Price Competition in U.S. Brewing
Author(s): Christian Rojas
Source: *The Journal of Industrial Economics*, Vol. 56, No. 1 (Mar., 2008), pp. 1-31
Published by: Wiley
Stable URL: http://www.jstor.org/stable/25483395
Accessed: 27-02-2017 20:43 UTC

THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821
Volume LVI March 2008 No. 1

PRICE COMPETITION IN U.S. BREWING*

CHRISTIAN ROJAS†

This study utilizes a brand-level dataset that captures a unique natural experiment, a 100% increase in the excise tax, to evaluate different pricing models in the U.S. beer industry. To assess the plausibility of different models, the increase in marginal cost resulting from the tax increase is exploited: observed prices in the post-increase period are compared to the prices that *should* be observed under various pricing models. Three types of models are analyzed: Bertrand-Nash, leadership, and collusion. Results indicate that extreme cases of collusion can be confidently ruled out while several models may explain the observed prices equally well.

I. INTRODUCTION

ECONOMISTS HAVE DEVOTED CONSIDERABLE EFFORT to the issue of identifying firms' pricing conduct when marginal cost data is unobserved. In homogeneous product markets, the typical approach is to estimate a conduct parameter that can lie in a continuum between competition and collusion (Bresnahan [1989]). Identification of conduct parameters in markets with many differentiated products, however, is difficult because the required product-level data is unlikely to exist (Nevo [1998]).[1]

The common approach used to study pricing conduct in differentiated products has been to consider a menu of plausible models of pricing conduct and rank them according to how well they fit the observed data. Measuring the fit of each model takes various forms. One alternative is to estimate directly different supply functions, one for each of the competing models, and construct pair-wise tests such as non-nested statistics (e.g., Gasmi, Laffont and Voung [1992]; Villas-Boas [2007]). This alternative is attractive

*I am indebted to Catherine Eckel for her encouragement, support and constructive comments. I thank Everett Peterson for his collaboration in related work and Victor Tremblay for his helpful comments. The Editor and two referees provided suggestions that have substantially improved this paper. I also thank Ronald Cotterill, Director of the Food Marketing Policy Research Center at the University of Connecticut for making IRI and LNA data available. Research and travel grants from the Department of Economics at Virginia Tech are acknowledged.

†Author's affiliation: Department of Resource Economics, 219 Stockbridge Hall, University of Massachusetts, Amherst, Massachusetts 01003, U.S.A.
e-mail: rojas@resecon.umass.edu

[1] In addition, there has been criticism on the conduct parameter approach at the theoretical and empirical level (e.g., Corts [1999]).

because of the straightforward interpretation of results, but supply needs to be modeled directly and it is often difficult to obtain rich enough and reliable supply data at the product level. In addition, with a large number of products and many competing models, constructing non-nested tests is a non-trivial task.

A second alternative is to use demand estimates to compute the price-cost margins implied by the competing models and compare them to observed price-cost margins (Nevo [2001]; Slade [2004]). A potential problem of this approach is that the crude measures of observed price-cost margins (i.e., average price-cost margins across products) that are available may yield unreliable inference.

This paper employs an alternative approach to evaluate pricing conduct by exploiting a unique natural experiment. In 1991, the excise tax on beer in the U.S. was doubled, effectively increasing the marginal cost of all domestic and imported beers. This study uses a brand-level dataset that captures both pre- and post-increase periods. Briefly put, the large increase in marginal cost is used to compute the 'predicted' prices that each of the competing models would yield when the tax is introduced. Model comparisons are then based on metrics that quantify the closeness of each model's predicted prices to the 'observed' prices in the post-increase period.

Previous research has used the excise tax rate to analyze firm conduct in the context of a homogeneous product model. Sumner [1981] and Sullivan [1985] exploit the state and time variation of the cigarette excise tax to identify the degree of competition in the industry. Sumner estimates the average pass-through rate of the excise tax and relates it to the average firm elasticity and its implied pricing conduct; Sumner concludes that the estimated average elasticity is consistent with competitive conduct. Sullivan employs a more flexible strategy and reaches a similar conclusion. Data limitations do not allow these authors to investigate the issue at the firm or product level, which is important because the strategic behavior among firms is a key factor in the pass-through rate of excise taxes (Anderson, de Palma and Kreider [2001]).

Studying firm conduct in the U.S. brewing industry is an interesting question in its own right. U.S. brewing has experienced a dramatic change from a fragmented industry to a highly concentrated oligopoly. The number of mass-producing brewers has declined from 350 in 1950 to 24 in 2000 with a corresponding increase in the Herfindahl-Hirschman index from 204 to 3,612 (Tremblay and Tremblay [2005]), making this industry one of the most concentrated in the U.S.[2]

[2] For comparison with other highly concentrated industries, the HHI's for cigarettes, breakfast cereals and automobiles are 3,100, 2,446 and 2,506, respectively. The average index for all manufacturing industries is 91 (U.S. Census Bureau, [1997] concentration ratios).

This rising concentration has often raised concerns about market power and non-cooperative behavior (Tremblay and Tremblay [2005]). In addition, Greer [1998] and Tremblay and Tremblay [2005], and references therein) identify Anheuser-Busch as a price leader especially through its heavily marketed brand Budweiser. Evidence supports the fact that by the 1990's, Anheuser-Busch had become the clear price leader (Tremblay and Tremblay [2005]: 171; Greer: 49–51).[3]

Earlier studies (Tremblay and Tremblay [1995] and [2005]) have found that the degree of market power in U.S. brewing is low; however, their analysis is limited to firm-level data. This limitation is particularly relevant in the U.S. brewing market where product differentiation is important and can give rise to prices above marginal costs even when deviations from competitive pricing (Bertrand-Nash) are non-existent.

Brand-level studies in other industries (e.g., Nevo [2001]; Slade [2004]) have considered Bertrand-Nash and collusion as the alternative modes of competition; this paper entertains two types of leadership models in addition to Bertrand-Nash and collusion.[4] Both leadership models are intended to reflect and formally test the forms of leadership reported in this industry. The first model is called 'collusive price leadership' in which followers match Budweiser's price changes. The second model is Stackelberg with two variants. In one variant, Budweiser acts as the price leader while in the other variant Anheuser-Busch leads other brands with its entire product line. As an additional reference, the hypothetical case of Bertrand-Nash with single-product firms (which can be thought of as the portion of mark-up due to product differentiation alone)[5] is compared to the other models.

The dataset is comprised of brand-level prices and quantities collected by scanning devices in 58 major metropolitan areas of the United States over a period of 20 quarters (1988–1992). The empirical strategy consists of four stages. First, a structural demand model for 64 brands is estimated. The demand model is based on the neoclassical 'representative consumer' approach rather than on the, more popular, 'discrete choice' approach. The discrete choice assumption seems appropriate for products like automobiles, but it appears less natural for beer. The major challenge of estimating

[3] An anecdotal example that supports this view is the statement by Robert Uihlein, Chairman of the Schlitz Brewery: 'A price increase is needed, but it will take Anheuser-Busch to do it.' (FORTUNE [November, 1975, p. 92])

[4] Two exceptions are Kadiyali, Vilcassim and Chintagunta [1996], and Gasmi, Laffont and Vuong who consider price leadership as an alternative mode of competition. These applications are limited to the Stackelberg model and a small number of products (4 and 2, respectively).

[5] The other portion of mark-ups for multi-product Bertrand-Nash competitors is due to concentration (or fewness of multi-product firms) and is equal to the difference between the mark-ups given by multi-product and single-product Bertrand-Nash competition (Nevo [2001]).

numerous substitution coefficients is dealt with by the Distance Metric (DM) method devised by Pinkse, Slade and Brett [2002]. This paper extends previous applications of the DM method (Pinkse and Slade [2004]; Slade [2004]) by also estimating advertising substitution patterns.[6]

In a second stage, the demand estimates are used to compute the marginal costs implied by each of the competing models during the period prior to the tax increase. The new marginal costs (pre-increase marginal costs plus the tax increase) are then used to find the new equilibrium, or predicted, prices that should be observed under each model when the tax is increased. Finally, the predicted prices are compared to the actual prices to study which model appears to be better supported by the data.

Results clearly rule out full collusion among all firms and the case of collusive price leadership. Some evidence indicates that both Stackelberg models may be better predictors of firm behavior. However, the prices predicted by Bertrand-Nash, Stackelberg leadership, and collusion among selected brands and firms, are not largely dissimilar. A discussion of these findings and their relation to previous work is presented in the conclusion.

II. THE INDUSTRY AND THE TAX INCREASE

Commercial brewing began during the colonial period. By 1810 there were 132 breweries producing 185,000 barrels of mainly English and Irish-type (ale, porter and stout) malt beverages. Lager beer was introduced in the mid-nineteenth century and today it accounts for over 90 per cent of the U.S. brewing industry's output.[7] Overall, total demand for beer in the U.S. has been constantly increasing since the mid-twentieth century. Between 1960 and 1980 strong consumption growth was observed, but for the last three, decades demand for beer has remained stagnant (180–210 million barrels per year). Per capita consumption has fluctuated but has stabilized at approximately 22 gallons per year.

Currently, the advertising-to-sales ratio for beer is 8.7 per cent compared to 2.9 per cent for cigarettes, and 7.1 per cent for other beverages (Advertising Age [2000]; cited in Tremblay and Tremblay [2005]). National brewers have taken advantage of the more cost-effective marketing channel: national TV. Larger national producers have driven many regional producers out of business partly because of this marketing disadvantage but also because of technological changes that required larger plants to achieve a minimum efficient scale (MES).

[6] Incorporating advertising into the demand system helps improve the validity of the price instruments (see section V(i)).

[7] A commonly used classification for beers sorts them into lagers and ales. Lagers are brewed with yeasts that ferment at the bottom of the fermenting tank. Ales are brewed with yeasts fermenting at high temperatures and at the top of the fermenting tank. Porter and Irish are darker and sweeter than ale, with minimal market share in the United States today.

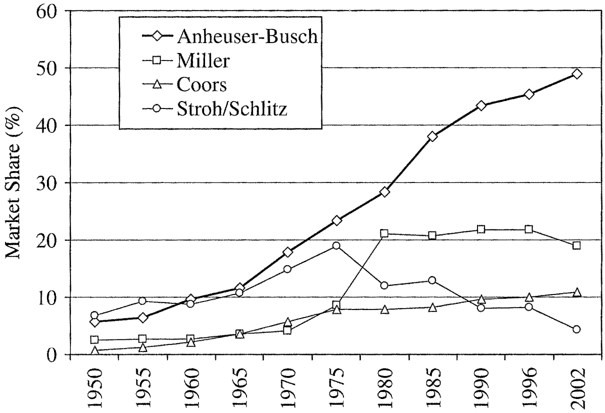

Figure 1
Market shares of largest brewers in the U.S.
Source: Greer (1998), Beer Marketer's Insights.
Note: Stroh exited the market in 1999. The 2003 number corresponds to Pabst's market share (the acquirer of some of Stroh's brands).

In 2003, nearly 80% of beer sales in the U.S. was concentrated among three firms: Anheuser-Busch (49.8%), SABMiller (17.8%) (formerly Miller and owned by Philip Morris) and Coors (10.7%). Anheuser-Busch has been the largest beer producer since 1960, with an ever increasing market share (Figure 1). Budweiser and Bud Light, Anheuser-Busch's two leading brands, currently capture approximately one third of beer sales nationwide.

The industry is characterized by numerous product introductions and, consequently, a large number of brands. An interesting fact is the increasing popularity of light beer. Since the successful introduction of Miller Lite in the 1970's, light beers have become the most popular beer type and now account for almost half the sales of beer in the U.S.

While imports and specialty beers have increased their combined market share from less than 1% in the 1970's to approximately 12% and 3%, respectively, their impact in the industry as a whole remains limited. The reason is that imports and specialty beers tend to compete less directly with traditional mass-producers since they target different types of consumers.

U.S. brewing remains as one of the most interesting industries because of its ramifications to other important issues such as health, taxation and regulation. Tremblay and Tremblay [2005] present the most comprehensive economic analysis of this industry to date.

The Federal Excise Tax Increase

In 1990, U.S. Congress approved an increase in the federal excise tax on beer from $9 to $18 per barrel. All brewers and importers were required to

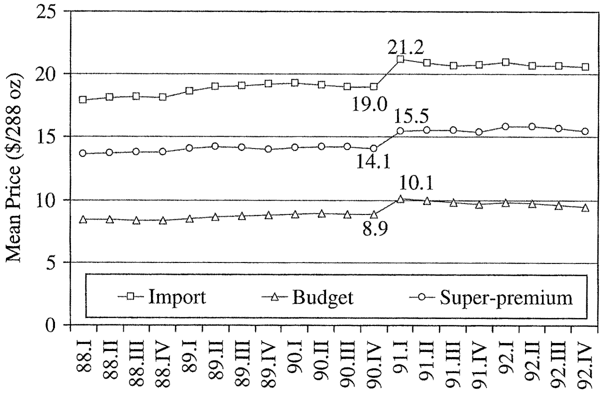

Figure 2
Quarterly mean prices, various beer segments (1988–1992)
Source: IRI Database, University of Connecticut

pay this tax on all produced units as of January, 1991. This increase, which was equivalent to an additional 64 cents in federal taxes per 288 ounces (a 24-pack), represented the largest federal tax hike for beer in U.S. history.

Figure 2 shows mean quarterly prices (over all cities) for three beer segments using the data set available for this paper. There is a clear shift in the mean price of all three categories in the first quarter of 1991. All mean increases are higher than the actual tax hike of 64 cents per 288 ounces: 220 cents for imports, 140 cents for super-premium beers and 120 cents for budget beers. These mean increases were 237%, 114%, and 84%, respectively, larger than the tax increase of 64 cents per case. This is consistent with the theoretical findings of Anderson, de Palma and Kreider who show that in oligopolies with differentiated products, an excise tax can be passed on to consumers by more than 100%.

III. THE EMPIRICAL MODEL

Comparison of different pricing models is carried out by exploiting the exogenous variation of an increase in the federal excise tax. Since all pricing models require estimates of own-price and cross-price elasticities, the first step is to estimate the demand for beer at the brand-level. With these estimates, the implied marginal costs of all brands are computed for each of the models. This computation is carried out in each quarter that preceded the tax increase. Using each brand's median marginal cost (over the pre-tax-increase period), the demand elasticity estimates, the pre-tax-increase values of the remaining variables, and the increase in marginal cost due to the tax increase, the post-tax increase equilibrium prices are computed for each pricing model. These predicted price increases are then compared to the actual price increases. This section provides details on demand, supply and the computation of marginal costs. Sections V(ii) and V(iii) present details on the computation of equilibrium prices and actual prices increases.

III(i). *Demand*

Let $\psi = \{1, \ldots, J\}$ be the product set, $\Xi = \{1, \ldots, T\}$ the set of markets (in this study a market is defined as a city-quarter pair), $q_t = \{q_{1t}, \ldots, q_{Jt}\}$ the vector of quantities demanded, $p_t = \{p_{1t}, \ldots, p_{Jt}\}$ the corresponding price vector and $x_t = \sum_j p_{jt} q_{jt}$ total expenditures. The linear approximation to the Almost Ideal Demand System (LALIDS) of Deaton and Muellbauer is used due its desirable theoretical properties:

$$(1) \qquad w_{jt} = a^*_{jt} + \sum\nolimits_k b_{jk} \log p_{kt} + d_j \log(x_t/P_t)$$

where $w_{jt} = \frac{p_{jt} q_{jt}}{x_t}$ is brand $j's$ sales share and log P_t is a price index approximated the loglinear analogue of the Laspeyeres index:[8]

$$(2) \qquad \log P_t \approx \sum\nolimits_j w^o_j \log(p_{jt})$$

where w^o_j is brand $j's$ 'base' share, defined as $w^o_j \equiv T^{-1} \sum_t w_{jt}$.[9]

Traditional advertising (e.g. television, radio and press) is considered the key advertising variable because of its crucial role in the development of the industry. Further, only the flow effects of advertising are considered with all lagged own- and cross-advertising terms being omitted for the demand equation.[10]

Advertising for brand k (A_k) is incorporated into equation (1) by defining the intercept term as: $a^*_{jt} = a_{jt} + \sum_k c_{jk} A^\gamma_{kt}$. The parameter γ is included to account for decreasing returns to advertising. Following Gasmi, Laffont and Voung, γ is set equal to 0.5. Substituting the redefined intercept into equation (1) and including an econometric error term gives:

$$(3) \qquad w_{jt} = a_{jt} + \sum\nolimits_k c_{jk} A^\gamma_{kt} + \sum\nolimits_k b_{jk} \log p_{kt} + d_j \log(x_t/P_t) + e_{jt}$$

Equation (3) is as a first-order approximation in prices and advertising to a demand function that allows unrestricted price and advertising parameters. In order to reduce the number of cross-price and cross-advertising coefficients that need to be estimated, the Distance Metric (DM) method of Pinkse, Slade and Brett is employed. This method specifies each

[8] Moschini [1995] explains how this price index can have superior approximating properties than the Stone price index of Deaton and Muellbauer.

[9] The 'fixed' base w^o_j moderates the problem of having an additional endogenous variable on the right hand side of (1).

[10] The existence of possible stock effects was investigated but the estimated coefficients on lagged advertising expenditures were found not to be statistically different from zero.

cross-coefficient (b_{jk} and c_{jk}) as a function of the distance between brands j and k in product space.

Distance measures may be either continuous or discrete. For example, alcohol content can be used to construct a continuous distance measure. Dichotomous variables that group brands into different market segments are used to construct discrete distance measures and take a value of 1 if brands j and k belong to the same grouping and zero otherwise. Continuous distance measures use an inverse measure of distance (closeness) between brands.

The terms b_{jk} and c_{jk} are specified as a linear combination of distance measures:

$$b_{jk} = \sum_{r=1}^{R} \lambda_r \delta_{jk}^r \tag{4}$$

$$C_{jk} = \sum_{S=1}^{S} \tau_S P_{jk}^S \tag{5}$$

where $\delta_{jk} = \{\delta_{jk}^1, ..., \delta_{jk}^R\}$ is the set of distance measures for cross-prices and $\mu_{jk} = \{\mu_{jk}^1, ..., \mu_{jk}^S\}$ the set of measures for cross-advertising; λ and τ are the coefficients to be estimated.[11] After replacing (4) and (5) into (3) and regrouping terms gives the empirical demand equation:

$$\begin{aligned} w_{jt} =& a_{jt} + b_{jj} \log p_{jt} + c_{jj} A_{jt}^{\gamma} + \sum\nolimits_{r=1}^{R} \left(\lambda_r \sum\nolimits_k \delta_{jk}^r \log p_{kt} \right) + \\ &+ \sum\nolimits_{s=1}^{S} \left(\tau_s \sum\nolimits_k \mu_{jk}^s A_k^{\gamma} \right) + d_j \log(x_t/P_t) + e_{jt} \end{aligned} \tag{6}$$

The estimated coefficients λ_r and τ_s and the distance measures between brands (δ_{jk} and μ_{jk}) are replaced into (4) and (5) to obtain cross-terms (b_{jk} and c_{jk}). Since distance measures are symmetric by definition, and to reduce the number of parameters to be estimated, symmetry (i.e., $b_{jk} = b_{kj}$ and $c_{jk} = c_{kj}$) is imposed by setting λ and τ to be equal across equations.

In principle, $(J-1)$ seemingly unrelated equations can be estimated. However, since J is very large, it becomes impractical to estimate such a large system. Alternatively, it is assumed that the own-price and own-advertising coefficients (b_{jj} and c_{jj}), and the price index coefficient (d_j), are equal across equations thereby reducing estimation to one equation (since symmetry is also imposed). Since this is too strong of an assumption, and following

[11] Various specifications of the semi-parametric estimator proposed by Pinkse, Slade and Brett were implemented to check that the parametric specification of h and g in (4) and (5) is not a restrictive functional form. See Rojas [2005].

Pinkse and Slade, the coefficients b_{jj}, c_{jj}, and d_j are specified as linear functions of brand j's characteristics.

Because of the large price increase in prices of all beers as a result of the tax increase, it is important to allow the budget share allocations x_t to change as the overall price of beer changes. This can be done by using a two-stage budgeting approach, where the bottom level of brand-level demand is given by (6) and the top-level demand for beer is modeled as (see Hausman, Leonard and Zona [1994]):

$$\log q_t = \beta_0 + \beta_1 \log Y_t + \beta_2 \log \Pi_t + \tilde{Z}_t \phi + \varepsilon_t \tag{7}$$

where Y_t is income, $\prod_t$ is a deflated price index for beer, $\tilde{Z}_t$ is a vector of city and time- and city-specific dummies, and $\beta = (\beta_0, \beta_1, \beta_2)$ and ϕ are vectors of parameters to be estimated. The main parameter of interest is β_2 because it measures the sensitivity of demand for 'all' beer to changes in the overall price of beer. Since x_t in (6) is a function of (7), *unconditional* elasticities and price derivatives can then be obtained by applying the chain rule to (6).

Continuous Distance Measures

The characteristics utilized are alcohol content (*ALC*), product coverage (*COV*), and container size (*SIZE*). Product coverage measures the fraction of the city in which a brand is present and is defined as the all commodity value (ACV) of stores carrying the product divided by the ACV of all stores in that city. Beers with low coverage may be interpreted as specialty brands that are targeted to a particular segment of the population.[12] Beer is sold in a variety of sizes (e.g., six and twelve packs), and the variable *SIZE* measures the average 'package size' of a brand. Higher volume brands (e.g., typical sales of twelve packs and cases) may compete less strongly with brands that are sold in smaller packages (e.g., six packs).

The characteristics of a brand determine its location in product space. Using these locations, inverse measures of distance (closeness) in one- and two-dimensional Euclidean spaces are computed for all pairs of brands.[13]

Discrete Distance Measures

Three different types of discrete distance measures are utilized. The first type focuses on various product groupings including product segment,

[12] For example, Miller Genuine Draft is in virtually all retail outlets of all cities whereas the imported Labatt's Blue is only present, on average, in 45% of a city's supermarkets.

[13] The inverse measure of distance between brands j and k is defined as: $1/[1 + 2(\text{Euclidean distance between } j \text{ and } k)]$. Euclidean distance in one-dimensional space is the absolute difference in the value of the characteristic between j and k. In n-dimensional space, Euclidean distance is equal to $\left[\sum_n (j_n - k_n)^2\right]^{1/2}$, where the subscript is the brand's coordinate in each of the n dimensions.

brewer identity, and national brand identity. With no clear consensus on product segment classifications, five different classifications are considered: (1) budget, light, premium, super-premium, and imports, (2) light and regular, (3) budget, light, and premium, (4) domestic and import, and (5) budget, premium, super-premium, and imports. Because brand competition may be stronger across brands from different brewers, a discrete measure is constructed to identify all brands by the same brewer. Similarly, brands that are national (regional) may compete more strongly with one another so a discrete distance measure that groups brands by whether they are national or regional is created.

Following PSB, two other types of discrete measures are constructed based on the nearest neighbor concept and whether products share a common boundary in product space. Brands j and k share a common boundary if there is a set of consumers that would be indifferent between both brands and prefer these two brands over any other brand in product space. The nearest neighbor and common boundary measures are computed for all pairs brands based on their location in alcohol content and coverage space, and coverage and container size space. A second set of nearest neighbor and common boundary measures are computed using both product characteristics and price thereby allowing consumers' brand choices to be influenced not only by distance in characteristics space but also by price (see Rojas for details).

Own-Price and Own-Advertising Interactions

Two product characteristics are interacted with own-price and own-advertising in the model:[14] the inverse of product coverage ($1/COV$) and the number of common boundary neighbors (NCB). The number of common boundary neighbors is a measure of local competition that determines the number of competitors that are closely located to a brand in product space. NCB is computed in product coverage-container size space and alcohol content-coverage space.

III(ii). *Supply*

Let F_n be the set of brands produced by firm n. Assuming constant marginal costs and linear additivity of advertising, the profit of firm n in a given market is expressed as:

$$\pi_n = \sum_{j \in F_n} (p_j - c_j) q_j(p, A) - \sum_{j \in F_n} A_j \tag{8}$$

[14] These were the interactions that yielded the largest explanatory power in several specifications.

where c_j is brand j's marginal cost, p_j is its price and A_j is firm n's advertising expenditures on brand j. Firm n's first order conditions can be expressed as:

$$(9) \qquad q_j(p, A) + \sum_{k \in F_n} (p_k - c_k) \overline{\frac{\partial q_k}{\partial p_j}} = 0, \text{ with respect to } p_j$$

$$(10) \qquad \sum_{k \in F_n} (p_k - c_k) \frac{\partial q_k}{\partial A_j} - 1 = 0, \text{ with respect to } A_j$$

where $\overline{\frac{\partial q_k}{\partial p_j}} = \frac{\partial p_k}{\partial p_j} + \sum_{m \notin F_n} \frac{\partial q_k}{\partial p_m} \frac{dp_m}{dp_j}$. Partial derivatives in (9) and (10) are the unconditional price derivatives obtained from demand estimates for equations (6) and (7). The term $\frac{dp_m}{dp_j}$ in $\overline{\frac{\partial q_k}{\partial p_j}}$, however, takes different values depending on the model of interest; this term is the 'conjecture' of firm n about how the price of product m will react to a change in the price of product j.

In principle, several games in advertising can also be considered. However, simulations of collusion, Bertrand-Nash and Stackelberg games in advertising produced equilibrium conditions that were essentially indistinguishable from each other; the reason is the small magnitude of advertising coefficients obtained from demand estimation. Consequently, price is treated as the main strategic variable of interest and it is assumed that firms compete in a Bertrand-Nash fashion in advertising.

Bertrand-Nash and Collusion

For Bertrand-Nash competition in prices, the conjecture takes a value of zero. The conjecture is also zero in the of single-product Bertrand-Nash competition and in the case of collusion, but the ownership sets (F_n) are modified to reflect the profit-maximizing conditions of single-product firms (i.e., F_n's are singletons) and colluding firms (i.e. joint profit-maximization), respectively.

Stackelberg Leadership

Two cases are considered, one in which Budweiser leads brands produced by firms other than Anheuser-Busch and the other in which Anheuser-Busch leads with its entire product line. For this game, the conjecture term $\left(\frac{dp_m}{dp_j}\right)$ takes a value of zero if j is a follower brand. If j is a leading brand, the conjecture term is computed from the first order conditions of followers by applying the implicit function theorem. Appendix A contains details of this procedure.

Collusive Price Leadership

In this case followers exactly match Budweiser's price changes. In this 'collusive price leadership' scenario only the first order conditions of the firm producing the leading brand (i.e., Anheuser-Busch) are relevant, since followers do not price via profit-maximization but by imitating the leader. The term $\frac{dp_m}{dp_j}$ in (9) is set to 1 in Budweiser's first order condition and to zero in Anheuser-Busch's remaining first order conditions.[15]

III(iii). *Marginal Costs*

In each market, there are two equations for each unknown marginal cost (c_j). After adding up (9) and (10) for each brand j, a solution for c_j is obtained in this new system.[16] The system in vector notation is:

$$(11) \qquad Q^o - \Delta(p - c) = 0$$

where Q^o and $(p - c)$ are $J \times 1$ vectors with elements $(q_j\ (p, A) - 1)$ and $(p_j - c_j)$, respectively; Δ is a $J \times J$ matrix with typical element $\Delta_{jk} = -\Delta^*_{jk}\left(\overline{\frac{\partial q_k}{\partial p_j}} + \frac{\partial q_k}{\partial A_j}\right)$, where Δ^*_{jk} takes a value of 1 if brands j and k are produced by same firm and zero otherwise. Applying simple inversion to Δ in (11) gives the implied marginal costs:

$$(12) \qquad c = p - \Delta^{-1}Q^o$$

Marginal costs in each market are computed using the demand estimates and the appropriate values of the conjectures $\left(\frac{dp_m}{dp_j}\right)$ for each of the models. Collusive possibilities (e.g., between specific products or firms) are investigated by appropriately modifying the elements Δ^*_{jk} (which determine the ownership sets F_n). For example, full collusion, or joint profit maximization, is equivalent to setting all Δ^*_{jk} elements to equal one. Appendix A provides details on the computation of marginal costs for the leadership models.

IV. DATA

Table I provides a description and summary statistics of the variable used. The main source is the Information Resources Inc. (IRI) Infoscan Database. The IRI data includes prices and total sales for several hundred brands for

[15] Appendix A contains details of computational problems (and the solutions adopted) that arise in both types of leadership models (Stackelberg and Collusive).

[16] Since this is a linear problem, the solution is unique. Moreover, if c_j is the same in both (9) and (10) (which it is by assumption) the solution will solve (9) and (10) individually. If, on the other hand, two different c_j's solve (9) and (10), the solution of the added system will be a linear combination of the two c_j's.

TABLE I
DATA DESCRIPTION SUMMARY OF STATISTICS

Variable	Description	Units	Mean	St dev	Min	Max
Price	Average Price per brand	$/288oz	12.1	3.87	0.82	28.9
Quantity	Volume Sold	288 oz	23.53	63.6	0.00	2652
SIZE	Quantity/Units Units = #of units sold, all sizes	N/A	0.38	0.117	0.08	1.30
Coverage (COV)	Sum of all commodity value (ACV) sold by stores carrying the product/ACV of all stores in the city	%	74.0	28.61	0.26	100
OVER50K	% of Households with income over $50,000/year	%	23.3	6.1	10.3	44.8
A	Quarterly national advertising expenditures	Mill of $	3.54	6.3	0	40.37
ALC	Alcohol Content	%/vol	4.48	0.94	0.4	5.25
R	1 if brand is regional, 0 otherwise	0/1	0.15	–	–	–
WAGES	Average wage of worker in retail sector	$/hour	7.3	1.17	3.58	12.3
DEN	Population per square mile	(000)	4.73	4.13	0.73	23.7
INCOME	Median Income	(000) of $	32.02	6.9	18.1	53.4

Source: IRI database, University of Connecticut; Bureau of Labor Statistics; Demographia; other sources.

up to 58 cities over 20 quarters (1988–1992).[17] Volume sales (Quantity) in each city are reported as the number of 288-ounce units sold each quarter by all supermarkets in that city and price is an average price for a volume of 288 oz. for each brand. To maintain focus on brands with significant market share, all brands with a local market share of less than 3% are excluded from the sample. This selection criterion provides a sample of 64 brands produced by 13 different brewers. Appendix B contains a table of all the brands chosen as well as other details of the database and the data selection procedure.

In addition to price and sales data, IRI has information on other brand specific and market variables. Because beer is sold in a variety of sizes (e.g., six and twelve packs), the variable *UNITS* provides the number of units, regardless of size, sold each quarter. An average size variable is created: $SIZE = \text{Quantity}/UNITS$. The variable *COV* measures the degree of city coverage for each brand. Lastly, the variable OVER50K, which is the fraction of households that have an income above $50,000 in each city-quarter pair, was also included in the estimation.

Advertising data (A) was obtained from the Leading National Advertising annual publication. These are quarterly data by brand comprising total national advertising expenditures for 10 media types. Alcohol content (ALC) was collected from various specialized sources.

Data for demand side instruments were collected from additional sources. A proxy for supermarkets labor cost ($WAGES$) is constructed

[17] The actual market definitions of these cities are broader than a single city and are usually referred to as 'metropolitan areas.' The term city here is used for simplicity. In general, the definition of these metropolitan areas is broader than the BLS definitions.

from data from the Bureau of Labor Statistics CPS monthly earning files. City density estimates (*DEN*), collected from Demographia and the Bureau of Labor Statistics, were included to proxy for cost of shelf space. *INCOME* from the IRI database was used to instrument for expenditures (x_t).

V. ESTIMATION

V(i). *Demand and Instruments*

Because firms are assumed to choose both price and advertising, these variables are treated as endogenous variables. To avoid simultaneity bias, an instrumental variables approach is used to consistently estimate both brand-level demand (6) and top-level demand (7).

Brand-level Demand

Let n_z be the number of instruments, Z the $(T \times J) \times n_z$ matrix of instruments, S the collection of right hand side variables in equation (6), θ the vector of parameters to be estimated and $\underline{w}$ sales shares in vector form. The generalized method of moments (GMM) estimator $\hat{\theta}_{GMM} = (S'P_zS)^{-1}S'P_z\underline{w}$ is employed. The consistent estimator for its asymptotic variance is defined as $A\text{var}(\hat{\theta}_{GMM}) = (S'P_zS)^{-1}$, where $P_z = Z(Z'\hat{\Omega}Z')^{-1}Z$ and Ω is a diagonal matrix with diagonal element equal to the squared residual obtained from a 'first step' 2SLS regression.

As in previous work, the instruments employed in this paper rely on the identification assumption that, after controlling for brand, city, and time specific effects, demand shocks are independent across cities. Because beer is produced in large plants and distributed to various states, the prices of a brand across different cities share a common marginal cost component, implying that prices of a given brand are correlated across markets. If the identifying assumption is true, prices will not be correlated with demand shocks in other markets and can hence be used as instruments for other markets. In particular, the average price of a brand in other cities is used as its instrument.

The data employed in this study are based on broadly defined markets. These broad market definitions, which are similar to those used by the Bureau of Labor Statistics, reduce the possibility of potential correlation between the unobserved shocks across markets. Furthermore, demand shocks that may be correlated across markets because of broad advertising strategies are controlled for by including national advertising expenditures in the demand equation. To further control of other potential unobserved national shocks, time dummies are included in the estimation.

Because advertising expenditures are only observed at the national level each quarter, lagged advertising expenditures are used as its instrument.

Expenditures (x_t), which is constructed with price and quantity variables, is also treated as endogenous and is instrumented with median income.

A final identification assumption, which is common practice in the literature, is that product characteristics are assumed to be mean independent of the error term. The validity of the proposed instruments is assessed by conducting a formal test. Additional instruments for price are created from city-specific marginal costs (i.e., proxies for shelf space and transportation costs, see Nevo, [2001]) and an overidentifying restrictions test is used to check the validity of instruments.

As observed by Berry [1994], an additional source of endogeneity may be present in differentiated products industries. Unobserved product characteristics (included in the error term), which can be interpreted as product quality, style, durability, status, or brand valuation, may be correlated with price and advertising and produce biases in the estimated coefficients. This source of endogeneity is controlled for with the inclusion of brand-specific fixed effects (Nevo [2001]). These fixed effects control for the unobserved product characteristics that are invariant across markets, reducing the bias and improving the fit of the model.

Top-level Demand

Equation (7) is estimated with two-stage least squares where the instruments are cost shifters: ingredients, packaging, city density (to proxy for shelf space) and labor cost (supermarket and industry).

V(ii). *Predicted Prices with Higher Excise Taxes*

Marginal costs (12) for the pre-tax-increase period are used to compute each model's predicted equilibrium prices after the tax change (i.e., the first quarter of 1991). Since excise taxes were increased for all beers at a uniform rate of E per unit, predicted prices in each city for quarter $y+1$ are computed by solving for p^{y+1} $(j = 1, \ldots, J)$ in the following system of non-linear equations:

$$q_j(p^{y+1}, A^y) - 1 + \sum_{k \in F_n} (p_k^{y+1} - c_k^y - E)\left[\overline{\frac{\partial q_k}{\partial p_j}} + \frac{\partial q_k}{\partial A_j}\right] = 0, \text{ for } j$$
$$= 1, \ldots, J$$

where the superscript y denotes the quarter prior to the tax increase: fourth quarter of 1990.[18] Because q_j and all derivatives are functions of price (p_j^{y+1}),

[18] To avoid sensitivity to potential outliers in quarter y, the median city-specific marginal cost of brand k over the period 1988–1990 is used for c_k^y. Results, however, are qualitatively the same if only marginal costs for the fourth quarter of 1990 are used.

the search includes these non-linear terms. Other variables (i.e., advertising, distance measures, product characteristics and total expenditures x_t) are held constant at time y values, while demand parameters are those obtained from estimation. The predicted prices are computed for every brand in each of the 46 cities for which data are available.[19]

Results are invariable to whether pre- or post-tax-increase advertising is used in the search. In some cities, a few brands (1 or 2) exited or entered the market between the fourth quarter of 1990 and the first quarter of 1991. In these cities, the search was performed for the subset of brands that were present in both quarters. The potential bias of this simplification is likely to be small as the ignored brands tend to be marginal in terms of sales.

V(iii). *Estimates of Actual Price Increases*

A straightforward way to compute an *average* estimate of the actual price increase across cities (and its confidence interval) is to estimate a separate regression for each brand of the following form (see Hausman and Leonard [2002]):

$$p_{yz} = \theta_z + \eta' I + \xi_{yz} \tag{13}$$

where p_{yz} is price in quarter y and city z (i.e., each city-quarter pair y, z corresponds to a market t), θ_z are city fixed effects, I is a vector of time dummy variables (one for each of 19 quarters -one omitted to avoid collinearity) and η its corresponding vector of coefficients. If the time dummy on the fourth quarter of 1990 is omitted (i.e., this is the reference quarter), the coefficient on the dummy for the first quarter of 1991 can be interpreted as the absolute mean price increase for that brand due to the tax increase (its standard error is used to construct confidence intervals). This coefficient, however, captures the mean effect on price of all city-invariant factors present in the first quarter of 1991 (i.e., other national shocks besides the tax increase). A dummy variable that takes a value of 1 in the first quarter of each year was included in (13) to control for a possible seasonality effect.

VI. RESULTS

VI(i). *Demand*

This section presents the estimation results for brand-level demand (eq. 6) and top-level demand (eq. 7).

[19] This system is solved by using the iterative Newton algorithm for large-scale problems provided by *Matlab*. Convergence is quickly achieved for the Bertrand-Nash and collusive models, but leadership models require several hours of computing power.

Brand-level Demand

Because the functional form of demand constitutes only a local approximation to any unknown demand function, demand parameters can potentially differ between the two regimes (pre- and post-tax-increase). However, aside from slightly larger standard errors, demand estimates with pre-increase data produced results that were essentially the same as those obtained with the full sample. Estimates are therefore robust to these two sample sizes. Demand estimates reported in this section were computed with the full sample.

The regressions below contain variables that consistently had the greatest explanatory power in different specifications. Table (II) reports the GMM regression results for two different models. The difference between models 1 and 2 is the inclusion of brand dummies. The two models contain time and city binary variables (coefficients not reported).

In the intercept, there is only one product-specific variable that varies by market: number of common boundaries in alcohol content-product coverage space (*NCBAC*). The negative coefficient on *NCBAC* shows that brands that share a common boundary with more neighbors in alcohol content-coverage space have a lower sales share.

The estimated coefficients for own-price, own-advertising, and their interactions with product characteristics are reported in the second group of variables in Table II. Because price and advertising are highly correlated with their corresponding interactions with product coverage, the inverse of this latter variable ($1/COV$) is used to avoid collinearity. The own-price and own-advertising coefficients are significantly different from zero at the 1% level and have the expected negative and positive signs. The negative coefficients on the interaction of price and advertising with the inverse of product coverage indicates that as the coverage of a brand increases, the own-price effect for that brand decreases (becomes less negative) while the own-advertising effect increases (becomes more positive). Thus, the sales of brands that are widely sold within a city are less sensitive to a change in price than are brands that are less widely available. Also, advertising is more effective for brands that are more widely sold. Finally, as the number of common boundaries increases, the own-price effect increases (becomes more negative) and the own-advertising effect decreases. This shows that higher brand competition is associated with more price responsive demand and less effective advertising.

Comparing models 1 and 2, the estimated own-price coefficient is nearly twice as large in absolute terms when brand dummies are included. Conversely, the own-advertising coefficient decreased by approximately 80 per cent in model 2. The better goodness-of-fit of model 2 and the magnitude of change on both price and advertising coefficients highlight the importance of accounting for endogeneity, resulting from unobserved

TABLE II
RESULTS OF GMM ESTIMATION OF DEMAND MODEL*

DEPENDENT VARIABLE: SALES SHARE (w_{jt})	Model 1		Model 2	
Variable; Description	Coeff.	(t-stat)**	Coeff.	(t-stat)**
Constant a_{jt}				
Brand Dummies	*no*	*no*	*yes*	*yes*
NCBAC = # common boundary neighbors, Alcohol content – Coverage space	−1.15	(−0.85)	−3.91	(−3.66)
OVER50K	−94.84	(−0.57)	−240.0	(−1.90)
Own Price (b) and Own-Advertising (c)				
logP	−122.40	(−9.82)	−252.90	(−5.71)
logP′ × (1/COV)	−0.56	(−2.38)	−1.09	(−3.46)
logP′ × NCBCSP; NCBCSP = # CB neighbors Coverage-Size-Price space	−4.82	(−7.28)	−7.14	(−11.35)
A^{γ}	8.48	(31.15)	1.32	(4.39)
A' × (1/COV)	−0.68	(−5.58)	−0.19	(−3.47)
A' × NCBCS; NCBCS = # common boundary (CB) neighbors, Coverage-Size space	−1.65	(−3.57)	−0.16	(−4.53)
Weighted Cross Price and Weighted Cross-Advertising Terms (λ_l and τ_m)				
Distance Measures for Price				
Alcohol Content – Product Coverage, two-dimensional product space	2.10	(13.66)	5.32	(11.00)
Nearest neighbors in Alcohol Content – Product Coverage space	−0.21	(−0.30)	8.87	(15.62)
Brewer identity	−12.18	(−5.38)	17.30	(5.31)
Product classification 2: Regular – light	52.39	(6.62)	93.56	(3.99)
National Identity	40.83	(5.85)	49.61	(5.39)
Distance Measures for Advertising				
Container Size, one-dimensional product space	0.17	(7.83)	0.16	(8.64)
Common boundary in product coverage – container size – price space	0.85	(15.50)	0.71	(15.23)
Nearest neighbors in product coverage – container size space	0.61	(14.70)	0.40	(12.24)
Product Classification 3: Budget, light, and premium	−2.78	(−14.58)	−3.22	(−9.10)
National Identity	−3.02	(−21.79)	5.30	(2.65)
Price Index (d)				
$\log(x_t/P_t^L)$	28.15	(1.08)	27.35	(1.38)
R^2 (centered, uncentered)	0.40, 0.58		0.66, 0.76	
J-Statistic (p-value)	0.90		0.50	

*Note: Based on 33,892 observations. Coefficients in table are original coefficients × 10^4. All specifications include, time and city dummies (not reported).
**Asymptotic t-statistics.

product characteristics, with the inclusion of brand dummies. Furthermore, the overidentification test in model 2 (*p-value* = 0.50) suggests that the choice of instruments is valid.

In model 2, the estimated coefficients on the weighted cross-price terms are all positive. Thus, brands that are closer in the alcohol content-product coverage space (both in terms of Euclidean distance and nearest neighbor), produced by the same brewer, belong to the same product segment, or have similar geographic coverage, are stronger substitutes than other brands. Intuitively, consumers will more likely switch to a brand located nearby in product space and/or produced by the same brewer than to more distant brands. Based on the magnitude of the estimated coefficients, the strongest substitution effects are for brands in the same product segment and with similar geographic coverage.

With the exception of product segment, the estimated coefficients on weighted cross-advertising terms are positive. This suggests the existence of cooperative effects across brands that are more closely located in product space and with the same geographic coverage. However, the negative coefficient for product segment indicates that there are predatory advertising effects for brands in the same product segment, thereby potentially offsetting some of the cooperative effects.

The estimated coefficient on real expenditures, $\log(x_t/P_t)$, is not statistically different from zero. Several attempts to interact product or market characteristics with real expenditures yielded statistically insignificant coefficients. This result implies that the brand-level budget elasticities are not statistically different from one.

Top-level Demand

Various sets of instrumental variables and specifications failed to produce statistically significant estimates of the overall price elasticity of beer, β_2 in equation (7). This is a somewhat unlikely result, as prior research has typically found a statistically significant price elasticity for beer. Evidence of this are the troublesome unconditional (brand-level) price elasticities obtained with this estimate of β_2: it is not uncommon to see own-price elasticities that are less than one in absolute value, which typically imply (unfeasible) price-cost margins of over 100%.

Hausman, Leonard and Zona suggest that the use of longer time series can produce more plausible estimates of the overall price elasticity of beer. The data used in this study appear to be contemporaneous to that of Hausman, Leonard and Zona, who use 16 years of data and instrumental variables (similar to those proposed here) to estimate an overall price elasticity for beer of -1.36 (s.e.: 0.21). This is the estimate employed to compute unconditional elasticities and price derivatives.

Elasticities

Unconditional elasticities were computed in each city-quarter pair using the estimates of ((6), model 2 above) and (7). All own-price elasticities are negative and statistically significant, with a median of -3.34.[20] Cross-price elasticities have a median value of 0.050 with 92% of them being positive; none of the negative cross-price elasticities is significant while 96% of the positive cross-price elasticities is significant. In general, median own-price elasticities are slightly smaller to those reported in Hausman, Leonard and Zona (-4.98), and Slade (-4.1). Cross-price elasticities are similar to those in Slade but an order of magnitude smaller than those reported by Hausman, Leonard and Zona. A reason for the larger cross-price elasticities in Hausman, Leonard and Zona is that a significantly smaller number of brands are considered in their study; the magnitude of cross-price elasticities needs to decrease as more brands are added, otherwise the elasticity matrix would cease to be dominant diagonal.

Median own-advertising elasticity is 0.024 with approximately 85% of them being positive. Cross-advertising elasticities have median of 0.021 with 88% of them being positive. However, not all advertising elasticities are significant: 15% of all negative advertising elasticities and 14% of all positive elasticities are not statistically significant. A sample of median price and advertising elasticities and a further discussion are provided in Rojas and Peterson [2008].

VI(ii). *Implied Price-Cost Margins*

For each model, implied marginal costs in the pre-tax-increase period are calculated according to details in section III(iii). Summary statistics of marginal costs can be informative about differences in the equilibrium predictions of the models; however, price-cost margins are more readily interpretable. Pre-tax-increase summary statistics of price-cost margins (PCM) as a percentage of price ($100 \times [p-c]/p$) are presented in Table III. Seven different models are considered: Bertrand-Nash; two Stackelberg scenarios: firm leadership by Anheuser-Busch and brand leadership by Budweiser; collusive leadership by Budweiser; and three collusive scenarios: collusion of the three leading firms (Anheuser-Busch, Coors and Miller), collusion of the leading regular brand produced by each of the three largest firms (Budweiser, Coors and Miller Genuine Draft), and full collusion. In addition, the hypothetical case of single-product Bertrand-Nash competition is also considered.

The mean PCMs in the two Stackelberg cases and in collusion among three brands are very similar to that of Bertrand-Nash. Collusion among three firms has slightly larger PCMs than Bertrand-Nash, and full collusion has the largest PCMs (although there are several cases in which unfeasible price-cost margins (over 100%) are detected). The PCMs under single-

[20] Signficance is determined with 95% confidence intervals (not shown), which were computed using 5,000 draws from the asymptotic distribution of the estimated demand coefficients.

TABLE III
SUMMARY STATISTICS OF PRICE COST MARGINS FOR DIFFERENT MODELS, (1988--1990)*

Model	Mean	Median	St dev
Single-Product Bertrand-Nash	34.63	31.06	25.18
Bertrand-Nash	37.26	35.48	23.73
Anheuser-Busch Stackelberg Leadership	37.47	35.85	23.79
Budweiser Stackelberg Leadership	37.27	35.48	23.74
Collusive Leadership (Budweiser)**	55.03	52.79	23.83
Collusion 3 firms§	39.53	41.77	23.23
Collusion 3 brands ±	37.20	35.72	23.44
Full Collusion	68.01	69.78	11.64

*Margins are defined as $100 \times (p-c)/p$. Based on 18,369 (brand-city-quarter) observations in the pre-increase period (1988–1990).
**Price-cost margins obtained for Anheuser-Busch brands only.
§Anheuser-Busch, Adolph Coors and Miller.
± Budweiser, Coors, Miller Genuine Draft.

product Bertrand-Nash are lower but close to multi-product Bertrand-Nash. This suggests that a large portion of the mark-up implied by multi-product Bertrand-Nash competition can be explained by product differentiation alone, while only a small remainder can be explained by concentration of brand ownership.

Since in the collusive price leadership scenario, PCMs are only computed for Anheuser-Busch brands, summary statistics for this case are not directly comparable with those of other models. However, PCMs are significantly larger for Budweiser (mean 97% vs. 78% in Bertrand-Nash, not shown) and similar to Bertrand-Nash PCMs for other Anheuser-Busch brands (mean 47% in both collusive price leadership and Bertrand-Nash, not shown).

In all models, PCMs vary considerably across brands. This heterogeneity is directly related to the price elasticities and the strategic behavior of firms and thus is plays an important role when comparing each model's predictive power.

One way to identifiy models of competition is to compare implied PCMs with observed PCMs. However, observed PCMs are unavailable. A raw measure of PCM is the gross margin (total shipments minus labor and materials) calculated from the Annual Survey of Manufacturers (as in Nevo [2001]). The average gross margin for the U.S. brewing industry in the pre-tax increase period (1988–1990) is 44.53% (27.5% for all food industries), which is somewhat larger than what is predicted by the models, except for collusive leadership and full collusion. The next section provides brand-level closeness measures between the observed prices and the prices predicted by the different models during the post-tax-increase period.

VI(iii). *Predicted vs. Actual Price Increases*

Here absolute price increases $\left(p_j^{y+1} - p_j^y\right)$ are compared to estimates of observed or 'actual' price increases (see sections V(ii) and V(iii)). A graphical

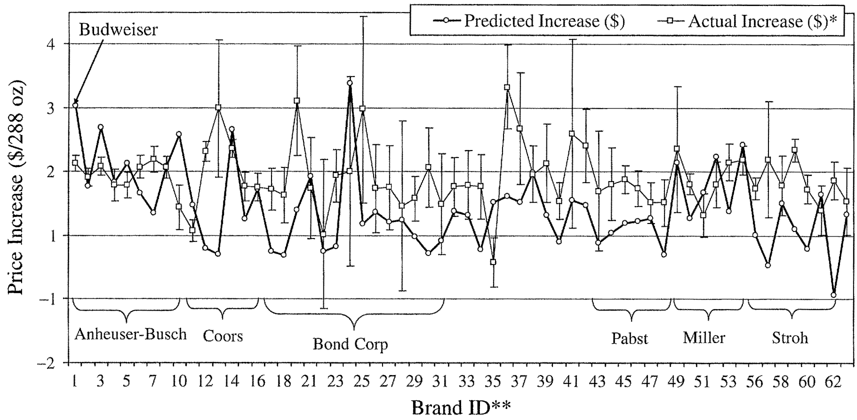

Figure 3
Predicted price increases by Bertrand-Nash behavior vs. actual price increases per brand after 100% hike in the federal excise tax (mean over 46 cities)
*Vertical Lines are 95% Confidence Intervals
**See Table V for Brand ID's

assessment of each brand's mean predicted price increase across cities revealed that full collusion and collusive price leadership produce unlikely high price increases (10 times or more the amount of the actual price increase) whereas the remaining models (including single-product Bertrand-Nash) yield predicted increases that are not substantially different from each other. For this reason, and because of space constraints, only predicted increases of Bertrand-Nash (the baseline case) are presented graphically for each brand. Figure 3 plots the means of the price increases as predicted by Bertrand-Nash as well as the means of the actual price increases. Mean predicted increases are averages across 46 cities while mean actual increases are computed according to details in section V(iii). 95% confidence intervals are displayed for the mean of actual price increases.[21]

There are several patterns in Figure 3. Price increases tend to be under-predicted (44 out of 63 brands).[22] Also, over-predicted prices appear to be more frequent among the two largest beer producers: Anheuser-Busch (7 out of 10) and Miller (4 out of 7). This is because these two firms produce the

[21] The non-linear systems for predicted price increases require several hours of computing time. Calculating confidence intervals for predicted mean price increases with a bootstrapping technique are hence extremely costly even with a modest number of draws.

[22] A similar number of under-predictions are detected for the other models not shown in the figure (except for full collusion and collusive price leadership in which all price increases are over-predicted).

TABLE IV
SUMMARY STATISTICS OF ACTUAL AND PREDICTED PRICE INCREASES, AND PERFORMANCE METRICS OF MODELS

	Summary Statistics[a]			Performance Metrics		
	Mean	Median	St. Dev.	# No-Reject[b]	Weighted Increase[c]	SSD[d]
Actual Increases	1.38	1.37	0.65	N/A	64.74	N/A
Predicted Increases						
Single-Product Bertrand-Nash	0.95	0.81	1.07	20	89.74	3587
Bertrand-Nash	1.02	0.90	1.02	21	86.91	3643
A-B Stackelberg Leader[e]	1.00	0.92	0.98	21	76.80	3441
Budweiser Stackelberg Leader	0.99	0.89	0.99	21	78.97	3442
Collusion 3 firms[f]	1.21	1.11	1.01	22	87.48	4328
Collusion 3 brands[g]	1.03	0.93	1.00	23	84.45	3621
Full Collusion	18.19	13.40	15.88	0	1014.95	> 1E6

[a]Computed with absolute price increases for each brand: the absolute price difference between the first quarter of 1991 and the fourth quarter of 1990, over 46 cities (1748 observations).
[b]Number of brands for which mean of predicted increases falls within the confidence intervals of mean of actual increases (as in Figure 3).
[c]Sum of weighted absolute price increases; weight = volume of brand sold in city/total volume of all brands in all cities in the first quarter of 1990 (1748 observations).
[d]Sum of squared deviations over all brands and all cities; deviation = predicted-actual (1748 obs.).
[e]A-B = Anheuser-Busch.
[f]Anheuser-Busch, Adolph Coors, Miller (Philip Morris).
[g]Budweiser, Coors, Miller Genuine Draft.

more price-inelastic brands and price elasticity is inversely associated with higher tax pass-through rates. Many brands have tight 95% confidence intervals around actual mean increases (around 15¢ and 20¢), indicating that price increases do not vary substantially across cities. This pattern can particularly be observed for brewers that tend to produce nationally: Anheuser-Busch, Coors, Pabst, Miller and Stroh.

Since graphical assessment is not very informative about the relative predictive power of the models, summary statistics and performance metrics are analyzed next. The left part of Table IV presents summary statistics of price increases (i.e., the absolute difference in prices between the two quarters).[23] Except for full collusion, the mean and median of predicted increases are similar across models and smaller than those of actual price increases; this is a consequence of most prices being under-predicted. The closeness in statistics between single-product Bertrand-Nash and multi-product Bertrand-Nash is consistent with the observation in section VI(ii) that a large portion of mark-ups may be explained by product differentiation alone, and a small portion by brand ownership concentration.

[23] Collusive price leadership is omitted from the table because convergence in the search for the predicted prices was not always achieved. When convergence was achieved, price increases were unlikely high. The lack of convergence is interpreted as further support for the improbable nature of this model.

The summary statistics of Table IV would suggest that collusion among 3 firms appears to be a better predictor of actual price increases. However, closer graphical inspection (discussed below) indicates that this is due to larger over-predictions for Anheuser-Busch's and Miller brands rather than by smaller under-predictions of other brands. Larger standard deviations of predicted increases with respect to standard deviations of actual price increases indicates less variability in actual price increases. The full collusion model can confidently be rejected given its unlikely predicted increases.

The right part of Table IV presents three performance metrics. The first metric (# Non-Rejections) is the number of brands for which predicted mean price increases fall within the confidence intervals of actual mean price increases shown in Figure 3. According to this metric, collusion among 3 brands explains firm conduct more precisely than the other models.

Two more rigorous metrics are considered. The first is the sum of weighted price increases, where the weight is given by each brand's market share. With this metric, accuracy in prediction is more important for more widely sold brands. Interestingly, with this criterion Stackelberg leadership by Anheuser-Busch, closely followed by Stackelberg leadership by Budweiser, outperforms other models. The second metric is the sum of squared deviations, where a deviation is defined as the difference between the predicted increase and the actual increase.[24] This criterion confirms that the two Stackelberg models may be better predictors of actual conduct.

To further understand why the last two metrics indicate a superior performance by Stackelberg leadership models, the differences between predicted increases and actual increases for a selected group of brands and models are analyzed. Figure 4 reports such differences for three models: Bertrand-Nash, Stackelberg leadership by Anheuser-Busch and 3-firm collusion. Stackelberg leadership by Budweiser and 3-brand collusion are not included because they can not be distinguished visually from Anheuser-Busch Stackelberg leadership (A-B leads) and Bertrand-Nash (B-N), respectively.[25] The selected brands are those for which the notable differences across models were observed; these brands belong to the three largest firms (A-B, Coors, Miller).

Proximity to zero in Figure 4 denotes greater accuracy; a positive number denotes over-prediction and a negative one under-prediction. Compared with Bertrand-Nash, it can be seen that A-B Stackelberg is more accurate for several A-B brands (both over-predicted and under-predicted), especially Budweiser, the most popular brand in the U.S. The 3-firm collusion case almost always predicts larger price increases than Bertrand-Nash and hence it only represents an improvement to Bertrand-Nash in under-prediction cases.

[24] The same conclusion is reached if each deviation is weighted.

[25] Single-product Bertrand-Nash is also excluded because of its closeness to multi-product Bertrand-Nash.

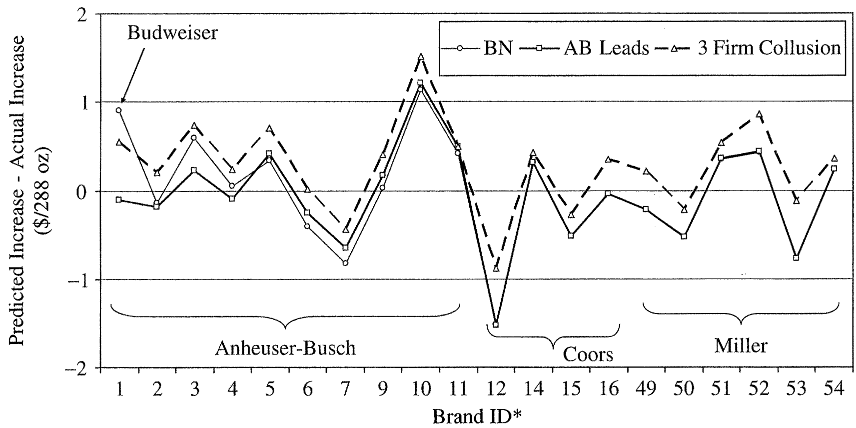

Figure 4
(Predicted price increase)—(actual price increases), after 100% hike in the Federal excise tax (mean over 46 cities)
*SeeTable B.I for Brand ID's

VII. CONCLUSION

The study combines a rich brand-level data set, recent demand estimation techniques and a unique natural experiment (a large increase in excise tax) to evaluate alternative models of firms' pricing conduct in the U.S. brewing industry. The strategy is to focus on the period when the tax increase became effective, January of 1991, and compare the observed prices with those predicted by different models of price competition.

Bertrand-Nash, and several variants of leadership and collusion are considered as possible pricing models. There are two cases of Stackelberg leadership, one where the largest firm, Anheuser-Busch, leads with its entire product line and another where it leads with its flagship brand Budweiser. Collusive scenarios consider the three largest firms, the three leading regular brands of beer and collusion among all brands. A case of collusive leadership is also considered, where all brands match Budweiser's price increases.

Several metrics of closeness between predicted price increases and observed price increases indicate that collusion among all brands and collusive price leadership can be confidently rejected as plausible models of firm conduct. Among the remaining models, both Stackelberg leadership variants appear to be slightly better predictors of firm behavior. However, this evidence is interpreted with caution as competing models' predicted increases are not largely dissimilar and it is likely that the better fit might not be statistically significant to warrant such a conclusion. The results are somewhat in line to those of Nevo [2001] and Slade [2004], who reject full collusion in favor of Bertrand-Nash. Results also indicate that single-product Bertrand-Nash predicts price increases that are similar to those of multi-product Bertrand-Nash, which suggests that mark-ups are mostly driven by product differentiation while brand ownership concentration appears to have a small role.

One possibility that emerges from this study is the existence of observational equivalence between several models of competition. The reason for this may be that in a complex environment with many strategic interactions, it is more difficult to achieve straightforward comparative static results than in the textbook duopoly case, where, for example, Stackelberg's equilibrium price is substantially larger than Bertrand-Nash's.

As in previous work, the inference conducted in this paper depends crucially on the precision of demand estimates. The distance metric method employed here is effective in reducing the number of cross-price and cross-price effects, but it relies heavily in the researcher's ability to have data on all product characteristics that effectively determine substitution effects. Results may change if there are important unobserved product characteristics.

There are several advantages and disadvantages to the approach used in this paper. Since the effective change in marginal cost for all beer producers is known, the comparison between models may be potentially more reliable than contrasting crude measures of observed price-cost margins with the implied price-cost margins. An additional advantage is the relative simplicity with which comparisons across models can be made. However, the approach employed here can not be generalized because natural experiments are not always present. Another potential drawback is that clear statistical comparisons, as in non-nested tests, may not be feasible.

There are several issues that this study does not address. First, because of the time aggregation of data (quarterly), it is not possible to study in more detail firms' price adjustment decisions which would prove useful in determining Anheuser-Buschs leadership and in analyzing the stock effects of advertising. The analysis here takes the form of a one-shot game. A dynamic environment, however, is important when future profits are not independent of the current state thereby making the static solution suboptimal. For example, it is likely that more successful firms like Anheuser-Busch have a longer horizon in mind (and hence a potentially different price than the static solution) than firms that are under financial stress like Stroh, Heilman and Pabst. Finally, detailed cost data at the manufacturer and retailer level can allow to extend the analysis to vertical aspects and also more rigorous econometric tests of the competing pricing models considered in this paper.

APPENDIX A: SUPPLY DETAILS

Derivation of $\frac{dp_m}{dp_j}$

Define a partition of the product set as $\psi = (\psi_f, \psi_l)$, where ψ_f is the set of follower brands and ψ_l is the set of leading brands, with J^F and J^L number of elements respectively. For each leader, a system of equations is constructed. Each l^{th} system of equations is used to compute the vector of all $\frac{dp_m}{dp_l}$ terms for leader l. An equation in system l is obtained by totally differentiating the price first order condition of all follower brands (9)[26] with respect to all followers' prices (p_f, for all $f \in \psi_f$) and the price of the l^{th} leader, p_l ($l \in \psi_l$):

$$(14)\quad \sum_{f \in \Psi_f} \underbrace{\left[\frac{\partial q_j}{\partial p_f} + \sum_{k \in \Psi_f}\left(\Delta^*_{kj}(p_k - c_k)\frac{\partial^2 q_k}{\partial p_j \partial p_f}\right) + \Delta^*_{fj}\frac{\partial q_f}{\partial p_j}\right]}_{g(j.m)} dp_f + \underbrace{\left[\frac{\partial q_j}{\partial p_l} + \sum_{k \in \Psi_f}\left(\Delta^*_{kj}(p_k - c_k)\frac{\partial^2 q_k}{\partial p_j \partial p_l}\right)\right]}_{h(j,m)} dp_l = 0;\ j,\ k, f \in \Psi_f$$

where Δ^*_{jk} takes the value of one if brands j and k are produced by the same firm and zero otherwise. Therefore, for a given leader l, there are J^F equations like (14). Let G be the $(J^F \times J^F)$ matrix that contains all g elements above and define the $(J^F \times 1)$ vectors D_s and H_l as:

$$D_s = \begin{bmatrix} dp_1 \\ \cdot \\ \cdot \\ \cdot \\ dp_{J^F} \end{bmatrix};\ H_l = \begin{bmatrix} -h(1,\ l) \\ \cdot \\ \cdot \\ \cdot \\ -h(J^F,\ l) \end{bmatrix}$$

For a given p_l, (14) is written in matrix notation as:

$$GD_s - H_l dp_l = 0$$

where dp_l is treated as a scalar for matrix operations. The J^F derivatives of the followers' prices with respect to a given p_l are computed as:

$$(15)\quad \frac{D_s}{dp_l} = G^{-1}H_l$$

Concatenating the $(J - J^F)$ vectors of dimension $(J^F \times 1)$ given in (15) (one vector for each p_l) gives $D = G^{-1}H$. The $J^F \times J^L$ matrix D has a typical element $\frac{dp_f}{dp_l}$, for $f \in \psi_f$ and $l \in \psi_l$.

[26] It is assumed that the first order condition with respect to advertising (10) does not play a role in deriving $\frac{\partial p_m}{\partial p_l}$. Without this assumption, inversion of matrix G below is not possible since it is not a square matrix. Results are unlikely to be sensitive to this assumption given the estimated small impact advertising has on demand.

Marginal Costs in Leadership Models

Stackelberg Model

While marginal costs are obtained by applying (12), the derivative $\frac{dp_m}{dp_j}$ needs to be computed first via equation (15). Several technical difficulties arise in this model. First, there is a large number of possible Stackelberg scenarios. Given the motivation in this paper, only the case in which Anheuser-Busch acts as a leader, both with all its brands as well as with Budweiser, are considered.

Second, since the term $\frac{dp_m}{dp_j}$ in the leaders' first order conditions is a function of followers' marginal costs (see equation (15)), these marginal costs are computed first. When Anheuser-Busch acts as a leader with all its brands, followers' marginal costs can be obtained by inversion of a smaller system of dimension J^F in (12). These marginal costs are used to compute $\frac{dp_m}{dp_j}$, which is afterwards used to calculate the marginal costs of the leading brands.

When Budweiser is a sole brand leader, the term $\frac{dp_m}{dp_j}$ is set to zero if m is produced by Anheuser-Busch, except for the brand Budweiser. Also, it is assumed that Budweiser only leads brands produced by rival firms (i.e., not by Anheuser-Busch).

Collusive Price Leadership Model

In this case, only Anheuser-Busch's marginal costs can be derived since first order conditions of other firms are not relevant (see section III(ii)). These marginal costs are also recovered by applying (12) to a system of dimension J^L (where J^L is the number of brands sold by Anheuser-Busch) and by setting $\frac{dp_m}{dp_j}$ to 1 in Budweiser's first order condition and zero in the remaining first order conditions.

APPENDIX B: DATA DETAILS

IRI is a Chicago based marketing firm that collects scanner data from a large sample of supermarkets that is drawn from a universe of stores with annual sales of more than 2 million dollars. This universe accounts for 82% of all grocery sales in the U.S. In most cities, the sample of supermarkets covers more than 20% of the relevant population. In addition, IRI data correlates well with private sources in the Brewing Industry (the correlation coefficient of market shares for the top 10 brands between data from IRI and data from the Modern Brewery Age Blue Book is 0.95). Brands that had at least a 3% local market share in any given city were selected. After selecting brands according to this criterion, remaining observations are dropped if they had a local market share of less than 0.025%. Brands that appear in less than 10 quarters are also dropped. Also, if a brand appears only in one city in a given quarter, the observation for that quarter is not included either. This is done because some variables in other cities are used as instruments. On average there are 37 brands sold in each city market with a minimum of 24 brands and a maximum of 48 brands. Table B.I contains a list of the brands used with information on country of origin and the corresponding brewers.

The original data set contained observations in 63 cities; five cities were dropped because of minimal number of brands or quantities. Overall, the number of cities increases over time; however, some cities appear only in a few quarters in the middle of

TABLE B.I
SELECTED BRANDS BY BREWER (ACRONYM AND COUNTRY OF ORIGIN)

Brewer	Brand	Brewer	Brand
Anheuser-Busch:	Budweiser	Grupo Modelo:	Corona
(AB, U.S.)	Bud Dry	(GM, Mexico)	
	Bud Light	Goya (GO, U.S.):	Goya
	Busch	Heineken:	Heineken
	Busch Light	(H, Netherlands)	
	Michelob	Labatt:	Labatt
	Michelob Dry	(LB, Canada)	Labatt Blue
	Michelob Golden Draft		Rolling Rock
	Michelob Light	Molson:	Molson
	Natural Light	(M, Canada)	Molson Golden
	Odoul's		Old Vienna
Adolph Coors:	Coors	Pabst:	Falstaff
(ADC, U.S.)	Coors Extra Gold	(P, U.S.)	Hamms
	Coors Light		Hamms Light
	Keystone		Olympia
	Keystone Light		Pabst Blue Ribbon
Bond Corp[a]:	Black Label		Red White & Blue
(B, U.S.)	Blatz	Miller/Phillip Morris:	Genuine Draft
	Heidelberg	(PM, U.S.)	Meister Brau
	Henry Weinhard Ale		Meister Brau Light
	Henry Weinhard P. R.		MGD Light
	Kingsbury		Miller High Life
	Lone Star		Miller Lite
	Lone Star Light		Milwaukee's Best
	Old Style	Stroh:	Goebel
	Old Style Light	(S, U.S.)	Old Milwaukee
	Rainier		Old Milwaukee Light
	Schmidts		Piels
	Sterling		Schaefer
	Weidemann		Schlitz
	White Stag		Stroh
Genesee:	Genesee	FX Matts:	Matts
(GE, U.S.)	Kochs	(W, U.S.)	Utica Club

[a]These brands correspond to G. Hieleman Brewing Co., which was acquired in 1987 by Australian Bond Corporation Holdings; it is classified as a domestic brewer because this foreign ownership was temporary.

the period. The average number of cities per quarter is 47. Brands are identified as regional or national as follows. First the percentage of cities in which each brand was present was averaged over time. Brands with an average percentage close to 100 are denoted national and brands with a percentage of (roughly) 50% or less are denoted regional. The variable *WAGES* was constructed by averaging the hourly wages of interviewed individuals from the Bureau of Labor Statistics CPS monthly earning files at the NBER. For a given city-quarter combination, individuals working in the retail sector were selected for that city over the corresponding three months. The average was then calculated over the number of individuals selected.

REFERENCES

Anderson, S.; de Palma, A. and Kreider, B., 2001, 'Tax Incidence in Differentiated Product; Oligopoly,' *Journal of Public Economics*, 81, pp. 173–192.

Anheuser-Busch, 2003, *Annual Report, available at http://www.anheuser-busch.com/annual/2003/Domestic.pdf*

Census Bureau, *Annual Survey of Manufacturers* various years.

Berry, S., 1994, 'Estimating Discrete-Choice Models of Product Differentiation,' *RAND Journal of Economics*, 25, pp. 242–262.

Bresnahan, T., 1989, 'Empirical Methods for Industries with Market Power,' in Schmalensee, R. and Willig, R. (eds.) *Handbook of Industrial Organization, Vol. II* (North-Holland, Amsterdam).

Bureau of Labor Statistics, *Current Population Survey Basic Monthly Files*, various issues, available at the National Bureau of Economic Research: *http://www.nber.org/data/cps_index.html*

Corts, K., 1999, 'Conduct Parameters and the Measurement of Market Power,' *Journal of Econometrics*, pp. 227–250.

Deaton, A. and Muellbauer, J., 1980, 'An Almost Ideal Demand System,' *American Economic Review*, 70, pp. 312–326.

Demographia. www.demographia.com

Gasmi, F.; Laffont, J. J. and Vuong, Q., 1992, 'Econometric Analysis of Collusive Behavior in a Soft-Drink Market,' *Journal of Economics and Management Strategy*, 1, pp. 277–311.

Greer, D., 1998, 'Beer: Causes of Structural Change,' in Duetsch, L. (ed.), *Industry Studies* (Prentice Hall, New Jersey).

Hausman, J. and Leonard, G., 2002, 'The Competitive Effects of a New Product Introduction: A Case Study,' *Journal of Industrial Economics*, 50, pp. 237–263.

Hausman, J.; Leonard, G. and Zona, D., 1994, 'Competitive Analysis with Differentiated; Products,' *Annales d'Économie et de Statistique*, 34, pp. 159–180.

Kadiyali, V.; Vilcassim, N. and Chintagunta, P., 1996, 'Empirical Analysis of Competitive; Product Line Pricing Decisions: Lead, Follow, or Move Together?,' *Journal of Business*, 69, pp. 459–487.

Moschini, G., 1995, 'Units of Measurement and the Stone Index in Demand System Estimation,' *American Journal of Agricultural Economics*, 77, pp. 63–68.

Nevo, A., 1998, 'Identification of the Oligopoly Solution Concept in a Differentiated Products Industry,' *Economics Letters*, 59, pp. 391–395.

Nevo, A., 2001, 'Measuring Market Power in the Ready-to-Eat Cereal Industry,' *Econometrica*, 69, pp. 307–342.

Pinkse, J.; Slade, M. and Brett, C., 2002, 'Spatial Price Competition: A Semiparametric; Approach,' *Econometrica*, 70, pp. 1111–1155.

Pinkse, J. and Slade, M., 2004, 'Mergers, Brand Competition, and the Price of a Pint,' *European Economic Review*, 48, pp. 617–643.

Rojas, C., 2005, *Demand Estimation with Differentiated Products: An Application to Price Competition in the U.S. Brewing Industry*, Ph.D. dissertation, Virginia Tech.

Rojas, C. and Peterson, E., 2008, 'Demand for Differentiated Products: Price and Advertising Evidence from the U.S. Beer Market,' *International Journal of Industrial Organisation*, 26, pp. 288–307.

Slade, M., 2004, 'Market Power and Joint Dominance in U.K. Brewing,' *Journal of Industrial Economics*, 52, pp. 133–163.

Sullivan, D., 1985, 'Testing Hypotheses about Firm Behavior in the Cigarette Industry,' *Journal of Political Economy*, 93, pp. 586–598.

Sumner, D., 1981, 'Measurement of Monopoly Behavior: An Application to the Cigarette Industry,' *Journal of Political Economy*, 89, pp. 1010–1019.

Tremblay, V. and Tremblay, C., 1995, 'Advertising, Price, and Welfare: Evidence from the U.S. Brewing Industry,' *Southern Economic Journal*, 62, pp. 367–381.

Tremblay, V. and Tremblay, C., 2005, *The U.S. Brewing Industry: Data and Economic Analysis*, (MIT Press, Cambridge, Massachusetts).

U.S. Census Bureau, *Concentration Ratios in Manufacturing*, 1997 Economic Census. *http://www.census.gov/prod/ec97/m31s-cr.pdf*

Villas-Boas, S., 2007, 'Vertical Relationships Between Manufacturers and Retailers: Inference With Limited Data,' *The Review of Economic Studies*, 47, pp. 625–652.