marginal cost can be directly observed.

The paper is organized as follows. In Section 2 we briefly describe both the demand and production approach to recover markups, where we make a distinction between the theoretical structure and the relevant econometric issues. Section 3 introduces both datasets, and we discuss the specifics of the implementation of the two approaches in Section 4. Section 5 presents the results. In the last section we offer a few concluding remarks and discuss potential avenues for future research.

2 Estimating markups

We distinguish between what we refer to as *behavioral* and *practical* assumptions that go into both approaches. In term of behavioral assumptions, the production approach imposes milder assumptions than the demand approach. The demand approach requires a specific model of how firms compete and a model of consumer behavior. All the production approach requires, in terms of behavioral assumptions, is that firms minimize costs and that there is at least one variable input in production. Cost minimization is implied by profit maximization, regardless of the model of competition.

The implementation of each approach comes, however, with a host of issues surrounding measurement of key variables, functional forms, and identifying assumptions (i.e., instruments for price in the demand approach and for inputs in the production approach). When it comes to these practical assumptions, it is difficult to say one approach requires stronger assumptions than the other, for both approaches require researchers to make a series of practical assumptions, and the assumptions needed for one approach have little to do with the assumptions associated with other.

2.1 Demand approach

The premise of this approach is that a researcher has data on prices, quantities, and other relevant product attributes for a specific product set across multiple markets (over time and/or over different places). We refer the reader to Ackerberg et al. (2007) or Train (2009) for a more detailed overview, and we briefly discuss some of the main assumptions we want to stress.

A consumer l derives utility from product j:

$$u_{lj} = \alpha_l P_j + \beta_l x_j + \xi_j + \varepsilon_{lj}, \tag{1}$$

where x_j is a vector of nonprice characteristics, P_j is the price of product j, ξ_j is a market-level shock, and ε_{lj} is a consumer-level shock. The coefficients (β_l, α_l) can be a function of either observed or unobserved consumer characteristics.

The distribution of consumer preferences can be written as $F(\alpha, \beta|d, \theta)$, where d is a vector of demographic characteristics, and θ parameterizes the distribution. In principle, there could be heterogeneous coefficients for P_j and every characteristic in x_j, and these coefficients could be correlated with each other or not. In practice, researchers parameterize the distribution F with a relatively small number of parameters, allowing for a few heterogeneous coefficients and restricting the correlation patterns across coefficients. The specification of consumer heterogeneity is important in that it dictates what sort of substitution patterns the demand model can accommodate. For instance, a model without heterogeneity and ε_{lj} distributed with a Type-I extreme value distribution is the familiar logit model, which implies highly restrictive substitution patterns.

Ultimately, the goal of demand estimation in this context is to estimate θ. Consumers are fully informed about the choice set $\mathcal{S}$ and buy the good to maximize utility. The error term ε_{lj} is typically required to follow a Type-I extreme value distribution. As shown by Berry (1994) and Berry et al. (1995), we can solve for the market demand shock ξ as a function of the observed market share s_j:

$$\xi(s_j, P_j, x_j, d, \theta). \tag{2}$$

The recovery of ξ is done for each market, indexed by t. After specifying a vector of instruments Z_{jt}^d, θ can be estimated using the following moments:

$$\mathbb{E}(\xi(s_{jt}, P_{jt}, x_{jt}, d_t, \theta) Z_{jt}^d) = 0. \tag{3}$$

The choice of instruments is important both for dealing with the potential endogeneity of price and to identify the heterogeneity in the demand model. Recently, Reynaert and Verboven (2014) and Gandhi and Houde (2015) provide some guidance in constructing instruments that are effective in identifying the distribution of heterogeneity. However, finding instruments which are strong and valid in the presence of price endogeneity remains a challenge without general solutions.

After estimating the distribution of consumer preferences (θ), the final step in recovering markups is to formulate a specific model of competition and conduct. [3] The estimated demand system implies a matrix of own- and cross-price elasticities at the observed prices and quantities. The model of competition implies a system of first-order conditions that relates these elasticities to marginal costs of production.

Note that the original Berry et al. (1995) study and most subsequent papers assume that firms sell directly to consumers. This would not be an appropriate assumption for the beer

[3] In Berry et al. (1995) and many subsequent papers, the model of conduct is specified before demand estimation and the first-order conditions for equilibrium are imposed as additional moments in estimation. Imposing these moments improves the efficiency of the estimator, if the notion of competition is correctly specified. We omit further discussion of these moments for the sake of brevity.

industry, where distributors and retailers stand in between brewers and consumers. Moreover, as discussed below, the production approach delivers markups for beer producers. In Section 4, we extend the basic demand approach to account for this vertical structure.

2.2 Production approach

The *production* approach relies on optimal input demand conditions obtained from cost minimization and the ability to identify the output elasticity of a variable input free of adjustment costs. These conditions relate the output elasticity of an input to the share of that input's expenditure in total sales and the markup. To obtain output elasticities, we need estimates of the production function, and therefore, just as estimating the demand system in the *demand* approach, we need to estimate the production function to recover markups.[4]

We can formally derive markups by considering how a firm can minimize costs while attaining production level Q_{ft}. In the current period, the firm takes as given the levels of dynamic inputs and relevant state variables, so we consider only the first order conditions for variable inputs (X), of which they are potentially many depending on the application.[5] In particular consider the associated Lagrangian function for producer f:

$$\mathcal{L}(X_{ft}, K_{ft}, \lambda_{ft}) = \sum_{X} P_{ft}^{X} X_{ft} + r_{ft} K_{ft} + \lambda_{ft}(Q_{ft} - Q_{ft}, (X_{ft}, K_{ft})), \tag{4}$$

where P_{ft}^{X} and r_{ft} denote the input price for variable inputs and capital (K), respectively.[6]

Taking the first order condition with respect to a variable input (X), rearranging terms and multiplying both sides by $\frac{X_{ft}}{Q_{fjt}}$, and let the markup be defined as $\mu_{ft} \equiv \frac{P_{ft}}{\lambda_{ft}}$, we obtain the following expression:

$$\frac{\partial Q_{ft}(X_{ft}, K_{ft})}{\partial X_{ft}} \frac{X_{ft}}{Q_{ft}} = \mu_{ft} \frac{P_{ft}^{X} X_{ft}}{P_{ft} Q_{ft}}. \tag{5}$$

This expression holds for *any* form of conduct among firms and static demand specification.[7] In order to calculate producer-level markups using production data, we require

[4] We refer the reader to De Loecker and Warzynski (2012) for a detailed exposition.

[5] They typically include intermediate inputs, electricity, and labor. If the researcher allows for multiple variable inputs, overidentifying restrictions can be imposed to estimate markups. Similarly, there can be multiple dynamic inputs in production, and here we use the notation K, since capital is a natural candidate.

[6] We use the subscript f to denote the producer. In our application we observe plants, and recover plant-level markups, which in the case of multi-product plants are a simple weighted average of product-specific markups. We can further aggregate the plant-level markups up to firm level, if we were interested in comparing firm-level markups directly. We are mostly concerned with comparing the mean markup across products, and so in the main analysis the plant is the relevant unit of observation. In addition, the multi-product dimension within plants is significantly reduced.

[7] In deriving this expression we have abstracted away from market power in input market or non-

estimates of the output elasticities of one variable input of production and data on that input's revenue share; where the latter is readily available in most production data sets.

In addition to specifying the production function, which can be kept fairly general of the form (in logs):

$$q_{ft} = f(\boldsymbol{x}_{ft}; \boldsymbol{\beta}) + \omega_{ft}, \tag{6}$$

we require an explicit form for the productivity process. This process is modeled as a first-order Markov process:

$$\omega_{ft} = g(\omega_{ft-1}) + \nu_{ft}. \tag{7}$$

The latter generates the moment conditions used to estimate the production function parameters:

$$\mathbb{E}(\nu_{ft}(\boldsymbol{\beta}) Z^q_{ft}) = 0, \tag{8}$$

where Z^q_{ft} are the instruments used for the various inputs of production.[8]

Therefore, with data on outputs and inputs for a panel of producers in a given industry we can calculate markups for a firm f as follows:

$$\mu_{ft} = \theta^X_{ft} \frac{R_{ft}}{C^X_{ft}}, \tag{9}$$

where $\theta^X_{ft} = \frac{\partial Q_{ft}}{\partial X_{ft}} \frac{X_{ft}}{Q_{ft}} = \theta(\boldsymbol{\beta}, \boldsymbol{X}, \boldsymbol{K})$, $R_{ft} = P_{ft} Q_{ft}$ and $C^X_{ft} = P^X_{ft} X_{ft}$.

The data requirements for the two approaches approach are thus radically different. The production approach relies on data for (upstream) producers; the demand approach is typically implemented with data on retail (downstream) purchases. The production approach relies on a panel of producers; the demand approach can be implemented using only a cross section (although many recent implementations rely on market-level panels).

We illustrate the approach with an example of the US automobile industry, and demonstrate the simplicity of this approach by calculating a simple average of the markup (for the *upstream* producers). Berry et al. (1996) report summary statistics for US car plants using US census data. In Table 2 the authors report three statistics for the year 1982 that are sufficient to calculate an average markup: material expenditure per unit of output ($7,493), unit price ($10,672) and assuming that the output elasticity of materials is 0.85[9] :

$$\mu = 0.85 \times \frac{10,672Q}{7,493Q} = 1.21, \tag{10}$$

linear input prices. However, we can adjust the expression accordingly. The expression then becomes $\mu_f = \theta^X \frac{PQ}{P^X X}(1 + \eta^{-1})$, where θ^X is the output elasticity of input X, as before, and η is the elasticity of the input price P^X with respect to input demand X.

[8] We discuss our choices for the instruments Z^q_{ft} and specification of the Markov process g in section 4.

[9] This is obtained by directly equalizing it to the cost share of materials, and it is not obtained by estimating the production function, as in the production approach we apply to the beer data. This example is merely to illustrate the approach and the essential data requirements.

which is roughly in line with the reported markups in Table 1, which are taken directly from BLP, albeit the latter is for the year 1990. Interestingly enough, BLP report an average markup across their entire sample of 1.31, which is in line with the production approach.[10] This calculation merely demonstrates the minimal requirements. For the sake of illustration, we have assumed that the output elasticity of materials is equal to its cost share. We cannot take for granted the equivalence of cost shares and output elasticities if we have both flexible and quasi-fixed inputs, or if we do not want to assume Cobb-Douglas production functions, so we generally require output elasticity estimates from a structural model of production.

3 Data

In this section, we describe the standard demand and production data sets we use for the US Brewing industry. The data sets overlap in 1992, the year we focus on for the comparison of markups. Further details about both data sets are available in the Appendix.

3.1 Demand data

On the demand side we rely on data from a large Chicago-based supermarket chain *Dominick's Finer Foods* .[11] This data has been used extensively for demand estimation, and in particular the recent work of Asker (2016) and Goldberg and Hellerstein (2013) use these data to study the beer market; we refer readers to their data descriptions for further details. Table 2 presents summary statistics.

We consider the period 1991-1995 and have a total of 23 firms covering over 100 distinct products. We measure quantities in six-pack-equivalent units and prices in dollars per six pack. It is important to note that we directly observe both the retail and wholesale price by product; the wholesale prices play an important role in accounting for the industry's vertical structure.

3.2 Production data

The data on production comes from the U.S. Census Bureau's Research Data Center Program. We use data on beer brewers, NAICS code 312120 , as part of the Census of Manufacturers (henceforth, CMF), and the Annual Survey of Manufacturers (henceforth, ASM) from 1972 to 1997.

[10] We obtain the average markup by converting BLP's reported Lerner index of 0.239 and the average reported price and price-cost difference. We do not expect the demand- and production-based markups to be identical, if only for the reason that foreign cars are included in the demand approach as well. Dealer margins could explain part of the difference as well.

[11] This data is provided by *Kilts Center for Marketing* at Booth School of Business.

This generates a panel of plants active during the period 1972-2012, and we observe standard production data covering: output, input use (labor, capital, intermediate inputs and energy), investment, and indicators of plant survival (entry/exit). Figure 1 plots the real value of aggregate shipments and intermediate input expenditures. We also include the number of active breweries on the right axis.[12]

4 Implementation

In this section, we describe the implementation of the demand and production approaches for the US Brewing industry.

4.1 Demand approach: model of consumer heterogeneity

We consider a logit specification and two mixed logit specifications in which consumer i has the following utility function for product j in period t:

$$u_{ijt} = \beta_{pi} P_{ijt} + \beta_{li} \text{light}_j + \beta_x x_j + \xi_{jt} + \varepsilon_{ijt}, \tag{11}$$

where

$$\begin{aligned} \beta_{pi} &= -\exp\left(\theta_p + \sigma_p * v_{i,1}\right) \\ \beta_{li} &= \beta_l + \sigma_l * v_{i,2}, \end{aligned}$$

and light_j is a dummy indicating a light beer, and x_j is a vector of other characteristics for product j. We include product-level dummy variables, which absorb β_l and $\beta_x x_i$. Furthermore, we allow for separate linear time trends in the utility function for each brand.

The main difference in the two mixed logit specifications is whether the heterogeneous characteristics v_1 and v_2 are correlated or not. In our Income-Mixed Logit specification, $v_{1i} = v_{2i} = inc_i$, and we simulate a distribution of inc_i that approximates the distribution of log income in the region. In our Random-Mixed Logit specification, v_{i1} and v_{i2} are independent and have a simulated distribution that approximates the standard normal (using Gauss-Hermite quadrature).

4.2 Demand approach: instruments

We consider five different sets of instrumental variables.

[12] It appears that Census in fact substantially overstates the actual number of domestic brewing companies in existence since it includes entities having establishments primarily engaged in manufacturing any kind of malt beverages, even those that are kept in minimal operation or that are experimental in nature. The Brewing Industry Survey records the active number of breweries – see `http://www.brewersassociation.org/statistics/number-of-breweries/` and it is this series that we plot in Figure 1. These statistics (and our production data) leave out establishments with a primary classification as something other than a brewery, such as brewpubs.

(1) Exogenous prices In specifications with exogenous prices, the instruments are price, price squared, price interacted with the light beer dummy, and price interacted with a dummy for the "Big Three" brewing companies (Anheuser-Busch, Miller, and Coors).

(2) Cost-based instruments The cost-based instruments are identical to the instruments with exogenous prices, but with the retailer's measure of the replacement cost replacing prices.

(3) Exogenous product sets The BLP instruments are the number of products in the choice set, the number of light beers, each of these two counts interacted with a light beer dummy, and each of the two counts interacted with the "Big Three" dummy.

(4) Promotion-based instruments The sales-based instruments are a dummy variable for being on sale (given by the retailer), the number of products which are on sale or promotion, the number of light beers on sale, the number of products on sale indicated with a light beer dummy, and the on sale dummy interacted with the light beer dummy.

(5) Optimal instruments For the exogenous price instrument set, we consider an approximation to optimal instruments following Reynaert and Verboven (2014). See Appendix A for more details.

4.3 Demand approach: model of vertical structure

We assume that brewers and distributors function as integrated firms, avoiding any double marginalization issues.[13] Furthermore, we assume that retail markups are exogenous – i.e., the difference between the retail price and wholesale price does not respond to changes in the wholesale prices. In particular, this model of price setting is similar to the following environment: A brewer's f's profits are given by

$$\sum_{j \in J_f} s_j \left(P^r\right) \left(P_j^w - c_j - \tau\right), \tag{12}$$

where J_f is the set of products produced by firm f and τ represents taxes levied on distributors. In equation (12), demand is a function of retail prices, but the brewer-distributor margins involve wholesale prices. Assuming that retail markups are exogenous, the brewer's profit maximization problem can be written quite simply:

$$\max_{P_f^w} \sum_{j \in J_f} s_j \left(P_f^w + \mu_f^r, P_{-f}^w + \mu_{-f}^r\right) \left(P_j^w - c_j\right), \tag{13}$$

[13] See Asker (2016) for more discussion.

where P_f^w is the vector of wholesale prices for products produced by f, P_{-f}^w is the vector of wholesale prices for other products, and μ^r denotes a vector of retail markups.

The Dominick's data includes a measure of the retailer's cost of replacing the product. We assume that the wholesale price is equal to this cost, implying that retail markups are simply the difference between the product's price and the reported replacement cost. The assumption of exogenous retail markups means that the μ^r terms in a brewer's profit function are simply exogenous constants; they may vary by product-market-year, but they do not respond to changes in upstream pricing.

The first-order conditions for equation (13) can be expressed as follows:

$$c_f = \Delta_f^{-1} s_f + P_f^w - \tau_f, \tag{14}$$

where c_f, s_f, P_f^w, and τ_f are all $|J_f| \times 1$ vectors with rows corresponding to firm f's products. Δ_f is a $|J_f| \times |J_f|$ matrix which is the Jacobian of s_f with respect to P_f^w (or P_f^r).

We apply this calculation separately to each market-week, effectively assuming that brewer-distributors are able to adjust their prices frequently and set different prices in each market. While it may be realistic that distributors adjust prices frequently and tailor prices to specific markets, brewers do not. In any case, we must consider a model with double marginalization before we can make a distinction between the two (more on this below).

A potential problem with this model of integrated brewer-distributor pricing is that it ignores distribution costs. With our production-based approach, we recover markups for brewers that do not include distributor costs. A simple way of dealing with this is to include distribution costs in the tax τ_f paid by distributors. In our baseline specification, we set $\tau_f = 0.44$, reflecting the federal and state taxes levied on beer (per six pack).[14] In a robustness check, we set $\tau_f = .69$, matching the average cost of taxes and shipping reported by Consumer Reports (1996). Increasing the value of τ_f in this way mechanically decreases the inferred price received and marginal cost (each by the same amount) for the brewer. As expected, this causes markups to increase, as shown in Table 5.[15]

4.4 Production approach: technology

We follow DLW closely and start from a general production function where production takes labor (l), capital (k), and intermediate inputs (m) to produce output (y), while relying on productivity ω_{ft}.

[14] The federal tax was \$18/barrel and the Illinois state tax was \$0.185/gallon.

[15] This correction for shipping costs is probably excessive; based on rough estimates of freight costs, twenty-five cents per six pack is an unrealistically high cost of shipping beer between large Midwestern breweries and the Chicago area in our demand data.

We consider an exhaustive list of empirical models to compute markups by varying the assumptions on i) technology, ii) which inputs to production are not subject to adjustment cost and are freely set by the plant's manager and iii) how productivity evolves. The technologies we consider are gross output, restricted profit value added, and structural value added production functions. For each technology type we also consider the standard Cobb-Douglas and Translog functional forms.

With variables measured in logs, a gross output production function with a Cobb-Douglas specification has a linear functional form:

$$y_{ft} = \beta_0 + \beta_m m_{ft} + \beta_l l_{ft} + \beta_k k_{ft} + \epsilon_{ft} + \omega_{ft}, \tag{15}$$

where ϵ_{ft} can be considered measurement error and ω_{ft} can be considered productivity (more on these error terms below). Translog specifications involve second-order terms and interactions of the inputs on the right-hand side of equation (15).

Both value added technologies drop materials from the right hand side of equation (15). In the restricted profit value added specification, the dependent variable is $\log(R_{ft} - M_{ft})$, where R_{ft} is sales and M_{ft} is materials expenditures. With restricted profit value added technology, the elasticity should be corrected as follows:

$$\theta_{ft}^X = \widetilde{\theta}_{ft}^X \frac{(R_{ft} - M_{ft})}{R_{ft}}, \tag{16}$$

where $\widetilde{\theta}_{ft}^X$ is the elasticity of value added production with respect to input X, and θ_{ft}^X is the overall output elasticity that we use for our markup calculations.

In structural value added specifications, the dependent variable is the same as in the gross output case, but materials are simply omitted from the right-hand side. This is justified by a production function with the following form:

$$Y_{ft} = \min\left[\gamma_m M_{ft}, F(L_{ft}, K_{ft}) \exp(\omega_{ft})\right] \exp(\epsilon_{ft}). \tag{17}$$

Here, labor and capital may be substitutable, but materials are a perfect complement to the combination of labor and capital. Assuming that materials and labor are variable inputs, a cost-minimizing firm will set

$$Y_{ft} = \gamma_m M_{ft} \exp(\epsilon_{ft}) = F(L_{ft}, K_{ft}) \exp(\omega_{ft} + \epsilon_{ft}), \tag{18}$$

justifying the regression of Y_{ft} on labor and capital while ignoring materials. A caveat is that, in theory, equation (18) may not be satisfied in certain situations. If labor and capital are quasi-fixed, and materials is a flexible input, then when output prices are sufficiently low relative to the price of materials, it will be better to set $M = 0$ and not produce at all. However, given that our data only includes actively producing breweries, we assume that equation (18) always holds.

While structural value added production functions allows us to ignore materials when regressing production on inputs, materials cannot be ignored when calculating markups. In our derivation of production-based markups above, we considered production functions that were differentiable with respect to a single variable input. Here, we must consider increasing both materials and labor together to have a well-defined marginal product. Marginal costs, by producer f and time t, can be expressed as follows:

$$\lambda_{Qft} = \lambda_{Fft} + P_{Mt}/\gamma_M, \tag{19}$$

where λ_F is the marginal cost of an additional unit of F, and λ_Q is the marginal cost of an additional unit of output. Accordingly, markups can be computed as follows:

$$\mu_{Qft} = \frac{1}{\mu_{Fft}^{-1} + \alpha_{Mft}}, \tag{20}$$

where μ_{Fft}^{-1} is the markup formula if we treat labor as the variable input and use the original markup formula, and α_{Mft} is the revenue share of materials expenditure.

4.5 Production approach: moments and estimation

Constructing moments for estimation begins with recovering productivity. Our strategy follows Ackerberg et al.'s (2015) control function approach, which builds on Olley and Pakes (1996) and Levinsohn and Petrin (2003). The control function for productivity includes capital, labor, materials, firm-level wages, and interactions. A first-stage control function regression allows us to purge measurement error from the production function.[16] That is, the production function can be decomposed as $y_{ft} = y_{ft}^* + \epsilon_{ft}$, and after the control function regression, we have measures of y_{ft}^* and ϵ_{ft}.[17]

Next, given a vector of production function parameters β, we can recover productivity $\omega_{ft}(\beta)$:

$$\omega_{ft}(\beta) = y_{ft}^* - x_{ft}'\beta. \tag{22}$$

Given estimated productivities, we can estimate a first-order Markov process for ω_{ft}. Given the process governing ω_{ft}'s evolution, we can compute innovations in the productivity

[16] The validity of the moments on the variable inputs has been a topic of interest in the recent literature on production function identification. By including wages – a serially correlated input price that varies across producers – in the control function, we can avoid Gandhi et al.'s (2013) non-identification result.

[17] Throughout all our specifications we allow for unanticipated shocks to output and measurement error, such that $y_{ft} = y_{ft}^* + \epsilon_{ft}$. This error captures the differences between expected output, given the information set of the producer, and the actual realization. As proposed by DLW, we correct the markup formula by eliminating measurement error as follows:

$$\mu_{ft} = \theta_{ft}^X \frac{R_{ft}(\exp(\epsilon_{ft}))^{-1}}{C_{ft}^X} \tag{21}$$

process:

$$\nu_{ft}(\beta) = \omega_{ft}(\beta) - \mathbb{E}(\omega_{ft}(\beta)|\omega_{ft-1}(\beta)), \tag{23}$$

It is these innovations $\nu_{ft}(\beta)$ that we use to form moments and compute a method of moments estimate of the production function parameters β.

The moments are products of the innovations and inputs. Roughly speaking, for variable inputs x we assume $\mathbb{E}(\nu_{ft}(\beta)x_{ft-1}) = 0$, while for fixed inputs k we assume $\mathbb{E}(\nu_{ft}(\beta)k_{ft}) = 0$. Contemporaneous capital is always included in the moments. As described in the presentation of the results, we either include contemporaneous or lagged labor, depending on the specification. Lagged materials expenditure is included in the moments for gross output production specifications; otherwise, materials expenditure is not included in the moments

5 Results

In order to put our markups estimates from the two approaches into context we present a reported breakdown of a nationally representative 6-pack of beer (Consumer Reports, 1996). Table 3 decomposes the final final price for a 6-pack into the various cost of production categories (ingredients, packaging, labor in production and advertising/management), distribution and retail. We use this case study to sketch a broad picture of the margins along the vertical chain, and in particular to get a sense of the markup at the production level.

The average price in our data is \$3.25 in 1995 and is slightly below the price of \$4.01 (for large brewers) used in this break-down. Treating advertising and management as non-variable expenditures and other costs as fully variable and proportional to output, the report implies a producer-level markup of 1.44 for mass-produced beers. If we also treat labor as a non-variable expenditure, the markup changes to 2.27. We view these as rough upper and lower bounds on the plausible range for the average markup in the industry.

5.1 Demand Approach

For each approach, we compute the share-weighted mean markup for US producers in the year 1992. Note that large brewers had a large majority of the market in 1992, so our mean markup estimates can be understood as mean markup estimates for mass-produced beer.

For our baseline model of competition and vertical structure in the industry (fixed retail markups and Nash-Bertrand competition among integrated brewer-distributors), Table 4 presents mean markup estimates, mean own-price elasticities, and the proportion of products with well-defined markups. Note that for a single-product firm, the markup is well-defined if and only if the own-price elasticity is less than -1.

We consider both the logit and mixed logit models of demand, and we consider various instruments used throughout the literature. Perhaps somewhat surprisingly, we obtain a mean markup estimate in the 1.6-1.7 range in several specifications, very much consistent with the range predicted by the Consumer Reports study.

In Table 5 we vary the model of competition to evaluate the sensitivity of the average markup across models of conduct. This exercise is in the spirit of Bresnahan's (1987) seminal work on market power in the US automobile industry. Interestingly, monopoly conduct only implies marginally higher markup estimates for the logit specifications and some of the mixed logit specifications with little heterogeneity.[18] This is a result of these specifications having relatively small cross-price elasticities, so taking into account the cross-price effects across all products does not make a huge difference. However, for mixed logit specifications with substantial heterogeneity, particularly those with substantial heterogeneity in the price coefficients, cross-price elasticities are relatively higher, and monopoly conduct implies much larger markups. See Table 10 for a comparison of some cross-price elasticities for a few selected products in two different specifications.

The only instrument set that systematically delivers implausible markup estimates are the instruments based on sales/promotions, which yield relatively elastic demand and implausibly low markup estimates (around 1.35), if we take seriously the lower bound of 1.44 suggested by Consumer Reports. This apparent failure of these instruments might be because the price decreases associated with sales are also accompanied by other promotional activities such as placing the product in a prominent display. This apparent failure could be rationalized with a similar story to Hendel and Nevo's (2006) account of laundry detergent demand. If brewers set prices with respect to long-run consumption elasticities, but short-run purchase elasticities associated with discounts are higher than consumption elasticities because consumers have some ability to store beer, then these instruments would lead to downward bias in markup estimates. However, this story should also bear on instrument sets (1) and (5), which treat prices as exogenous, for the discounts show up in the price variation. One possibility is that this source of bias is more apparent when using the sales/promotion-based instruments. Perhaps more plausibly, the sales/promotion instruments could be associated with another source of bias due to product placement. That is, a product's being "on sale" may involve both discounts and placement of the product in a more prominent position within the store, an thus the quantities sold during these sales exaggerate the effect of the discount.

[18] Table 9 presents the demand parameter estimates. Some of the mixed logit specifications converged to zero variances for the heterogeneous coefficients.

5.2 Production approach

Table 6 presents the markup estimates across different production function specifications, labor or materials (when appropriate) as the variable input, and instruments for the innovations in productivity.

For some of the gross output specifications, we obtain very large standard errors on markups, reflecting the difficulty of identifying the production function parameters for variable inputs, especially in the translog case. Within the gross-output estimates, only the case of Cobb-Douglas technology with materials as the variable input has reasonably precise standard errors, and those estimates (1.65) are very much consistent with the demand estimates and fall within the range predicted by the Consumer Reports study.

The results from the value added specifications are more striking. The structural value added point estimates (1.59-1.69) are very close to the point estimates on the demand side, and the standard errors are reasonably small (.08). However, the restricted profit value added specifications deliver point estimates which are implausibly high. In the Cobb-Douglas restricted profit value added specifications, the bottom end of a 90% confidence interval would be near the top end of the plausible range predicted by Consumer Reports.

These apparently problematic results from the restricted profit value added specifications should not be surprising. Gandhi et al. (2013) and Ackerberg et al. (2015) caution against using restricted profit value added specifications. While these specifications can be defended as a local approximation, the approximation may break down and deliver biased elasticity estimates when estimating based on cross-section variation in which different plants have very different input levels. Consistent with these claims, Gandhi et al. (2013) argue that restricted profit value added specifications deliver implausibly high output elasticities in their application using Columbian and Chilean data. Ackerberg et al. (2015) recommend using only the structural version of value added production functions.

Looking at the production function coefficients in Table 8, one notable observation is that the coefficients on labor are estimated with relatively small standard errors for value added specifications, but labor and materials coefficients both have large standard errors in the gross output specifications, reflecting high correlation between labor and materials inputs.

We surmise that, at least in the context of modeling breweries, there is much to recommend structural value added specifications. First, the apparent lack of independent identifying variation in materials and labor inputs makes it difficult to estimate a gross output production function with both materials and capital on the right hand side. This lack of independent identifying variation is also entirely consistent with a structural value added production function (equation (17)). Even assuming that all breweries share the same proportionality γ_M between materials and output, there would still be some inde-

pendent variation in labor and materials due to productivity variation. However, this independent variation will not be identifying variation – given the placement of the error terms in the functional form, materials should be a strictly better predictor of output than labor.[19]

Second, materials plausibly do come in a fixed proportion with beer output, as there is little substitutability for beer ingredients. It's worth noting that our estimation strategy does not impose that the proportionality γ_M between materials and output need not be common across breweries, or even within a brewery over time, meaning that our approach is compatible with different breweries buying inputs with different prices. This proportionality only enters into our framework when computing markups through the α_{Mit} term – the ratio of materials expenditure to sales – and this is compatible with the proportionality γ_{Mit} varying over breweries and time.[20]

Table 7 presents results for different identifying assumptions, focusing on the case of structural value added production functions. At least for Cobb-Douglas specifications, estimated mean markups are completely robust; it makes almost no difference whether we use contemporaneous labor l_t or lagged labor l_{t-1} as an instrumental variable, or whether the Markov process for productivity evolution is assumed to be linear or quadratic in the lagged value of productivity. When lagged labor is used as an instrument in translog specifications, the production coefficients are poorly identified, and consequently standard errors for markups are large, too, but the point estimates remain similar. All together, the estimated mean markups are remarkably stable across identifying assumptions, and we find that varying the identifying assumptions never makes a statistically significant difference in the estimated mean markups.

5.3 Comparisons

Both approaches are thus in broad agreement on the overall level of market power in this industry, with markup estimates falling within the 1.5-1.9 range in most specifications with reasonably precise standard errors.[21]

Table 11 presents a more quantitative comparison of the two approaches. For each of the demand specifications, we compare the mean markup estimates assuming Nash-

[19] We could make this point stronger by placing the productivity term in equation (17) outside the minimum function. In this case, labor and materials will be perfectly collinear, so there is obviously no hope of identifying coefficients on both variables. In the data, labor and materials are not perfectly collinear, so we prefer the form in equation (17), which rationalizes both the *independent variation* in the two variables and lack of *independent identifying variation*.

[20] Heterogeneous input requirements also allow us to capture the fact that some brewers purchase malt directly while others have malting equipment and purchase grains. This is the only significant possibility for substitution between materials expenditure and the other inputs that we are aware of.

[21] The markups obtained from the production approach are a national average, as we rely on the US census beer producers, whereas the demand approach relies on a more specific market – i.e., the Chicago area.

Bertrand pricing in wholesale prices to the estimates from a structural value added production specification.[22] We present two pieces of information for each comparison. First, a z-test comparing the two means at a 5% (two-sided) significance level.[23] However, failure to reject the null hypothesis of similar means is not highly informative in cases where we have poor identification on the demand side. Therefore, we also compute the Battacharya coefficient, a measure of how much two distributions overlap, for the two asymptotic distributions. (Zero means the two distributions do not overlap; unity means that the two distributions are identical.) When the two estimates are compatible and the degree of overlap of the two distributions is high, the agreement between the two approaches is meaningful. Table 11 illustrates that we have meaningful agreement between the two approaches in many cases.

On the other hand, we see that there are some cases of strong disagreement, particularly when using instruments based on sales/promotions. This points to the possibility of using the two approaches together to select among models and assumptions. Indeed, the demand approach is unavoidable in certain contexts, for the production approach can not be used to perform prospective merger simulations. However, the production approach could be used to evaluate the demand approach and ensure that the demand model delivers plausible levels of market power.

6 Conclusion

While inferring markups from demand data is common practice, the estimation relies on specific assumptions on consumer choice and how firms compete in a market. Alternatively, markups can be inferred from production data, relying on the assumption that firms minimize costs with respect to a specified variable input. In this paper, we directly compare markup estimates from both strategies for the case of the US beer industry. We rely on standard data sources for both approaches – i.e., market-level demand data and industry-wide production data for the US beer industry.

For most specifications that deliver reasonably precise estimates, we obtain mean markup estimates falling within the 1.5-1.9 range for both the demand and production approaches. The results indicate fairly broad agreement between the two approaches while highlighting potential pitfalls. We argue that the only well-identified results from either approach that are significantly at odds with the other approach can plausibly be attributed to erroneous assumptions or misspecification. In particular, we find that demand-based markup

[22] The specification selected for the comparisons has a Cobb-Douglas functional form in capital and labor and uses lagged labor as an instrument, but as Table 7 shows, the markup estimates from different version of structural value added specifications are quite robust

[23] Note that the estimated asymptotic distribution of the demand-based markups is often highly non-normal, so this test should be seen as illustrative

estimates are sensitive to the choice of instruments, and that instruments on discounts and promotional activity lead to implausibly low markups. On the production side, restricted profit value added specifications lead to implausibly high markups, consistent with Gandhi et al.'s (2013) claims about bias in these specifications.

Structural approaches to studying market power have recently been criticized by Angrist and Pischke (2010), in essence, for relying on a host of hard-to-test assumptions. The results from this paper show that two very different approaches give compatible results on the degree of market power in the US beer industry. At a very minimum, our results support the claim that these structural approaches give plausible estimates when implemented well, and they suggest that combining the demand and production approaches may assist researchers in selecting assumptions and evaluating structural models.

References

Ackerberg, D., L. Benkard, S. Berry, and A. Pakes (2007). Econometric tools for analyzing market outcomes. *The Handbook of Econometrics*, Chapter 63.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

Angrist, J. and J.-S. Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspective 24*(2), 3–30.

Arkolakis, C., A. Costinot, D. Donaldson, and A. Rodriguez-Clare (2015). The elusive procompetitive effects of trade. *MIT mimeo*.

Asker, J. (2016). Diagnosing foreclosure due to exclusive dealing. *Journal of Industrial Economics 3*(64), 375–410.

Berry, S., S. Kortum, and A. Pakes (1996). Environmental change and hedonic cost functions for automobiles. *Proceeding of the National Academy of Sciences 93*(23), 12731–12738.

Berry, S., J. Levinsohn, and A. Pakes (1995, July). Automobile prices in market equilibrium. *Econometrica 63*(4), 841–90.

Berry, S., J. Levinsohn, and A. Pakes (2004). Differentiated products demand systems from a combination of micro and macro data: The new car market. *The Journal of Political Economy 112*(1), 68–105.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics 25*(2), 242–262.

Björnerstedt, J. and F. Verboven (2013). Does merger simulation work? evidence from the swedish analgesics market. *American Economic Journal: Applied Economics*.

Bresnahan, T. F. (1987). Competition and collusion in the american automobile market: The 1955 price war. *Journal of Industrial Economics XXXV*(4), 457–482.

Bresnahan, T. F. (1989). Empirical studies of industries with market power. *Handbook of industrial organization 2*, 1011–1057.

Chamberlain, G. (1987). Asymptotic efficiency in estimation with conditional moment restrictions. *Journal of Econometrics 34*(3), 305–334.

Consumer Reports (1996, June). Can you judge a beer by its label? *Consumer Reports*, 10–12.

De Loecker, J. (2011). Recovering markups from production data. *International Journal of Industrial Organization 29*, 350–355.

De Loecker, J., P. Goldberg, A. Khandelwal, and N. Pavcnik (2016). Prices, markups and trade reform. *Econometrica 84*(2), 445–510.

De Loecker, J. and F. Warzynski (2012). Markups and firm-level export status. *American Economic Review 102*(6), 2437–2471.

Gandhi, A. and J.-F. Houde (2015). Measuring substitution patterns in differentiated products industries the missing instruments. Technical report, Mimeo, University of Wisconsin-Madison and University of Pennsylvania.

Gandhi, A., S. Navarro, and D. Rivers (2013). On the identification of production functions: How heterogeneous is productivity. *mimeo University of Wisconsin-Madison*.

Genesove, D. and W. P. Mullin (1998). Testing static oligopoly models: Conduct and cost in the sugar industry, 1890-1914. *The RAND Journal of Economics 29*(2), 355–377.

Goldberg, P. K. and R. Hellerstein (2013). A structural approach to identifying the sources of local-currency price stability. *Review of Economic Studies 80*(1), 175–210.

Hall, R. E. (1988). The relation between price and marginal cost in us industry. *Journal of Political Economy 96*(5), 921–947.

Hellerstein, R. (2008). Who bears the cost of a change in the exchange rate? pass-through accounting for the case of beer. *Journal of International Economics 76*(1), 14 – 32.

Hendel, I. and A. Nevo (2006, November). Measuring the implications of sales and consumer inventory behavior. *Econometrica 74*(6), 1637–1673.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing TFP in china and india. *The Quarterly Journal of Economics 124*(4), 1403–1448.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *Review of Economic Studies 70*(2), 317–341.

Miller, N. and M. Weinberg (2016). The market power effects of a merger: Evidence from the us brewing industry.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications industry. *Econometrica 64*(6), 1263–1297.

Peeters, M. (2014). Heterogeneous mark-ups, growth and endogenous misallocation. *mimeo Yale University*.

Peters, C. T. (2003). Evaluating the performance of merger simulation: Evidence from the us airline industry. *DOJ Antitrust Division Economic Analysis Group Discussion Paper* (03-1).

Reynaert, M. and F. Verboven (2014). Improving the performance of random coefficients demand models: the role of optimal instruments. *Journal of Econometrics 179*(1), 83–98.

Rojas, C. (2008). Price competition in us brewing. *The Journal of Industrial Economics 56*(1), 1–31.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge UP.

Verboven, F. (2012). Empirical industrial organization and competition policy. *mimeo KU Leuven*.

Weinberg, M. C. (2011). More evidence on the performance of merger simulations. *The American Economic Review 101*(3), 51–55.

Weinberg, M. C. and D. Hosken (2013). Evidence on the accuracy of merger simulations. *Review of Economics and Statistics 95*(5), 1584–1600.

Wolfram, C. D. (1999). Measuring duopoly power in the british electricity spot market. *American Economic Review*, 805–826.

A Appendix

A.1 Demand estimation

In the terminology of BLP, the nonlinear parameters for our models are β_p, σ_p, and σ_l. We search over these parameters using a global optimization algorithm initially (differential evolution), and then we switch to Nelder-Meade.

A.2 Demand data

We aggregate demand over time and products. We define a product as all UPC codes that share a volume, value for the light beer dummy, and brand. The revenues and quantities (in units of six packs) are added up across such UPC codes for each month, and then prices are recovered by dividing revenues by quantities.

We aggregate stores according to Dominick's pricing zones. See Hellerstein (2008) for details.

At the UPC-week level, we define an "on-sale" dummy that is equal to unity if and only if any sort of promotional activity was recorded in the Dominicks' data for that UPC and week. Then, when aggregating the data across stores, weeks, and UPCs, we take the mean value of this dummy as our product-level measure of sales/promotional activity. The "number of products on sale" instrumental variable is the sum of our product-level sales/promotion measure over products within a market.

A.3 Optimal instruments

In our demand estimation, we consider an approximation to optimal instruments following Reynaert and Verboven's (2014) adaptation of Chamberlain's (1987) optimal instruments. After estimating the model with the baseline set of instruments for each of the above cases, we compute the derivatives of the error terms ξ with respect to the parameters:

$$\frac{\partial \xi_{jt}\left(\bar{p}\right)}{\partial \theta}. \tag{24}$$

The error term should be evaluated not at the actual prices, but at the expected prices given the instruments $\bar{p}$. Since we construct these optimal instruments only for a set of baseline instruments that includes the actual price, the projection $\bar{p}$ ends up being the actual price.

The derivatives in equation (24) are then used as instruments to estimate demand.

A.4 Production data

The CMF sends a questionnaire to all manufacturing plants in the United States with more than 5 employees every five years, while the ASM is a four-year rotating panel with re-

placement, sent to approximately a third of manufacturing plants, with large plants being over-represented in the sampling scheme. An industry is defined as a four digit SIC code. Labor is measured using the total number of employees at the plant. Materials are measured using total cost of parts and raw materials.

Capital is constructed in two ways. For the majority of plants, including all plants in the CMF, capital is measured using a question on total assets – be they machines or buildings – at the plant. For the remaining observations, capital is constructed using the perpetual inventory method, using industry-specific depreciation rates and investment deflators from the Bureau of Economic Analysis and the National Bureau of Economic Research. Capital, materials and sales are deflated using the NBER-CES industry-level deflators into 1997 dollars.

Our output measure is the total value of product shipments (deflated). Labor inputs are measured by total production hours. Materials inputs are measured by total expenditure on materials, electiricity, and fuel.

Tables

Table 1: Estimated Markups for Average and Selected Cars

(a) Demand Approach: BLP

Model	Markup $(P-c)$	Markup(P/c)
Mazda 323	\$801	1.19
Ford Escort	\$1,077	1.23
Lexus LS400	\$9,030	1.49
BMW 735	\$10,975	1.41
Mean	\$3,753	1.31

(b) Production Approach: DLW

Material expenditure	Unit price	Material cost share
\$7,493	\$10,672	0.85
	Markup $(P-c)$	Markup(P/c)
	\$1,852	1.21

Notes: Panel (a) is based on Table VIII from Berry et al. (1995) and reports price-cost margins for the year 1990 but in 1982 dollars. The last column is based on authors own calculations, and reports the price-cost margin as a ratio to compare with the production approach discussed below. Panel (b) reports unit price and material cost for the average car in the sample of Berry et al. (1996) for the year 1982, and the share of material cost in total cost of production. The last row reports the average markup using the FOC on cost minimization as given by equation (9).

Table 2: Demand data: summary statistics

SUMMARY STATISTICS							
year	store-months	products	brands	firms	HHI	p_{retail}	$p_{wholesale}$
1991	14	103.9	36	21	3285	2.92	2.59
1992	24	107.3	36	21	3038	3.04	2.72
1993	24	110.	37	22	2978	2.97	2.81
1994	24	108.9	38	23	3302	3.07	2.94
1995	18	108.4	38	23	3089	3.25	3.02
Total	104	108.	38	23	3127	3.05	2.82

Notes: Number of products, brands, and firms are means across store-months. The Herfindahl-Hirschman Index was computed for each store-month, then averaged. The mean retail price $\bar{p}_r$ and wholesale price $\bar{p}_w$are weighted by sales volume.

Table 3: Decomposing retail prices

	Mass Brewers		Craft Beer		Overall	
Component	Item	cum	Item	cum	Overall	cum.
Retail	0.80	4.01	1.29	6.45	1.05	5.23
Distributor	0.66	3.21	1.19	5.16	0.93	4.19
Tax/shipping	0.69	2.55	0.62	3.97	0.66	3.26
Brewer profit	0.24	1.86	0.67	3.35	0.46	2.61
adv/manag	0.33	1.62	0.54	2.68	0.44	2.15
labor	0.47	1.29	1.06	2.14	0.77	1.72
packaging	0.66	0.82	0.83	1.08	0.75	0.95
ingredients	0.16	0.16	0.25	0.25	0.21	0.21

Markups			
Brewer-low(*)	1.44	1.57	1.52
Brewer-high(**)	2.27	3.10	2.74
Distributor	1.26	1.30	1.28
Retailer	1.25	1.25	1.25

Notes: Source: Consumer Reports, 1996, (*) Excluding advertizing and management cost – treated as fixed cost. (**) Also treating labor as fixed cost.

Table 4: Estimated markups: instruments and demand specification

	(1)	(2)	(3)	(4)	(5)
Logit					
$\bar{\mu}_{1992}$	1.65	1.95	1.27	1.35	1.65
	(1.60,1.71)	(1.64,2.89)	(1.17,1.48)	(1.32,1.39)	(1.60,1.71)
mean $\left(\frac{d \ln q_j}{d \ln p_j}\right)$	-3.49	-2.94	-6.30	-5.21	-3.49
	(-3.65,-3.33)	(-3.52,-2.46)	(-9.34,-4.25)	(-5.58,-4.86)	(-3.65,-3.33)
$Pr\left(\mu_{1992,j}\right.$ well-defined)	1	1	1	1	1
	(1,1)	(0.99,1)	(1,1)	(1,1)	(1,1)
Income-Mixed Logit					
$\bar{\mu}_{1992}$	1.69	6.02	1.35	1.36	1.62
	(1.44,2.25)	(1.74,1.40e2)	(1.19,1.80)	(1.32,1.40)	(1.56,1.69)
mean $\left(\frac{d \ln q_j}{d \ln p_j}\right)$	-4.20	-2.33	-5.24	-5.25	-3.60
	(-5.28,-3.36)	(-4.12,-0.64)	(-8.54,-3.21)	(-5.70,-4.83)	(-3.81,-3.40)
$Pr\left(\mu_{1992,j}\right.$ well-defined)	1	0.89	1	1	1
	(0.99,1)	(0.01,1)	(1,1)	(1,1)	(1,1)
Random-Mixed Logit					
$\bar{\mu}_{1992}$	1.66	2.05	1.23	1.35	1.65
	(1.57,1.80)	(1.68,5.86)	(1.12,1.51)	(1.32,1.39)	(1.59,1.71)
mean $\left(\frac{d \ln q_j}{d \ln p_j}\right)$	-3.87	-2.78	-7.26	-5.21	-3.51
	(-4.40,-3.55)	(-3.47,-1.96)	(-12.90,-4.08)	(-5.58,-4.86)	(-3.68,-3.34)
$Pr\left(\mu_{1992,j}\right.$ well-defined)	1	1	1	1	1
	(1,1)	(0.94,1)	(1,1)	(1,1)	(1,1)

Notes: 90% confidence intervals in parentheses. Mean markups are share-weighted within domestically produced products with well-defined markups. Mean own-price elasticities are share-weighted within domestically produced products and evaluated at wholesale prices. Markups may be undefined, e.g. when a single-product firm's own price elasticity is less than unity in absolute value. Markups calculated assuming Nash-Bertrand pricing by brewer-distributors and fixed retail markups. Instruments based on: (1) Exogenous retail prices; (2) Exogenous wholesale prices; (3) Exogenous product set (BLP instruments); (4) Exogenous sales/promotions; (5) Approximate optimal instruments with exogenous retail prices.

Table 5: Estimated markups: model of competition

	(1)	(2)	(3)	(4)	(5)
Logit					
Nash-Bertrand (NB)	1.65	1.95	1.27	1.35	1.65
	(1.60,1.71)	(1.64,2.89)	(1.17,1.48)	(1.32,1.39)	(1.60,1.71)
NB With Transport	1.93	2.54	1.34	1.43	1.93
	(1.81,2.23)	(1.90,4.79)	(1.20,1.62)	(1.39,1.48)	(1.81,2.23)
NB No Retail	1.53	1.71	1.23	1.30	1.53
	(1.49,1.57)	(1.52,2.05)	(1.14,1.39)	(1.27,1.32)	(1.49,1.57)
Monopoly	1.85	2.32	1.32	1.42	1.85
	(1.76,1.99)	(1.83,4.46)	(1.19,1.59)	(1.38,1.47)	(1.76,1.99)
Income-Mixed Logit					
Nash-Bertrand (NB)	1.69	6.02	1.35	1.36	1.62
	(1.44,2.25)	(1.74,1.40e2)	(1.19,1.80)	(1.32,1.40)	(1.56,1.69)
NB With Transport	1.94	8.03	1.44	1.44	1.85
	(1.56,3.19)	(2.03,2.66e2)	(1.22,2.34)	(1.39,1.51)	(1.74,2.09)
NB No Retail	1.54	4.20	1.29	1.30	1.50
	(1.37,1.83)	(1.58,61.00)	(1.16,1.61)	(1.27,1.34)	(1.46,1.55)
Monopoly	4.75	12.30	1.42	1.46	1.79
	(2.03,17.30)	(2.49,2.65e2)	(1.22,2.08)	(1.40,1.54)	(1.70,1.94)
Random-Mixed Logit					
Nash-Bertrand (NB)	1.66	2.05	1.23	1.35	1.65
	(1.57,1.80)	(1.68,5.86)	(1.12,1.51)	(1.32,1.39)	(1.59,1.71)
NB With Transport	1.91	2.70	1.28	1.43	1.91
	(1.73,2.37)	(1.92,11.20)	(1.14,1.68)	(1.39,1.48)	(1.80,2.20)
NB No Retail	1.53	1.79	1.20	1.30	1.52
	(1.46,1.61)	(1.54,4.02)	(1.10,1.41)	(1.27,1.32)	(1.48,1.57)
Monopoly	5.88	3.06	1.28	1.42	1.84
	(1.90,19.70)	(1.97,41.00)	(1.15,1.63)	(1.38,1.47)	(1.75,1.97)

Notes: 90% confidence intervals in parentheses. Estimates are share-weighted means in 1992 within domestically produced products with well-defined markups. Markups may be undefined, e.g. when a single-product firm's own price elasticity is less than unity in absolute value. Instruments based on: (1) Exogenous retail prices; (2) Exogenous wholesale prices; (3) Exogenous product set (BLP instruments); (4) Exogenous sales/promotions; (5) Approximate optimal instruments with exogenous retail prices.

Table 6: Estimated markups: production technologies

	(1)	(2)	(3)
Technology	Gross output	Value added	
		Restricted Profit	Structural
Cobb-Douglas			
Variable M	1.65		
	(0.35)		
Variable L	1.53	2.91	1.69
	(1.35)	(0.50)	(0.08)
Translog			
Variable M	1.43		
	(0.76)		
Variable L	2.13	1.68	1.59
	(3.40)	(0.62)	(0.08)

Notes: Production-weighted mean markups across all breweries for 1992. Standard errors in parentheses. All specifications assume productivity innovations are uncorrelated with (m_{t-1}, l_t, k_t) and a linear first-order Markov process for productivity.

Table 7: Estimated markups: identification of production

	(1)	(2)	(3)	(4)
Instrument	l_t	l_t	l_{t-1}	l_{t-1}
Markov Process	Linear	Quadratic	Linear	Quadratic
Cobb-Douglas	1.69	1.69	1.83	1.83
	(0.08)	(0.08)	(0.06)	(0.06)
Translog	1.59	1.60	1.92	2.07
	(0.08)	(0.09)	(1.84)	(5.10)

Notes: Production-weighted mean markups across all breweries for 1992. Standard errors in parentheses. All specifications assume productivity innovations are uncorrelated with (m_{t-1}, k_t). Productivity innovations are also assumed uncorrelated with l_{t-1} or l_t, as listed, and productivity either follows a linear or quadratic first-order Markov process.

Table 8: Production function parameter estimates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
betal	0.125	0.298	0.436	1.390	0.538	1.615	0.749	0.766
	(0.110)	(0.185)	(0.075)	(0.216)	(0.058)	(0.241)	(0.118)	(0.359)
betam	0.700	0.717						
	(0.148)	(0.248)						
betak	0.177	0.164	0.538	-0.152	0.489	-0.106	0.300	0.235
	(0.061)	(0.124)	(0.098)	(0.196)	(0.065)	(0.200)	(0.123)	(0.259)
betal2		0.042		0.127		0.168		0.271
		(0.048)		(0.028)		(0.027)		(0.272)
betam2		0.054						
		(0.023)						
betak2		0.008		0.094		0.102		0.084
		(0.025)		(0.019)		(0.019)		(0.076)
betalm		-0.124						
		(0.056)						
betamk		-0.048						
		(0.046)						
betalk		0.052		-0.230		-0.279		-0.287
		(0.044)		(0.042)		(0.045)		(0.267)
betalmk		0.001						
		(0.002)						
Specification	GO	GO	RPVA	RPVA	SVA	SVA	SVA	SVA
Technology	CD	TL	CD	TL	CD	TL	CD	TL
Labor IV	l_t	l_t	l_t	l_t	l_t	l_t	l_{t-1}	l_{t-1}

Notes: Standard errors in parentheses. Technology is either Cobb-Douglas or translog. Specification is either gross output, restricted profit value added, or structural value added. Instruments include k_t and either l_t or l_{t-1}. Instruments also include m_{t-1} for gross output specifications and interactions for translog specifications. All specifications shown here assume that expected productivity is linear in lagged productivity.

Table 9: Demand system parameter estimates

	(1)	(2)	(3)	(4)	(5)
			Logit		
θ_p	0.288	0.118	0.879	0.689	0.288
	(0.028)	(0.107)	(0.236)	(0.041)	(0.028)
			Income-Mixed Logit		
θ_p	0.649	-0.406	0.703	0.866	0.366
	(0.166)	(0.618)	(0.295)	(0.083)	(0.037)
σ_p	-2.868	-1.473	0.023	0.291	0.116
	(0.819)	(1.637)	(0.024)	(0.105)	(0.017)
σ_l	-1.476	-7.966	0.470	-1.504	0.136
	(3.317)	(3.706)	(0.281)	(1.543)	(0.130)
			Random-Mixed Logit		
θ_p	1.107	0.342	1.030	0.689	0.293
	(0.311)	(0.370)	(0.357)	(0.041)	(0.029)
σ_p	1.141	0.580			
	(0.325)	(0.508)			
σ_l	15.910		3.150		
	(78.440)		(2.641)		

Notes: Standard errors in parentheses. Price coefficient is $-\exp(\theta_p + \sigma_p v_{i,1})$. Light beer dummy coefficient is $\theta_l + \sigma_l v_{i,2}$, where θ_l is absorbed by product dummies. In the Random-Mixed Logit, $v_{i,1}$ and $v_{i,2}$ have independent discrete distributions that approximate the Standard Normal Distribution. In the Income-Mixed Logit, $v_{i,1} = v_{i,2} = \log(income_i) - 11$, using a discrete distribution for $income_i$ that approximates the local income distribtion. Missing entries correspond to coefficients that converged to zero and were dropped from the model. Instruments based on: (1) Exogenous retail prices; (2) Exogenous wholesale prices; (3) Exogenous product set (BLP instruments); (4) Exogenous sales/promotions; (5) Approximate optimal instruments with exogenous retail prices.

Table 10: Cross-price elasticities

	Budweiser	Bud Light	Miller	Miller Lite
Logit				
Budweiser	-3.97	4.09e-3	0.04	0.03
Bud Light	0.01	-3.99	0.04	0.03
Miller	0.01	4.09e-3	-3.94	0.03
Miller Lite	0.01	4.09e-3	0.04	-3.95
Income-Mixed Logit				
Budweiser	-3.41	0.02	0.17	0.13
Bud Light	0.05	-3.81	0.17	0.14
Miller	0.05	0.02	-3.43	0.14
Miller Lite	0.05	0.02	0.17	-3.80
Income-Mixed Logit				
Budweiser	-3.92	4.26e-3	0.17	0.03
Bud Light	0.01	-4.27	0.04	0.46
Miller	0.05	4.26e-3	-3.85	0.03
Miller Lite	0.01	0.06	0.04	-3.97

Notes: Cross-price elasticities (with respect to retail prices) for selected six-pack products. Elasticities of quantity in row with respect to price in column. Elasticities evaluated at wholesale price. Median values across markets in 1992. Estimates based on exogenous retail prices.

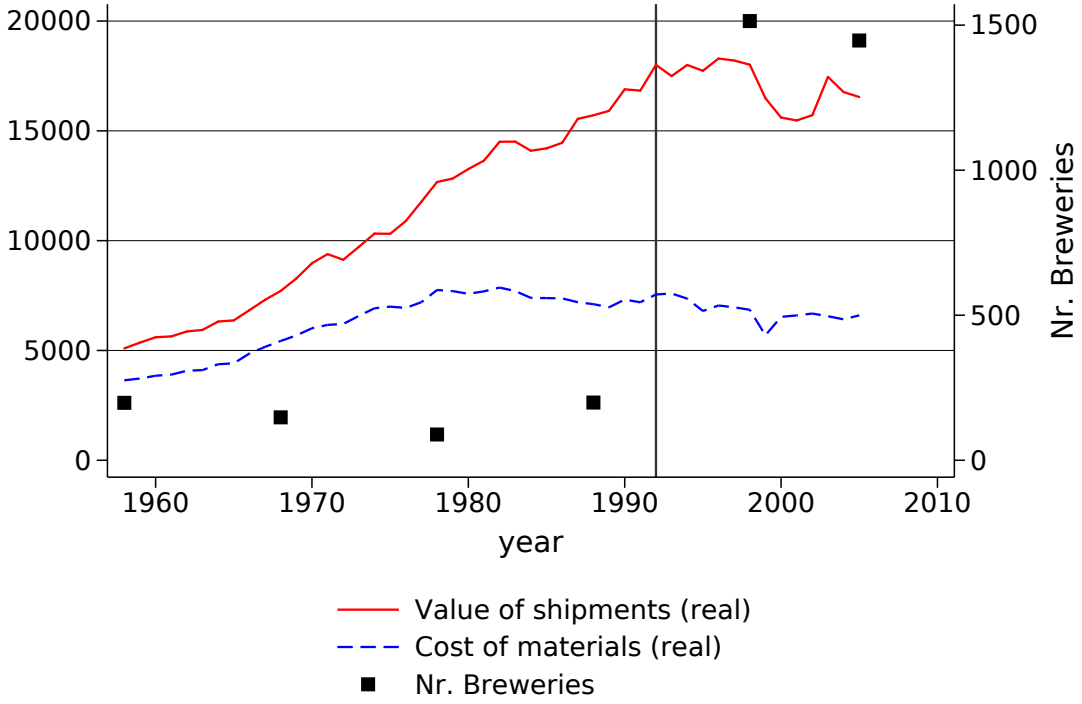
Table 11: Markups comparison

	(1)	(2)	(3)	(4)	(5)
Logit					
$\mu_{demand} = \mu_{production}$	OK	OK	reject	reject	OK
Battacharya coef	0.83	0.60	0.15	0.01	0.83
Income-Mixed Logit					
$\mu_{demand} = \mu_{production}$	OK	OK	OK	reject	OK
Battacharya coef	0.81	0.26	0.41	0.02	0.76
Random-Mixed Logit					
$\mu_{demand} = \mu_{production}$	OK	OK	reject	reject	OK
Battacharya coef	0.96	0.52	0.19	0.01	0.82

Notes: Comparison of share-weighted mean markups for domestic producers in 1992. Means compared using two-sided z-test at 5% significance, assumes mean markup estimate is asymptotically normal. Battacharya coefficient is $\int \sqrt{q_D(\mu)\, q_P(\mu)} d\mu$, where q_D and q_P are densities of the asymptotic distributions of the demand- and production-based mean markup estimate. The production specification relies on structural value added Cobb-Douglas techonology, assumes that productivity innovations are orthogonal to lagged materials and contemporaneous labor and capital, and assumes a linear first-order Markov process for productivity. Demand specifications assume static Nash-Bertrand price competition. Instruments based on: (1) Exogenous retail prices; (2) Exogenous wholesale prices; (3) Exogenous product set (BLP instruments); (4) Exogenous sales/promotions; (5) Approximate optimal instruments with exogenous retail prices.

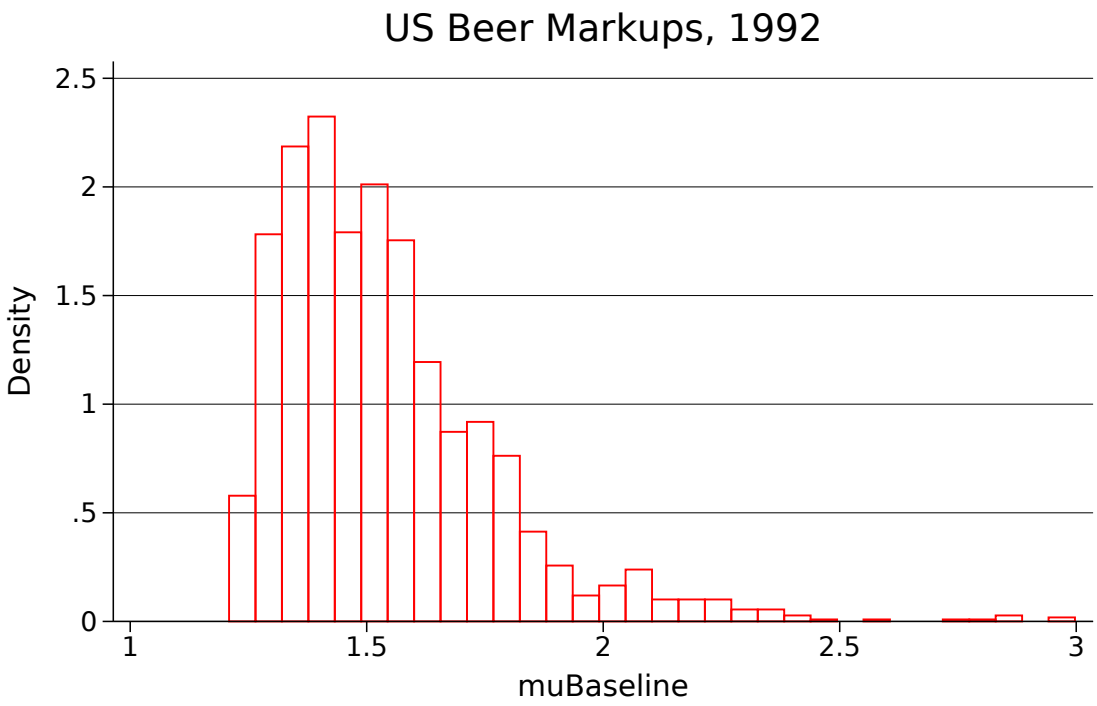
Figures

Figure 1: Shipments, Intermediate inputs and number of breweries



Notes: This figure plots the annual (real) value of shipments and intermediate inputs as reported by the NBER Manufacturing Database for NAICS 312120. The squares denote the number of active breweries (on the right vertical axis) as reported by the US Brewers Association. Vertical line indicates the year our production and demand samples overlap.

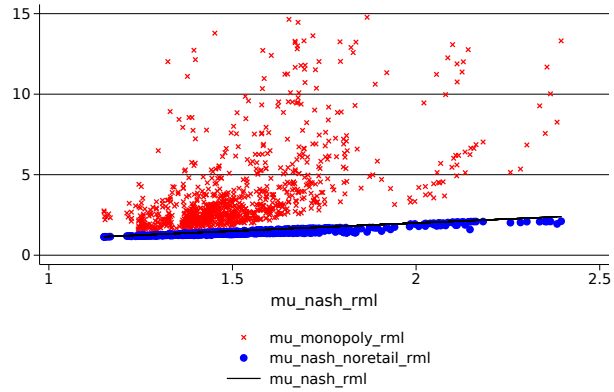
Figure 2: Distribution of Producer-level markups: demand approach



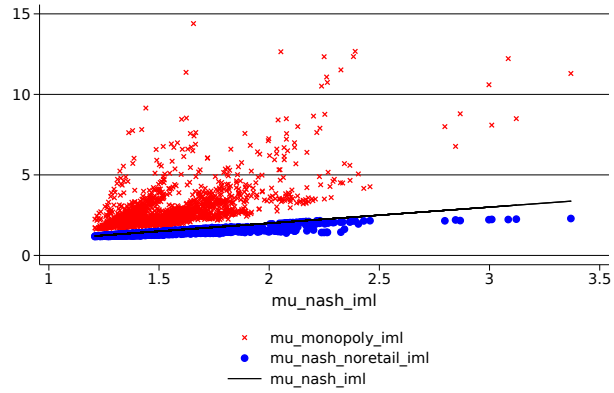
Notes This figure plots the histogram of product-level markups based on an income-mixed logit specification, exogenous retail prices, and Nash-Bertrand competition.

Figure 3: Markups across demand and competition specifications

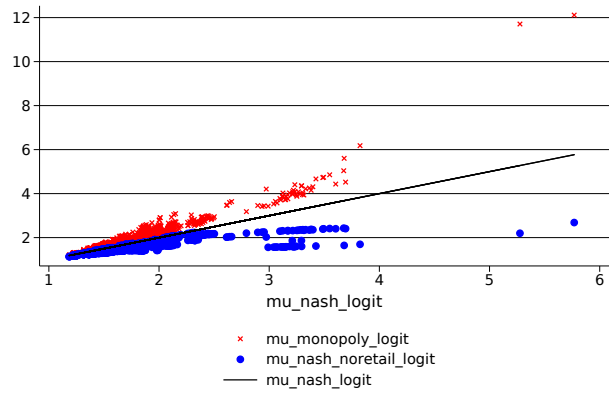
Demand specification: random mixed-logit



Demand specification: income mixed-logit



Demand specification: logit mixed-logit



Notes: Each panel is a scatter plot of estimated markups (across products and time) for a demand specification with exogenous retail prices, across three models of competition: Nash-Bertrand (our benchmark on the horizontal), Nash-Bertrand without retail, and monopoly. We trim the markup distribution from below at 1 and from above at 15 for each markup distribution.