The Market Power Effects of a Merger: Evidence from the U.S. Brewing Industry*

Nathan H. Miller[†]
Georgetown University

Matthew C. Weinberg[‡]
Drexel University

July 25, 2016

Abstract

We provide an empirical analysis of the Miller/Coors merger in the U.S. brewing industry. We document an abrupt increase in retail prices just after the merger for MillerCoors and its closest competitor, Anheuser Busch (ABI), both in absolute terms and relative to imported brands. We reject the hypothesis that the price increases can be explained by movement from one Nash-Bertrand equilibrium to another. The results are consistent with the merger facilitating coordination between MillerCoors and ABI. The results also support substantial marginal cost efficiencies, that consumer surplus loss arises due to post-merger coordination, and that total surplus increases.

Keywords: market power; mergers; unilateral effects; coordinated effects; antitrust policy; merger enforcement; brewing industry
JEL classification: K21; L13; L41; L66

*This paper previously circulated under the title "Mergers Facilitate Tacit Collusion: Evidence from the U.S. Brewing Industry." We thank Jonathan Baker, Dan Hosken, Robert Porter, Ted Rosenbaum, Chuck Romeo, Gloria Sheu, Jonathan Williams, and seminar participants at the Board of Governors of the Federal Reserve System, Clemson University, Cornell University, Drexel University, Georgetown University, Indiana University (Kelley School of Business), University of Colorado - Boulder, University of Illinois - Urbana-Champaign, University of Kentucky, University of Notre Dame, University of Oklahoma, University of Pennsylvania (Wharton School of Business), University of Texas - Austin, and Yale University for helpful comments. We have benefited from conversations with John Asker, Orley Ashenfelter, Michael Baye, Hank Farber, Allan Collard-Wexler, Michael Grubb, Ali Hortaçsu, J.F. Houde, Aviv Nevo, and Mike Waldman, among others. Conor Ryan provided research assistance. All estimates and analyses in this paper based on SymphonyIRI Group, Incorporated data are by the authors and not by SymphonyIRI Group, Inc.

†Georgetown University, McDonough School of Business, 37th and O Streets NW, Washington DC 20057. Email: nathan.miller@georgetown.edu.

‡Drexel University, Gerri C. LeBow Hall, 3220 Market Street, Philadelphia PA 19104. Email: mcw325@drexel.edu.

1 Introduction

Theory indicates that repeated interaction within oligopolies can support collusive equilibria if there are few enough firms (e.g., Friedman (1971), Abreu (1988)). Consistent with this, the Merger Guidelines of the United States Department of Justice (DOJ) and Federal Trade Commission (FTC) emphasize that mergers in concentrated markets can facilitate coordination among the remaining competitors, and almost 60 percent of the merger Complaints filed by the DOJ and FTC over 1990-2014 allege coordinated effects (Gilbert and Greene (2015)). The empirical literature of industrial organization, by contrast, focuses much more on the unilateral effects of mergers that arise from the internalization of competition between the two merging firms. Often coordinated effects are ruled out implicitly through an assumption of one-shot Nash-Bertrand competition both before and after the merger (e.g., Berry and Pakes (1993); Hausman, Leonard and Zona (1994); Werden and Froeb (1994); Nevo (2000a)).

We study the economic effects of MillerCoors, a joint venture between SABMiller PLC and Molson Coors Brewing that combines the operations of these brewers in the United States. The joint venture underwent antitrust review as a merger between the second- and third-largest firms in the U.S. brewing industry. It was approved on June 5, 2008 by the DOJ on the basis that merger-specific cost reductions likely would outweigh any anticompetitive effects. Normal course documents of Anheuser-Busch Inbev (ABI) – the closest competitor of MillerCoors – released publicly in later antitrust litigation indicate that the goals of the company's pricing practices include "yielding the highest level of followership in the short-term" and "improving competitor conduct over the long-term." While business documents must be interpreted carefully, this language raises the question of whether the Miller/Coors merger helped facilitate coordination in the brewing industry.

We start with a descriptive analysis of retail price data that span 39 geographic regions over the period 2001-2011. Inflation-adjusted prices are stable around a small downward trend over the seven years preceding the merger. The prices of MillerCoors and ABI then increase abruptly in the Fall of 2008, just after the Miller/Coors merger, and these higher prices persist through the end of the sample. Descriptive regressions place the magnitude of the price increases at roughly six percent, both in absolute terms and relative to import prices, controlling for changes in the macroeconomic environment. Prices increase somewhat more in regions where the merger has larger effects on market concentration, and somewhat less in regions where merger-specific cost reductions likely are larger. These modifying considerations nearly offset on average, leaving unexplained a sudden four percent increase

in MillerCoors and ABI prices that is common across all regions.

To explore the possibility of coordinated effects, we estimate a structural model of demand and supply that allows for post-merger departures from Nash-Bertrand competition. The supply-side of the model incorporates a parameter that determines the extent to which MillerCoors and ABI internalize their pricing externality during the post-merger periods. The other competitors in the model – Modelo and Heineken – are assumed to compete a la Bertrand both against each other and against ABI/MillerCoors. This demarcation between the domestic and import competitors is supported both in the data and in qualitative evidence that we summarize later in the paper. The demand-side of the model is standard. We use a discrete choice random utility model that allows for the estimation of reasonable consumer substitution patterns with aggregated data on market shares and prices. Similar models have been applied to the beer industry (e.g., Hellerstein (2008); Goldberg and Hellerstein (2013); Asker (2016); Romeo (2016); Sweeting and Tao (2016)).

The model rejects Nash-Bertrand competition if the post-merger prices of ABI exceed what can be explained by unilateral effects. The focus on ABI allows us to flexibly capture merger-specific cost reductions for MillerCoors. Because even large increases in ABI's prices could be rationalized by some unobserved shock, restrictions must be placed on the structural error terms. The key identifying assumption is that changes in ABI's unobserved demand and costs, before versus after the merger, are not systematically different from changes in the unobserved demand and costs of Modelo and Heineken. We interpret the merger itself as a plausibly exogenous shifter of the competitive environment (Berry and Haile (2014)) and use this assumption to form moment conditions. The strategy benefits from the presence of competitors outside of the coordinating group, which allows us to control for unobserved demand and cost changes that are common across firms.

The results are consistent with the Miller/Coors merger having coordinated effects. The governing supply-side parameter is statistically different than zero, and robust across a number of modeling choices. The model thus rejects Nash-Bertrand competition in the post-merger periods. Strictly interpreted, the point estimate on our preferred specification implies that ABI and MillerCoors internalize 26 percent of their pricing externalities after the merger. Using counterfactual simulations, we determine that the observed post-merger prices of these firms are 6-8 percent higher than they would have been under Nash-Bertrand competition, and markups are 17-18 percent higher. We develop a number of additional results, including: (i) merger-specific cost reductions are large and roughly counter-balance unilateral effects; (ii) consumer surplus loss is due to post-merger coordination; and (iii) the merger increases total surplus due to the magnitude of marginal cost reductions.

Our analysis is subject to a number of caveats and limitations. We highlight two here. First, the empirical results support post-merger coordination but do not provide direct evidence that the merger enabled this coordination. One alternative explanation is that the recession enabled a ABI/MillerCoors coalition to support higher prices (Rotemberg and Saloner (1986)). Some empirical evidence cuts against this possibility – for example, the price increases of ABI and MillerCoors persist during the economic recovery – but it is difficult to rule out entirely. Second, we could wrongly attribute price increases to coordination if the model understates ABI price increases under Nash-Bertrand competition. This motivates a bevy of robustness analyses that we present in the body of the paper and in a lengthy appendix. Nonetheless, some alternative theories could explain the data. If price matching behavior arises in equilibrium due to some set of retailer beliefs, then a greater share of the unilateral effects likely would fall onto the ABI prices. Another possibility is that firms have uncertainty about each others' costs and play a dynamic signaling game (Sweeting and Tao (2016)).[1] It is impossible to rule out all alternatives, although the available documentary evidence provides some support for the coordinated effects interpretation.

Our research relates to the industrial organization literature in a number of areas. The methodology is closest to Ciliberto and Williams (2014), which models the airline industry as a differentiated-products pricing game. There deviations from Nash-Bertrand competition are proportional to the amount of multi-route contact between carriers. Inference depends on whether prices are higher on routes that feature more multi-route contact, relative to what would arise in Nash-Bertrand equilibrium. Our identification strategy also is similar to Porter (1983) and Igami (2015), which focus on regime shifts in markets with homogeneous products. The former article examines the Joint Executive Committee in the 1880s railroad industry, and identifies reversions between the Nash-Cournot and collusive output levels. The latter article examines a cartel in the coffee bean industry. Inferences are made based on the magnitude of the price decreases that occur after the cartel collapses, under the assumption that post-cartel competition is Nash-Cournot.

A growing number of articles examine the *ex post* effects of mergers (see Ashenfelter, Hosken and Weinberg (2014) for a survey). Most commonly these "merger retrospectives" employ program evaluation techniques to estimate price changes.[2] A handful compare these

[1] Sweeting and Tao (2016) show that if prices are strategic complements then each firm has an incentive to shade prices upward in order to signal high costs. This incentive is weak when there are many firms, but can be strong with few firms. An implication of the model is that consolidation can amplify the extent to which signaling raises prices above Nash-Bertrand equilibrium.

[2] One such paper provides evidence on how prices changed across different geographic markets after the Miller/Coors merger (e.g., Ashenfelter, Hosken and Weinberg (2015)). The article explains variation in price

changes to the predictions obtained under the assumption of Nash-Bertrand competition before and after the merger (e.g., Peters (2006); Weinberg and Hosken (2013); Houde (2012); Bjornerstedt and Verboven (2015)). One recent working paper examines a merger in the ready-to-eat cereal industry and seeks to identify departures from Nash-Bertrand competition (Michel (2016)). The analysis of whether mergers lead to coordinated effects remains novel in this literature, and could help account for discrepancies between the predictions of merger simulation and how prices actually change after mergers.

Finally, our paper relates to articles that measure market power using price and quantity data and at most incomplete data on costs. Because our test for post-merger coordination is based on whether changes in ABI prices can be explained with unilateral effects, it is crucial that the strategic complementarity of MillerCoors and ABI prices is captured in a reasonable manner. For this reason, we follow Nevo (2001) and implement the test after estimating a random utility model of demand that allows for flexible consumer substitution patterns. Further, as in Bresnahan (1987) and Nevo (2001), we assess the plausibility of different models of competition by examining their implied unit costs of production. If we impose Nash-Bertrand competition after the merger, taking into account the unilateral effects, a 13 percent increase in ABI costs are needed to rationalize the data. Publicly-available company documents do not provide support for such a cost increase.

The rest of the paper is organized as follows. Section 2 provides background information on the U.S. brewing industry and the datasets used in the analysis. Section 3 examines variation in changes in retail prices before and after the merger, and summarizes a body of qualitative evidence regarding coordination between MillerCoors and ABI. Section 4 presents several sets of demand estimates, which are key inputs into the model of competition. The supply-side model of brewer competition is presented in Section 5, and Section 6 quantifies the economic importance of deviations from Nash-Bertrand competition using counterfactual simulations. Section 7 discusses some distinctive features of the U.S. brewing industry that may have led the merger to soften competition beyond what can be explained with unilateral effects. Robustness analysis and extensions are available in an online appendix.

---

changes across regions using how the merger would increase local market concentration and reduce shipping distances. It finds a negative relationship between prices and shipping distances and a positive relationship between prices and concentration. The same empirical patterns exist in our data.

2 Industry Background

2.1 Market structure

As do most firms in branded consumer product industries, brewers compete in prices, new product introductions, advertising and periodic sales. The beer industry differs from typical retail consumer product industries in its vertical structure because of state laws regulating the sales and distribution of alcohol. Large brewers are prohibited from selling beer directly to retail outlets. Instead, they typically sell to state-licensed distributors, who in turn sell to retailers. Payments along the supply-chain cannot include slotting fees, slotting allowances, or other fixed payments between firms.[3] While retail price maintenance is technically illegal in many states, in practice distributors are often induced to sell at wholesale prices set by brewers (Asker (2016)).

Table 1 shows revenue-based market shares at two-year intervals over 2001-2011, based on retail scanner data that we describe later in this section. The brands of five brewers – ABI, SABMiller, Molson Coors, Grupo Modelo, and Heineken – account for approximately 80% of total retail revenue, and there is no obvious downward trend in this revenue share despite the recent growth of microbreweries. ABI accounts for about 35 percent of retail revenue, and MillerCoors account for around 30%. Modelo and Heineken, both importers, together account for about 15% of revenues. The national HHIs are in the range that characterizes "moderately concentrated" markets in the DOJ/FTC Merger Guidelines. However, many regions exhibit greater concentration due to heterogeneity in supply and demand conditions. For example, 23 of the 39 regions in our sample have HHIs above 2,500 in the year 2011, which is in the range that characterizes "highly concentrated" markets in the Guidelines.

Consolidation in the industry has continued since the Miller/Coors merger. ABI acquired Modelo in 2013, after our sample period. The DOJ obtained a settlement in which the rights to produce and distribute Modelo brands in the U.S. were divested to Constellation, a large distributor and producer of wine and spirits. More recently, ABI is acquiring SAB-Miller in a deal worth $106 billion. The merging firms are expected to divest SABMiller's stake in the MillerCoors joint venture to Molson Coors as part of a settlement package.

[3] The relevant statutes are the Alcoholic Beverage Control Act and the Federal Alcohol Administration Act, both of which are administered by the Bureau of Alcohol, Tobacco and Firearms (ATF). See the 2002 advisory posted by the ATF: `https://www.abc.ca.gov/trade/Advisory-SlottingFees.htm`, last accessed by the authors on November 4, 2014.

Table 1: Revenue Shares and HHI

Year	ABI	MillerCoors	Miller	Coors	Modelo	Heineken	Total	HHI
2001	0.37	.	0.20	0.12	0.08	0.04	0.81	2,043
2003	0.39	.	0.19	0.11	0.08	0.05	0.82	2,092
2005	0.36	.	0.19	0.11	0.09	0.05	0.79	1,907
2007	0.35	.	0.18	0.11	0.10	0.06	0.80	1,853
2009	0.37	0.29	.	.	0.09	0.05	0.80	2,350
2011	0.35	0.28	.	.	0.09	0.07	0.79	2,162

Notes: The table provides revenue shares and the Herfindahl-Hirschman Index (HHI) over 2001-2011. Firm-specific revenue shares are provided for ABI, Miller, Coors, Modelo, Heineken. The total across these firms also is provided. The HHI is scaled from 0 to $10,000$. The revenue shares incorporate changes in brand ownership during the sample period, including the merger of Anheuser-Busch (AB) and Inbev to form A-B Inbev (ABI), which closed in April 2009, and the acquisition by Heineken of the FEMSA brands in April 2010. All statistics are based on supermarket sales recorded in IRI scanner data.

2.2 Data Sources

Our primary data source is retail scanner data from the IRI Academic Database (Bronnenberg, Kruger and Mela (2008)). The data include revenue and unit sales by UPC code, by week and store, for a sample of supermarkets over 2001-2011. We restrict the regression samples to 39 distinct geographic regions and 13 flagship brands. These brands include Bud Light, Budweiser, Michelob, Michelob Light, Miller Lite, Miller Genuine Draft, Miller High Life, Coors Light, Coors, Corona Extra, Corona Extra Light, Heineken, and Heineken Light. The most popular brands that we omit either are regional brands (e.g., Yuengling), or are in the so-called "subpremium" category and sell at much lower price points.

Beer is sold in different package sizes, and hereafter we refer to brand×size combinations as "products." We focus on 6-packs, 12-packs, 24-packs, and 30-packs. Thus, "Bud Light 12-Pack" is one product in the sample. We combine 24-packs and 30-packs in the construction of our products because whether 24-packs or 30-packs are sold tends to depend on region-specific historical considerations. We exclude 18-packs and promotional package sizes, which are much less popular. Following standard practice, we measure market shares based on 144-ounce equivalent units, the size of a 12-pack. This means, for example, that the sale of a 6-pack is down-weighted by 50% in the construction of market shares. Prices then are defined as the ratio of revenue to equivalent unit sales. Typically the larger package sizes are less expensive on an equivalent unit basis. In total, 12-packs produce the greatest number of unit sales and 24-packs account for the greatest sales volume.

For computational reasons, we aggregate the data from the store-week level to the region-month and region-quarter levels. A potential concern with our static approach to

estimating demand is that sales and consumer stockpiling could cause a bias that understates unilateral incentives to raise prices (Hendel and Nevo (2006)). While aggregating over time reduces this bias only in special cases, we vary the periodicity of the sample in this way to give some assurance that dynamic considerations do not drive the results. The identification strategy does not require week-to-week variation in the data, so this aggregation may even be helpful insofar as it reduces random measurement error. We revisit our descriptive regressions in the online appendix using store-level observations and show that the main empirical patterns are not created by changes in the store-level composition of the IRI data. There are 167,695 observations at the product-region-month-year level, spanning 2001-2011.

We use household demographics from the Public Use Microdata Sample (PUMS) of the American Community Survey to estimate demand. The PUMS data are available annually over 2005-2011. Households are identified as residing within specified geographic areas, each of which has at least 100,000 residents based on the 2000 Census. We merge the PUMS data to the IRI scanner data by matching on the counties that compose the IRI regions and the PUMS areas. In estimation, we take 500 draws on households per region-year, and obtain household income as total income divided by the number of household members. When using the PUMS, we necessarily restrict attention to the 2005-2011 period. We also discard data from the first year following the Miller/Coors merger to allow for the realization of cost reductions. There are 94,656 qualifying observations at the product-region-month-year level and 31,784 observations at the product-region-quarter-year level.

We obtain the driving miles between each IRI region and the nearest brewery for each product in our sample using Google Maps. For imported brands, we define the miles traveled based on the nearest port into which the beer is shipped.[4] We construct a notion of "distance" based on the interaction of driving miles and diesel fuel prices, which we obtain from the Energy Information Agency of the Department of Energy. This allows us to capture variation in transportation costs that arises both cross-sectionally, based on the location of regions and breweries, and inter-temporally, based on fluctuations in fuel costs. It also helps us estimate the distributional cost-savings of the Miller/Coors merger. All prices and incomes are deflated using the CPI and are reported in 2010 dollars.

[4]We obtain the location of Heineken's primary ports from the website of BDP, a logistics firm hired by Heineken to improve its operational efficiency. See `http://www.bdpinternational.com/clients/heineken/`, last accessed on February 26, 2015. The ports include Baltimore, Charleston, Houston, Port of Long Beach, Miami, Seattle, Oakland, Boston, and New York. We measure the shipping distance for Grupo Modelo brands as the driving distance from each retail location to Ciudad Obregon, Mexico.

3 Retail Prices

3.1 Time-series variation

Figure 1 plots average log retail prices over 2001-2011 for each firm's best selling 12-pack: Bud Light, Miller Lite, Coors Light, Corona Extra, and Heineken. The vertical line drawn at June 2008 signifies the consummation of the Miller/Coors merger. Horizontal ticks are placed at October because brewers typically adjust their prices in early autumn. Retail prices trend downward before the merger for all five products, a period spanning more than seven years in the data. After the merger, the prices of Bud Light, Miller Lite and Coors Light increase by about 8% and there is no obvious continuation of the downward trend. The prices of Corona Extra and Heineken do not exhibit any persistent increase and instead continue along a downward trend. The price gap between the cheaper domestic beers and the more expensive imports shrinks over time in the post-merger periods.

The most theoretically interesting aspects of the figure are that (i) the price of Bud Light increases by roughly the same amount as the prices of Miller Lite and Coors Light, and (ii) Modelo and Heineken prices do not increase, at least not persistently. Post-merger coordination between ABI and MillerCoors is one possible explanation. Alternatively, the data could be explained solely from unilateral effects, under a particular set of demand elasticities that produces strong strategic complementarity among the prices of domestic beers, and weak strategic complementarity between the prices of domestic and imported beers. Specific institutional practices also could be important. As one example, retailers could set equal prices for Bud Light, Miller Lite and Coors Light, regardless of external circumstances, due to some beliefs about the market or because of pressure from the brewers. Changing macroeconomic conditions also are relevant because the merger coincides with the onset of the Great Recession. Income losses could decrease the demand elasticities of domestic beer for a variety of reasons, including down-market substitution.[5]

We use difference-in-differences regressions to quantify the changes and extend inference to the other flagship brands. The regression equation specifies the log retail price of product

[5]The observed price patterns probably are not due exclusively to down-market substitution, however, because the sales of the domestic brands decrease both in absolute and relative terms with the recession.

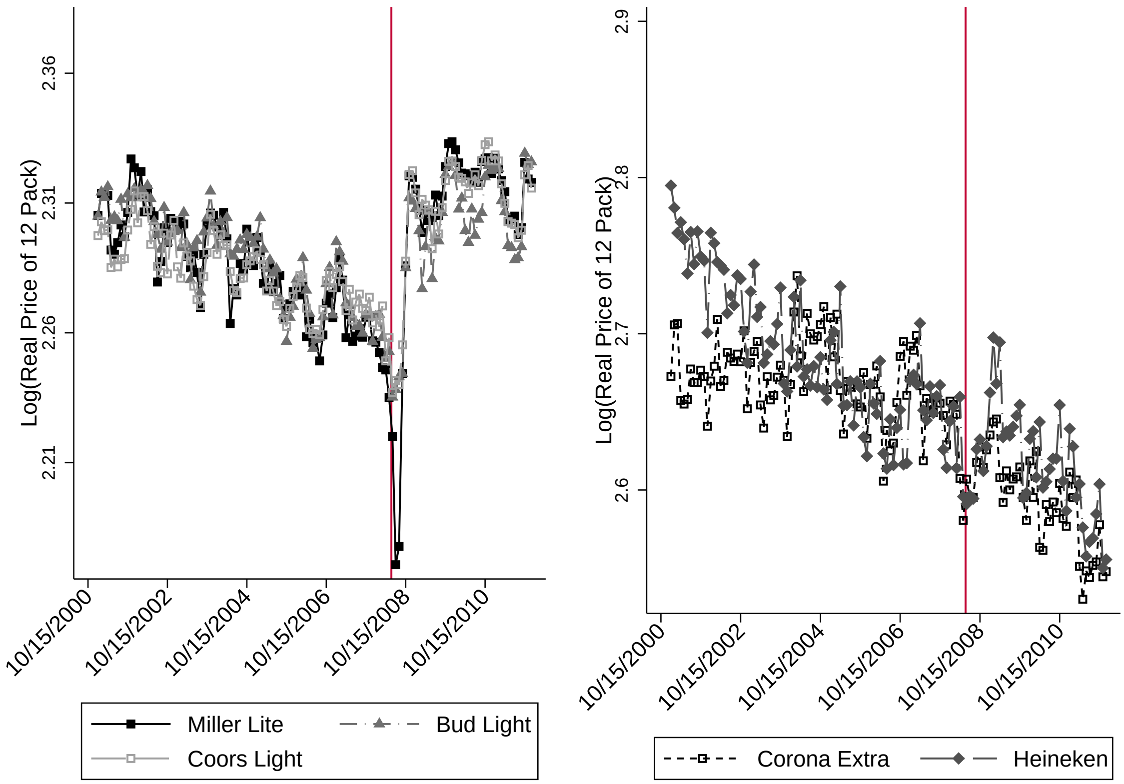

Figure 1: Average Retail Prices of Flagship Brand 12-Packs

Notes: The figure plots the national average price of a 12-pack over 2001-2011, separately for Bud Light, Miller Lite, Coors Light, Corona Extra and Heineken. The vertical axis is the natural log of the price in real 2010 dollars. The vertical bar drawn at June 2008 signifies the consummation of the Miller/Coors merger.

j in region r in period t according to

$$\begin{aligned} \log p_{jrt}^R &= \beta_1 \mathbb{1}\{\text{MillerCoors}\}_{jt} \times \mathbb{1}\{\text{Post-Merger}\}_t \\ &+ \beta_2 \mathbb{1}\{\text{ABI}\}_{jt} \times \mathbb{1}\{\text{Post-Merger}\}_t \\ &+ \beta_3 \mathbb{1}\{\text{Post-Merger}\}_t + \phi_{jr} + \tau_t + \epsilon_{jrt} \end{aligned} \tag{1}$$

which includes indicator variables for (i) MillerCoors brands in the post-merger periods, (ii) ABI brands in the post-merger periods, and (iii) all products in the post-merger periods. We absorb cross-sectional variation with product × region fixed effects (the ϕ_{jr} parameters). We use a linear time trend (τ_t) to account for the secular downward trend in prices. In some specifications, we expand on equation (1) by allowing the linear trend to vary freely across products, and by adding quarterly employment levels and average weekly earnings from the

Table 2: Changes in Retail Prices by Firm

	(i)	(ii)	(iii)	(iv)
$\mathbb{1}$\{MillerCoors\} $\times\mathbb{1}$\{Post-Merger\}	0.098	0.050	0.047	0.069
	(0.007)	(0.004)	(0.005)	(0.007)
$\mathbb{1}$\{ABI\}$\times\mathbb{1}$\{Post-Merger\}	0.087	0.040	0.038	0.062
	(0.007)	(0.005)	(0.005)	(0.007)
$\mathbb{1}$\{Post-Merger\}	-0.031	-0.007	-0.005	0.008
	(0.005)	(0.004)	(0.004)	(0.006)
log(Employment)	-	-	-0.051	0.131
	-	-	(0.080)	(0.081)
log(Earnings)	-	-	0.156	0.152
	-	-	(0.029)	(0.035)
Pre-Merger Average Price	11.75	11.14	11.14	11.14
Product Trends	No	No	Yes	Yes
Covariates	No	No	Yes	Yes
# Observations	25,740	167,695	167,695	151,525

Notes: Estimation is with OLS. The dependent variable is log real retail price. Observations are at the brand-size-region-month-year level. Column (i) contains 12-packs of Bud Light, Coors Light, Miller Lite, Corona Extra, and Heineken. Columns (ii) and (iii) contain 6, 12, and 24-packs of these brands plus Budweiser, Michelob Light, Michelob Ultra, Coors, Miller Genuine Draft, Miller High Life, Corona Light, and Heineken Premium Light. The estimation sample spans 39 regions from 2001-2011, except Column (iv) which excludes June, 2008 through May, 2009. All regressions include a linear time trend and product (brand×size) fixed effects interacted with region fixed effects. Standard errors are clustered at the region level and shown in parentheses.

Quarterly Census of Employment and Wages to help capture local economic conditions.

Table 2 presents the regression results. The sample in column (i) includes 12-packs of Bud Light, Coors Light, Miller Lite, Heineken and Corona, and so corresponds precisely to Figure 1. The coefficients indicate that MillerCoors and ABI prices increase by an average of 10.3 percent (because $\exp(0.098)-1 = 0.103$) and 9.1 percent, respectively, relative to imported brands. The absolute price increases are 6.8 and 5.7 percent. The difference between the MillerCoors and ABI coefficients is not statistically significant. Column (ii) incorporates the other brands and package sizes in the sample. Absolute and relative price increases for MillerCoors and ABI are then about four and five percent, respectively. The third column allows the time trend to vary freely by brand and package size and includes regional quarterly employment rates and average weekly earnings. The results are essentially unchanged. The final column excludes one year of data immediately following the merger. The estimates again reflect the pattern shown in Figure 1 and increase in magnitude. Ultimately, average prices for MillerCoors and ABI brands increased by between six and seven percent.

3.2 Cross-sectional Variation in Price Increases

Empirical analyses of mergers sometimes assume that price effects are proportional to the predicted change in HHI (or ΔHHI) induced by the merger (e.g., Dafny, Duggan and Ramanarayanan (2012); Ashenfelter, Hosken and Weinberg (2015)). For differentiated-product markets, there is a theoretical justification for this assumption if consumers substitute to other products in proportion to their market shares, because the relevant diversion ratios then can be approximated as a function of ΔHHI.[6] In this section, we exploit cross-sectional variation in the data to evaluate whether the price patterns developed above are well explained by the region-specific ΔHHI caused by the merger.

The regression equation specifies the log retail price of product j in region r in period t according to

$$\begin{aligned} \log p_{jrt}^{R} &= \alpha_1 \Delta\text{HHI}_r \times \mathbb{1}\{\text{Post-Merger}\}_t \\ &+ \alpha_2 \Delta\text{DIST}_r \times \mathbb{1}\{\text{Post-Merger}\}_t \\ &+ \alpha_3 \mathbb{1}\{\text{Post-Merger}\}_t + \phi_{jr} + \tau_t + \epsilon_{jrt} \end{aligned} \tag{2}$$

where ΔHHI_r is calculated based on data from the 18 months preceding the merger (scaled to be between zero and one) and ΔMILES_r is the reduction in miles from the brewery to the region experience by the Coors brands (in thousands). The error structure incorporates time effects and product × region fixed effects. We estimate equation (2) separately for MillerCoors, ABI, and Modelo/Heineken.[7]

The results are shown in Table 3. The price increases of MillerCoors and ABI are higher in regions with a greater ΔHHI, and lower in regions that experience a greater reduction in Coors' shipping distances. This is consistent with unilateral effects theory under proportional substitution. The net effect of greater concentration and lower shipping distances is close to zero on average.[8] Thus, the post-merger indicator accounts for most of the overall

[6]Consider a merger that involves two products with pre-merger market shares s_j and s_k, respectively. The predicted HHI change is ΔHHI= $2s_js_k$. If consumer substitution is proportional to market shares then diversion from product j to product k equals $s_k/(1-s_j)$, and can be approximated by $s_k(1+s_j)$ for small s_j. Diversion from k to j is analogous, meaning that the sum of the approximate diversion ratios is $s_j+s_k+2s_js_k$ or, equivalently, $s_j+s_k+\Delta$HHI. We first encountered these mathematics in Shapiro (2010). Miller, Remer, Ryan and Sheu (2016) provide Monte Carlo evidence that ΔHHI is highly correlated with unilateral price effects in the specific setting of proportional substitution.

[7]This replicates the analysis in Ashenfelter, Hosken and Weinberg (2015), which estimates equation (2) with proprietary IRI data that spans 2007-2011 and 47 geographic markets. Similar results are obtained.

[8]The average increase in concentration across markets is 0.02, which is associated with a price increase of 2.0 percent ($0.997 \times 0.02 = 0.020$). The average reduction in miles is 360, which implies a price reduction of 1.5 percent ($0.36 \times -0.042 = 0.015$).

Table 3: Cross-sectional Variation in Price Increases

	Pooled	MillerCoors	ABI	Imports
ΔHHI $\times$ $\mathbb{1}$\{Post-Merger\}	0.997	1.172	1.503	-0.005
	(0.454)	(0.542)	(0.531)	(0.534)
ΔMILES $\times$ $\mathbb{1}$\{Post-Merger\}	-0.042	-0.040	-0.053	-0.028
	(0.013)	(0.016)	(0.013)	(0.014)
$\mathbb{1}$\{Post-Merger\}	0.037	0.049	0.040	0.019
	(0.012)	(0.014)	(0.013)	(0.014)
# Observations	167,695	75,315	50,810	41,570

Notes: Estimation is with OLS. The dependent variable is log real retail price. Observations are at the brand-size-region-month-year level. The estimation sample spans 39 regions from 2001-2011. All regressions include a linear trend and product (brand $\times$size) fixed effects interacted with region fixed effects. Standard errors are clustered at the region level and shown in parentheses.

price increases shown previously; these are estimated to be 4.9 percent for MillerCoors and 4.0 percent for ABI. There is no statistically significant relationship between ΔHHI and import prices, though import prices are negatively related to ΔMILES. The magnitude and statistical significance of the post-merger indicator variable in these regressions suggests that unilateral effects may not fully account for the estimated price increases. Only weak inferences can be drawn, however, because the extent to which ΔHHI captures unilateral price effects depends how closely consumer substitution is proportional to market share. This helps motivate the additional structure that we place on the model, which allows us to estimate demand elasticities from data and account for unilateral effects directly.

3.3 Documentary record

There is some documentary evidence in the public domain that supports coordinated pricing by ABI and MillerCoors. The DOJ Complaint filed to enjoin the acquisition of Grupo Modelo by ABI alleges that ABI and MillerCoors announce (nominal) price increases each year in late summer to take effect in early fall. In most geographic areas, ABI is the market share leader and announces its price increase first; in other areas MillerCoors announces first. The price increases usually are matched by the follower, and if not they are rescinded. The Complaint quotes from the normal course documents of ABI:

> The specifics of ABI's pricing strategy are governed by its "Conduct Plan," a strategic plan for pricing in the United States that reads like a how-to manual for successful price coordination. The goals of the Conduct Plan include "yielding

> the highest level of followership in the short-term" and "improving competitor conduct over the long-term."
>
> ABI's Conduct Plan emphasizes the importance of being "Transparent – so competitors can clearly see the plan;" "Simple – so competitors can understand the plan;" "Consistent – so competitors can predict the plan;" and "Targeted – consider competition's structure." By pursuing these goals, ABI seeks to "dictate consistent and transparent competitive response."

The Complain does not identify the date at which ABI adopted its Conduct Plan, but some inferences can be made from the annual reports of the companies. The 2005 SABMiller annual report describes "intensified competition" and an "extremely competitive environment." The 2005 Anheuser-Busch report states that the company was "collapsing the price umbrella by reducing our price premium relative to major domestic competitors." SABMiller characterizes price competition as "intense" in its 2006 and 2007 reports. The tenor of the annual reports changes around the time of the merger. In its 2009 report, SABMiller attributes increasing earnings before interest, taxes, and amortization expenses to "robust pricing" and "reduced promotions and discounts." In its 2010 and 2011 reports, it references "sustained price increases" and "disciplined revenue management with selected price increases."[9]

The record supports that any coordination is limited to ABI and MillerCoors. The DOJ Complaint alleges that Modelo did not join the price increases and instead adopted a "Momentum Plan" that was designed to "grow Modelo's market share by shrinking the price gaps between brands owned by Modelo and domestic premium brands." The practical consequence of the Momentum Plan is that the nominal prices of Modelo remain flat even as ABI and MillerCoors prices increase. This limited the ability of ABI and MillerCoors to raise price due to the greater substitution of consumers to Modelo. The Complaint does not address the pricing practices of Heineken, though in the retail sales data we examine, the prices of Heineken's beers are similar to those of Corona.

[9]See the SABMiller Annual Report in 2005 (p. 13), 2006 (p. 5), 2007 (pp. 4 and 8), 2009 (p. 9 and 24), 2010 (pp. 29) and 2011 (p. 28), and the Anheuser-Busch Annual Report in 2005 (p. 5). The ABI annual reports in the post-merger years are more opaque.

4 Consumer Demand

4.1 Model

We use the random coefficient nested logit (RCNL) model to estimate consumer demand. The RCNL model has been applied in a number of recent empirical articles (e.g., Grennan (2013); Ciliberto and Williams (2014); Conlon and Rao (2016)), and similar discrete choice random utility models have been applied to the beer industry (e.g., Hellerstein (2008); Goldberg and Hellerstein (2013); Romeo (2016); Asker (2016)).

Suppose we observe $r = 1, \ldots, R$ regions over $t = 1, \ldots, T$ time periods. There are $i = 1, \ldots, N_{rt}$ consumers in each region-period combination. Each consumer purchases one of the observed products ($j = 1, \ldots, J_{rt}$) or selects the outside option ($j = 0$). We refer to observed products as "inside goods." The conditional indirect utility that consumer i receives from the inside good j in region r and period t is

$$u_{ijrt} \quad = \quad x_j \beta_i^* - \alpha_i^* p_{jrt} + \sigma_j^D + \tau_t^D + \xi_{jrt} + \bar{\epsilon}_{ijrt} \tag{3}$$

where x_j is a vector of observable product characteristics, p_{jrt} is the retail price, σ_j^D allows the mean valuation of unobserved product characteristics to vary freely by product, τ_t^D allows the mean valuation of the indirect utility from consuming the inside goods to vary freely over time, ξ_{jrt} is a region-period specific unobserved quality valuation, and $\bar{\epsilon}_{ijrt}$ is a stochastic term.

The observable product characteristics include a constant (i.e., an indicator that equals one for an inside good), calories, package size, and an indicator for whether the product is imported. Calories is highly correlated with alcohol content and serves to distinguish the "light" beers. We control for σ_j^D and τ_t^D using product and period fixed dummy variables, respectively. ξ_{jrt} is left as a structural error term. We specify the consumer-specific coefficients as $[\alpha_i^*\ , \beta_i^*]' = [\alpha, \beta]' + \Pi D_i$ where D_i is consumer income. Because the observable product characteristics are invariant over time, the mean consumer valuations for observables are absorbed by the product fixed effects in estimation. Many choice models allow consumer tastes to depend on random draws from a normal distribution (e.g., Berry, Levinsohn and Pakes (1995) and Nevo (2001)), which captures unobserved heterogeneity. These draws have no statistical power in our application and we opt for the simpler specification.

We decompose the stochastic term using the distributional assumptions of the nested logit model, following Berry (1994) and Cardell (1997). Define two groups, $g = 0, 1$, such

that group 1 includes the inside goods and group 0 is the outside good. Then

$$\bar{\epsilon}_{ijrt} = \zeta_{igrt} + (1-\rho)\epsilon_{ijrt} \tag{4}$$

where ϵ_{ijrt} is i.i.d extreme value, ζ_{igrt} has the unique distribution such that $\bar{\epsilon}_{ijrt}$ is extreme value, and ρ is a nesting parameter ($0 \leq \rho < 1$). Larger values of ρ correspond to a greater correlation in preferences for products of the same group, and thus less consumer substitution between the inside and outside goods. To close the model, we normalize the indirect utility of the outside good such that $u_{i0rt} = \epsilon_{i0rt}$, and normalize market size to be 50% greater than the maximum observed unit sales within the region.

In estimation, it is useful to decompose indirect utility such that

$$\begin{aligned} u_{ijrt} &= \delta_{jrt}(x_j, p_{jrt}, \sigma_j^D, \tau_t^D, \xi_{jrt}; \alpha, \beta) + \mu_{ijrt}(x_j, p_{jrt}, D_i; \Pi) + \zeta_{igrt} + (1-\rho)\epsilon_{ijrt} \\ \delta_{jrt} &= x_j\beta - \alpha p_{jrt} + \sigma_j^D + \tau_t^D + \xi_{jrt} \\ \mu_{ijrt} &= [p_{jrt}, x_j]' * \Pi D_i \end{aligned} \tag{5}$$

where $\delta_{jrt}(x_j, p_{jrt}, \sigma_j^D, \tau_t^D, \xi_{jrt}; \alpha, \beta)$ is the mean consumer valuation of product j in region r and period t, and the consumer-specific deviations are contained in $\mu_{ijrt}(x_j, D_i; \Pi) + \zeta_{igrt} + (1-\rho)\epsilon_{ijrt}$. Suppressing function arguments, the market share of good j in region r and period t is given by

$$s_{jrt} = \frac{1}{N_{rt}} \sum_{i=1}^{N_{rt}} \frac{\exp((\delta_{jrt} + \mu_{ijrt})/(1-\rho))}{\exp(I_{igrt}/(1-\rho))} \frac{\exp I_{igrt}}{\exp I_{irt}} \tag{6}$$

where I_{igrt} and I_{irt} are the McFadden (1978) inclusive values. The normalization on the mean indirect utility of the outside good yields $I_{i0rt} = 0$, while the inclusive value of the inside goods is $I_{i1rt} = (1-\rho)\log\sum_{j=1}^{J_{rt}} \exp((\delta_{jrt} + \mu_{ijrt})/(1-\rho))$ and the inclusive value of all goods is $I_{irt} = \log(1 + \exp I_{i1rt})$.

This RCNL model reduces to the nested logit model if $\Pi = 0$. This yields a equation that is linear in the parameters:

$$\log(s_{jrt}) - \log(s_{0rt}) = x_j\beta - \alpha p_{jrt} + \sigma_j^D + \tau_t^D + \rho\log(\bar{s}_{jrt|g}) + \xi_{jrt} \tag{7}$$

where $\bar{s}_{jrt|g} = s_{jrt}/\sum_{j=1}^{J_{rt}} s_{jrt}$ is the conditional share of product j among the inside goods. In this formulation, the nesting parameter ensures that the estimated elasticities are not overly sensitive to the market size normalization. The model nonetheless retains the property that

substitution patterns among inside goods are a function only of market shares. The full RCNL model relaxes this restriction by allowing consumer income to affect relative choice probabilities (it also allows the recession to affect demand in a natural way).

4.2 Estimation and instruments

We estimate the demand model using the nested fixed point procedure of Berry, Levinsohn and Pakes (1995). This approach derives a GMM estimator from the population moment condition $E[Z' \cdot \omega(\theta_0^D)] = 0$, where ω is a vector of stacked structural error terms, $\theta_0^D = (\alpha, \Pi, \rho)$ is the vector of parameters, and Z is a conformable matrix of instruments. The GMM estimate is

$$\widehat{\theta}^D = \arg\min_{\theta} \omega(\theta)' Z A^{-1} Z' \omega(\theta) \tag{8}$$

where A is a positive definite weighting matrix. For each candidate parameter vector, a contraction mapping identifies the vector of mean utility levels, δ^*, that solve the implicit system of equations $s(x, p, \delta^*; \Pi, \rho) = S$, where $s(\cdot)$ is a vector of market shares with elements defined by equation (6), p is a vector of all prices, and S is a vector of observed market shares. The structural error term then is calculated as $\omega_{jrt} = \delta^*_{jrt}(x, p, S; \Pi, \rho) + \alpha p_{jrt}$.

We employ a standard two step procedure for GMM estimation (e.g., Hansen (1982)). In the first step, we set $A = Z'Z$. In the second step, we reestimate the model using an optimal weighting matrix that employs an Eicker-White-Huber cluster correction to correct for heteroskedasticity, autocorrelation and within-region correlations (Bhattacharya 2005). We concentrate the price parameter, α, out of the optimization problem using 2SLS to reduce the dimensionality of the nonlinear search.

Identification requires at least one instrument for price and each nonlinear parameter. Prices are likely to be correlated with the structural error term because firms set prices with knowledge of product and market-specific consumer valuations. The creates a standard endogeneity problem that can be overcome if an instrument provides sufficient exogenous variation. Further, as highlighted in Berry and Haile (2014), the presence of heterogeneity in consumer preferences for product characteristics introduces a simultaneity problem that arises from the interaction of unknown demand parameters with market shares. In our specification of the RCNL, this heterogeneity is due to the income terms (i.e., the Π parameter) and the nested logit term (ρ).

The first set of instruments that we use address the endogeneity of prices. It includes

the distance between the brewery and the region (miles $\times$ diesel index) and an indicator equal to one for ABI and MillerCoors products after the merger. Both instruments arise from the supply-side of the model; distance shifts marginal costs and the indicator captures a change in the competitive structure of the industry. That the indicator has power is suggested by the estimated increases in prices after the Miller/Coors merger. Given the presence of product and time fixed effects, it is valid if the changes in the structural error terms of ABI and MillerCoors following the merger are not systematically different from the changes in the structural error terms of Modelo and Heineken.[10]

The second set of instruments helps identify the nested logit parameter, which governs the degree of correlation in unobserved preferences for the inside goods. What is required is exogenous variation in the conditional shares of the inside goods (i.e., variation in $\overline{s}_{jrt|g} = s_{jrt} / \sum_{j=1}^{J_{rt}} s_{jrt}$). This is made clear in the linear formulation of the nested logit model in equation (7). We use as instruments the number of products in the market and distance summed across all products in the market. The effect of these variables on choice probabilities need not be uniform in the sample, and to add flexibility we incorporate interactions with indicators for ABI and Miller/Coors products. The number of products is a standard "BLP" instrument and should be negatively correlated with the conditional share. Total distance captures variation in the marginal costs of competing products and should be positively correlated with conditional share. Validity requires that the structural error term is uncorrelated with the number of products.

Finally, we use mean income interacted with the observed product characteristics (a constant, calories, package size, and an import dummy). Under the assumption that mean income is orthogonal to the structural error, these instruments provide exogenous variation in the derivative of the structural error term with respect to the random coefficients, and thereby aid identification. Romeo (2014) provides evidence that similar instruments improve the numerical performance of random coefficient logit estimates. There are 12 instruments in total. We evaluate the relevance of these instruments, along with related considerations, in the appendix materials.

4.3 Results of demand estimation

Table 4 presents the results. Column (i) corresponds to the nested logit demand model, which we estimate with 2SLS in order to provide a simple benchmark. The remaining

[10]The indicator for ABI and MillerCoors products in the post-merger periods belongs to a class instruments proposed in Berry and Haile (2014). We refer readers to that article for a formal analysis.

Table 4: Baseline Demand Estimates

Demand Model:		NL-1	RCNL-1	RCNL-2	RCNL-3	RCNL-4
Data Frequency:		monthly	monthly	quarterly	monthly	quarterly
Variable	Parameter	(i)	(ii)	(iii)	(iv)	(v)
Price	α	-0.1377	-0.0887	-0.1087	-0.0798	-0.0944
		(0.0933)	(0.0141)	(0.0163)	(0.0147)	(0.0146)
Nesting Parameter	ρ	0.6067	0.8299	0.7779	0.8079	0.8344
		(0.0994)	(0.0402)	(0.0479)	(0.0602)	(0.0519)
Demographic Interactions						
Income×Price	Π_1		0.0007	0.0009		
			(0.0002)	(0.0003)		
Income×Constant	Π_2		0.0143	0.0125	0.0228	0.0241
			(0.0051)	(0.0055)	(0.0042)	(0.0042)
Income×Calories	Π_3		0.0043	0.0045	0.0038	0.0031
			(0.0016)	(0.0017)	(0.0018)	(0.0015)
Income×Import	Π_4				0.0039	0.0031
					(0.0019)	(0.0016)
Income×Package Size	Π_5				-0.0013	-0.0017
					(0.0007)	(0.006)
Other Statistics						
Median Own Price Elasticity		-3.76	-4.74	-4.33	-4.45	-6.10
Median Market Price Elasticity		-1.09	-0.60	-0.72	-0.60	-0.69
Median Outside Diversion		29.86%	12.96%	16.98%	13.91%	11.82%
$J-$Statistic			13.94	13.75	13.91	14.15

Notes: The table shows the baseline demand results. Estimation is with 2SLS in column (i) and with GMM in columns (ii)-(v). There are 94,656 observations at the brand-size-region-month-year level in columns (i), (ii), and (iv), and 31,784 observations at the brand-size-region-year-quarter level in columns (iii) and (v). The samples exclude the months/quarters between June 2008 and May 2009. All regressions include product (brand×size) and period (month or quarter) fixed effects. The elasticity and diversion numbers represent medians among all the brand-size-region-month/quarter-year observations. Standard errors clustered by region and shown in parentheses.

columns correspond to the full RCNL model. We use two main specifications. In columns (ii) and (iii), consumer income affects preferences for price, the inside good constant, and calories. In columns (iv) and (v), income affects preferences for the constant, calories, imports, and package size. Both specifications break the logit substitution patterns between domestic and imported beers, albeit with different mechanisms. The units of observation are brand-size-region-year-month combinations in columns (i), (ii), and (iv), and brand-size-region-year-quarter combinations in columns (iii) and (v). All regressions include product (i.e., brand×size) and period fixed effects.

The coefficients are precisely estimated and take the expected signs. The median

own price elasticities range from -4.33 to -6.10 with the RCNL models.[11] The market price elasticities are much lower, indicating that most substitution occurs within the inside goods, rather than between the inside goods and the outside good. This can be recast in terms of diversion: the outside good is the second-best choice for 12%-17% of consumers. The interactions reveal that consumers with higher incomes are less sensitive to price and tend to prefer the inside goods, imported brands, more calories, and smaller package sizes. Whether the monthly or quarterly data are used in estimation matters little. Finally, the Sargan-Hansen J-statistics are asymptotically χ^2 distributed with either eight (columns (ii) and (iii)) or seven (columns (iv) and (v)) degrees of freedom under the null hypothesis that the models are valid. The models cannot be rejected at the 95% confidence level.

Table 5 presents more detail on the elasticities that arise in the RCNL-1 specification. We provide the full elasticity matrix for 12-packs, along with aggregated cross-elasticities that summarize substitution from the 12-packs to selected categories of beer. One noticeable pattern is that own price elasticities tend to be somewhat higher for more expensive products. (This is not imposed in RCNL-1 due to the income×price interaction.) The logit restriction that consumers substitute to other products in proportion to their market shares is substantially relaxed. This can be seen by observing the heterogeneity that exists within a single column. For instance, column 1 shows that consumers of Bud Light 12-packs substitute disproportionately toward similar beers like Budweiser, Coors Light, and Miller Lite, and also toward the larger package sizes. Column 5 shows that, by contrast, consumers of Corona Extra substitute disproportionately toward Heineken and smaller package sizes.

[11] The own-price elasticities are somewhat greater than the own-price elasticities reported in Romeo (2016), and somewhat smaller than the own-price elasticities reported in Hellerstein (2008). Most similar are the elasticities of Slade (2004) and Pinske and Slade (2004), obtained for the U.K. beer industry.

Table 5: Mean Elasticities for 12-Pack Products from Specification RCNL-1

	Brand/Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
							Product-Specific Own and Cross Elasticities							
(1)	Bud Light	-4.389	0.160	0.019	0.182	0.235	0.101	0.146	0.047	0.040	0.130	0.046	0.072	0.196
(2)	Budweiser	0.323	-4.272	0.019	0.166	0.258	0.103	0.166	0.047	0.039	0.121	0.043	0.068	0.183
(3)	Coors	0.316	0.154	-4.371	0.163	0.259	0.102	0.167	0.046	0.038	0.119	0.042	0.066	0.180
(4)	Coors Light	0.351	0.160	0.019	-4.628	0.230	0.100	0.142	0.047	0.041	0.132	0.047	0.073	0.199
(5)	Corona Extra	0.279	0.147	0.018	0.137	-5.178	0.108	0.203	0.047	0.035	0.104	0.035	0.061	0.158
(6)	Corona Light	0.302	0.151	0.018	0.153	0.279	-5.795	0.183	0.048	0.037	0.113	0.039	0.065	0.171
(7)	Heineken	0.269	0.145	0.018	0.131	0.311	0.108	-5.147	0.047	0.035	0.101	0.034	0.059	0.153
(8)	Heineken Light	0.240	0.112	0.014	0.124	0.210	0.086	0.138	-5.900	0.026	0.089	0.028	0.051	0.135
(9)	Michelob	0.301	0.140	0.015	0.146	0.208	0.089	0.135	0.042	-4.970	0.116	0.036	0.061	0.175
(10)	Michelob Light	0.345	0.159	0.019	0.181	0.235	0.101	0.146	0.047	0.041	-5.071	0.046	0.072	0.196
(11)	Miller Gen. Draft	0.346	0.159	0.019	0.182	0.235	0.101	0.146	0.047	0.040	0.130	-4.696	0.072	0.196
(12)	Miller High Life	0.338	0.159	0.019	0.177	0.242	0.102	0.153	0.047	0.040	0.127	0.045	-3.495	0.191
(13)	Miller Lite	0.344	0.159	0.019	0.180	0.237	0.101	0.148	0.047	0.040	0.129	0.046	0.071	-4.517
(14)	Outside Good	0.016	0.007	0.001	0.009	0.011	0.005	0.006	0.002	0.002	0.006	0.002	0.003	0.009
							Total Cross Elasticities by Category							
	6-Packs	0.307	0.152	0.018	0.155	0.275	0.104	0.180	0.047	0.038	0.115	0.039	0.065	0.174
	12-Packs	0.320	0.154	0.019	0.163	0.250	0.102	0.161	0.047	0.039	0.121	0.042	0.068	0.183
	24-Packs	0.356	0.160	0.019	0.189	0.222	0.099	0.136	0.047	0.041	0.134	0.048	0.073	0.201
	Domestic	0.349	0.160	0.019	0.184	0.229	0.100	0.142	0.047	0.040	0.131	0.047	0.072	0.197
	Imported	0.279	0.147	0.018	0.138	0.301	0.108	0.200	0.047	0.035	0.104	0.035	0.061	0.158

Notes: The table provides mean elasticities of demand for 12-packs based on the RCNL-1 specification (column (ii) of Table 4). The cell entry in row i and column j is the percentage change in the quantity of product i with respect to the price of product j. Means are calculated across the year-month-region combinations.

5 Supply

5.1 Model

We employ a model of price competition among firms of differentiated products that departs from the standard approach only in that ABI and MillerCoors may (partially or fully) internalize their pricing externalities in the post-merger periods. The vector of equilibrium prices in each region-period satisfies the first order condition

$$p_t = mc_t - \left[\Omega_t(\kappa) \circ \left(\frac{\partial s_t(p_t;\theta^D)}{\partial p_t}\right)^T\right]^{-1} s_t(p_t;\theta^D) \tag{9}$$

where Ω_t is the ownership matrix, s_t is vector of market shares, and the operation $\circ$ is element-by-element matrix multiplication. We suppress region subscripts for brevity. The $(j,\ k)$ element of the ownership matrix equals one if products j and k are produced by the same firm. Otherwise the $(j,\ k)$ element equals κ if products j and k are sold by ABI and MillerCoors and the period postdates the merger. Otherwise the $(j,\ k)$ element equals zero. This generates Nash-Bertrand in the post-merger periods if $\kappa = 0$ and joint profit maximization for ABI and MillerCoors if $\kappa = 1$.

To set intuition, consider a hypothetical region in a pre-merger period t_1^* and a post-merger period t_2^*, and suppose that there are $j = 1, ...4$ products sold by ABI, Miller, Coors, and Modelo, respectively. The ownership matrices are given by

$$\Omega_{t_1^*} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \Omega_{t_2^*} = \begin{bmatrix} 1 & \kappa & \kappa & 0 \\ \kappa & 1 & 1 & 0 \\ \kappa & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{10}$$

The pre-merger ownership matrix at t_1^* is diagonal because competition is Nash-Bertrand and each firm sells a single product in the hypothetical region. The post-merger matrix reflects that Miller and Coors fully internalize how their prices affect each other, and the κ parameter dictates the extent to which the ABI and MillerCoors internalize the effect of their prices. An important restriction is that Modelo and Heineken price a la Nash, which we motivate on the basis of qualitative evidence discussed earlier.

To finish the supply-side model, we parameterize the marginal cost of product j in

region r and period t as follows:

$$mc_{jrt} = w_{jrt}\gamma + \sigma_j^S + \tau_t^S + \mu_r^S + \eta_{jrt}, \tag{11}$$

where w_{jrt} is a vector that includes the distance (miles $\times$ diesel index) between the region and brewery, and an indicator for MillerCoors products in post-merger periods. This allows the Miller/Coors merger to affect marginal costs through the rationalization of distribution and through residual cost synergies that are unrelated to distance. Unobserved costs depend on the product, region, and period-specific effects, σ_j^S, μ_c^S, and τ_t^S, which we control for with fixed effects, as well as on η_{jrt}, which we leave as a structural error term.[12]

5.2 Estimation and instruments

We estimate the supply-side of the model taking as given the demand results. Collect the supply-side parameters to be estimated in the vector $\theta_0^S = (\kappa, \gamma)$. For each candidate parameter vector, $\widetilde{\theta}^S$, we calculate the markups and observed marginal costs, and obtain the structural error as a function of the parameters:

$$\eta_{rt}^*(\widetilde{\theta}^S; \widehat{\theta}^D) = p_{rt} - w_{jrt}\widetilde{\gamma} - \sigma_j^S - \tau_t^S - \mu_r^S - \left[\Omega_t(\widetilde{\kappa}) \circ \left(\frac{\partial s_t(p_{rt}; \widehat{\theta}^D)}{\partial p_{rt}}\right)^T\right]^{-1} s_t(p_{rt}; \widehat{\theta}^D) \tag{12}$$

Identification rests on the population moment condition $E[z' \cdot \eta^*(\theta_0^S)] = 0$ where $\eta^*(\theta_0^S)$ is a stacked vector of structural errors and z is a conformable vector that contains an excluded instrument. The method-of-moments estimate is

$$\widehat{\theta}^S = \arg\min_{\theta} \eta^*(\theta; \widehat{\theta}^D)' z z' \eta^*(\theta; \widehat{\theta}^D) \tag{13}$$

We concentrate the fixed effects and the marginal cost parameters out of the optimization problem using OLS to reduce the dimensionality of the nonlinear search. We cluster the standard errors at the region-level, and make an adjustment to account for the incorporation of demand-side estimates following Wooldridge (2010).

[12] The slope of the marginal cost function influences the magnitude of price changes that arise from unilateral effects. Suppose that ABI has an upward-sloping marginal cost function. Then as consumers shift to ABI in response higher prices from MillerCoors, ABI has both demand-side and cost-side incentives to raise price. Because the identification strategy is based on whether prices differ from what would be predicted on the basis of unilateral effects (accounting for changes in demand/costs), the estimated κ parameter would be biased upward unless the specification accounts for increasing costs. We assume a constant marginal cost function. There is little evidence that ABI experienced capacity constraints over the sample period.

The markup term in equation (12) is endogenous because unobserved costs enter implicitly through price. We instrument with an indicator that equals one for ABI and MillerCoors in the post-merger periods. The power of the instrument is supported by the descriptive regression results. Validity holds if the unobserved costs of ABI are orthogonal to the instrument. Given the specification of the marginal cost function, this will be the case if changes in the unobserved costs of ABI, before vs. after the merger, are not systematically different from changes in the unobserved costs of Modelo and Heineken. This is because the product and time fixed effects absorb level effects in the marginal cost function. Further, the MillerCoors post-merger indicator allows the merger to shift marginal costs of MillerCoors and thus isolates the comparison between ABI and Modelo/Heineken.

We comment briefly on the empirical variation that identifies the coefficients. The κ estimate is positive if the post-merger prices of ABI exceed what can be explained by the unilateral effects of the Miller/Coors merger. However, a positive κ affects the prices of both ABI and MillerCoors. Thus, the estimate of the parameter on the post-merger MillerCoors indicator in the marginal cost function is negative if post-merger MillerCoors prices are lower than what can be explained by unilateral effects and the κ estimate. Consider a simple numerical example. Demand is logit and competition is (initially) Nash-Bertrand among three symmetric firms. Prices are 1.00, marginal costs are 0.70, and market shares are 0.20. The share of the outside good is 0.40. This is sufficient to calibrate the demand system. Allow the first two firms to merge. Prices under post-merger Nash-Bertrand competition are (1.06,1.06,1.01). If instead $\kappa = 0.50$ the post-merger prices are (1.10, 1.10, 1.07). If $\kappa = 0.50$ and the merging firms reduce their marginal costs to 0.50 then post-merger prices are (1.09, 1.09, 1.08). How the prices of ABI and MillerCoors change with the merger drive the estimates of the supply-side parameters.

5.3 Relationship to the conduct parameter literature

Our identification strategy differs in important ways from the conduct parameter literature described in Bresnahan (1989). The reason is that the empirical variation needed to identify changes in an equilibrium concept is different than the empirical variation needed to identify the level of conduct. To illustrate, consider the model of Nevo (1998), which incorporates conduct parameters into a differentiated-products setting. Equilibrium prices are determined by a first order condition that is identical to equation (9), except that a more flexible ownership matrix is specified. The (j, k) element equals one if products j and k are produced by the same firm and κ_{jk} otherwise. This allows all firms to engage in non-Nash

pricing competition. Estimation requires instruments for the markup terms. If marginal costs are constant then the most likely candidates are demand-shifters; otherwise demand rotators can be used (Bresnahan (1982) and Lau (1982)).

Thus, the empirical variation that pins down the conduct parameters relates prices to demand. This has intuitive appeal. For example, if an industry is highly competitive then prices are likely to be near marginal cost regardless of demand conditions. By contrast, a monopolist usually has a incentive to adjust its prices in response to demand movements. Inference based on the demand-price relationship, however, is complicated by the observation that oligopoly supergames can support many equilibria. This is the famous "Corts critique" of the conduct parameter literature (Corts (1999)). To see the problem, suppose that firms set a high price regardless of demand conditions along the equilibrium path. There would be no relationship between demand and price in the data, which means that the researcher would mistakenly infer that the market is highly competitive market. Indeed, conduct is identified from the demand-price relationship only under quite special circumstances. In our application, supply-side identification does not rely on the relationship between demand and price, so the Corts critique does not apply.

5.4 Supply-side estimation results

Table 6 presents the supply-side results. As described, each column corresponds to one of the baseline demand specifications. The marginal cost functions incorporate product (i.e., brand$\times$size), period (month or quarter), and region fixed effects in all cases. As shown, the estimates of κ are positive and statically significant. The null of Nash-Bertrand competition in the post-merger periods is easily rejected. With the RCNL demand specifications, the estimates range from 0.249 to 0.342. Strictly interpreted, this corresponds to ABI and MillerCoors internalizing between roughly a quarter and a third of their price effects on the others profits in the post-merger periods.

Brewer markups can be obtained from the κ estimates and the structure of the model.[13] Table 7 provides the average markup for each product in the data, both before and after the Miller/Coors merger, based on the RCNL-1 specification (column (ii) in Tables 4 and 6). Across all the 94,656 brand-size-month-region observations, the average markup is \$3.60 on an equivalent-unit basis, and this accounts for 34% of the retail price. The average markups

[13]Equation (9) allows for the retail price to be decomposed into brewer markups and marginal costs. We show in the online appendix that the magnitude of marginal costs is sensitive to the incorporation of retail market power, which has an effect that is economically similar to a per-unit tax that must be paid by the brewers. The brewer markups are largely unaffected by the incorporation of retail market power.

Table 6: Baseline Supply Estimates

Demand Model:		NL-1	RCNL-1	RCNL-2	RCNL-3	RCNL-4
Data Frequency:		monthly	monthly	quarterly	monthly	quarterly
Variable	Parameter	(i)	(ii)	(iii)	(iv)	(v)
Post-Merger Internalization	κ	0.376	0.264	0.249	0.286	0.342
of Coalition Pricing Externalities		(0.032)	(0.073)	(0.087)	(0.042)	(0.054)
Marginal Cost Parameters						
MillerCoors×PostMerger	γ_1	-0.616	-0.654	-0.649	-0.722	-0.526
		(0.040)	(0.050)	(0.060)	(0.042)	(0.040)
Distance	γ_2	0.142	0.168	0.163	0.169	0.148
		(0.046)	(0.059)	(0.059)	(0.060)	(0.049)

Notes: The table shows the baseline supply results. Estimation is with the method-of-moments. There are 94,656 observations at the brand-size-region-month-year level in columns (i), (ii), and (iv), and 31,784 observations at the brand-size-region-year-quarter level in columns (iii) and (v). The samples exclude the months/quarters between June 2008 and May 2009. All regressions include product (brand×size), period (month or quarter), and region fixed effects. Standard errors clustered by region and shown in parentheses.

Table 7: Brewer Markups from RCNL-1

	6-Packs		12-Packs		24-Packs	
Brand	Pre	Post	Pre	Post	Pre	Post
Bud Light	3.63	4.34	3.52	4.24	3.43	4.13
Budweiser	3.79	4.49	3.66	4.38	3.55	4.25
Coors	2.70	4.39	2.56	4.31	2.44	4.18
Coors Light	2.47	4.21	2.36	4.14	2.28	4.04
Corona Extra	3.30	3.18	3.04	2.91	3.04	3.03
Corona Light	3.02	2.91	2.75	2.65	2.87	2.80
Heineken	3.20	3.14	2.98	2.92	3.22	3.33
Heineken Light	2.87	2.81	2.61	2.50	2.75	2.69
Michelob	3.69	4.47	3.62	4.38	3.34	4.28
Michelob Light	3.61	4.34	3.53	4.23	3.46	4.06
Miller Gen. Draft	2.89	4.26	2.77	4.16	2.68	4.09
Miller High Life	2.91	4.28	2.80	4.20	2.74	4.13
Miller Lite	2.89	4.25	2.78	4.18	2.69	4.07

Notes: The table provides average markups for each brand-size combination, separately for the pre-merger and post-merger periods, based on the RCNL-1 specification shown in column (ii) of Tables 4 and 6.

on ABI 12-packs tend to be higher in the post-merger periods by about $0.70. This reflects the higher retail prices previously shown in Figure 1. The markups on Miller 12-packs increase by $1.40 and the markups on Coors products increase by $1.80. Those changes can be attributed to the combined impact of higher retail prices and lower marginal costs. The markups on imported beers do not change much over the sample period.

Turning to the marginal cost shifters, the estimated distance parameters range from 0.148 to 0.169 with the RCNL demand models. The magnitude of the coefficient indicates

that distance costs account for two to three percent of the retail price on average. Total distribution costs may be partially absorbed by the fixed effects. (Tremblay and Tremblay (2005, p. 162) peg taxes and shipping at 17 percent of the retail price in the year 1996.) The marginal cost specification allows for the Miller/Coors merger to produce efficiencies both through a reduction in shipping distance and a downward shift in marginal costs that common to all regions. The estimates of the latter effect range from \$0.66 to \$0.70 with the RCNL demand models. This most likely reflects distributional savings that are not captured in the distance between the brewery and region. Our estimates imply a reduction in the marginal cost of Coors Light of about 14 percent, which can be compared against the 11 percent reduction predicted in the trade press (e.g., van Brugge et al (2007)).

Figure 2 explores the cross-sectional variation in the marginal cost reductions due to Miller/Coors merger, based on the RCNL-1 specification. Scatter plots are shown for Coors Light 12-packs (left panel) and Miller Lite 12-packs (right panel). The horizontal dimension shows the change in marginal cost due to the merger. The vertical dimension is the corresponding change in the retail price, which we obtain by recomputing equilibrium under the counterfactual that the merger does not reduce marginal costs. The cost reductions for Coors Light range from \$0.60 to \$1.30, depending on the region in question. The average pass-through is about an \$0.80 price reduction per \$1.00 cost reduction, and similar pass-through arises for other products. Miller Lite displays less cross-sectional variation due to the more limited distance reductions. The results are broadly consistent with Ashenfelter, Hosken and Weinberg (2014b), which finds that shipping efficiencies reduce prices by 1.8 percent in the average market. By point of comparison, a counterfactual simulation in which we allow the merger to change costs only through shipping distance reductions implies prices that are 2.4 percent lower than they would be with the initial distances.

We now discuss two subjects related to supply-side identification. First, if post-merger competition is Nash-Bertrand then higher post-merger marginal costs of ABI would be needed to rationalize the observed prices, relative to the marginal costs implied by the baseline model. We obtain both sets of marginal costs to make this comparison explicit. These are identical pre-merger but diverge post-merger. We then regress the natural log of costs on (i) an indicator for the post-merger periods; (ii) and indicator for Bertrand-Nash competition in the post-merger periods. We control for product$\times$region fixed effects, and include a linear trend to reflect the fact that real costs are falling throughout the sample.[14]

Table 8 presents the results. The first column indicates that it would take approxi-

[14]The exercise is in the spirit of Bresnahan (1987).

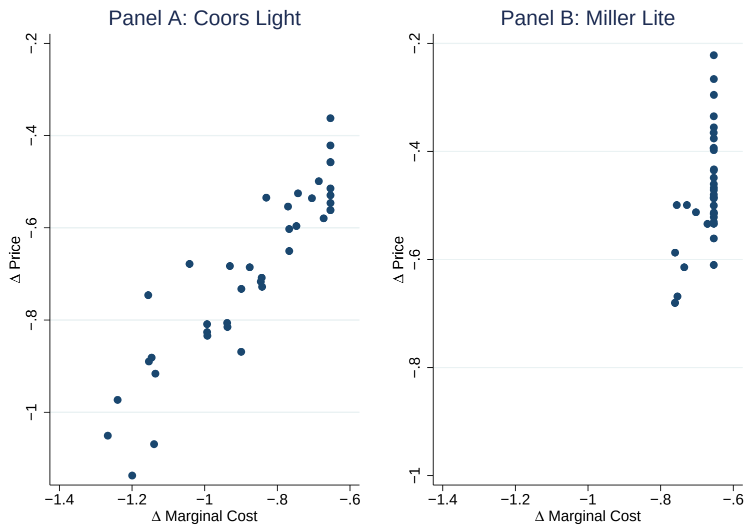

Figure 2: Change in Price against Change in Marginal Cost

Notes: The figure plots average regional difference in counterfactual price with no efficiencies and observed price against average regional difference in marginal cost for 12-Packs of Coors Light and Miller Lite. Each dot is a region-average in the year 2011 and is based on the RCNL-1 specification.

Table 8: Changes in ABI Log Costs with Bertrand and Coordination

	Budweiser	Bud Light	Michelob Light	Michelob Ultra
$\mathbb{1}${Post-Merger and Bertrand}	0.122	0.120	0.089	0.102
	(0.006)	(0.006)	(0.004)	(0.007)
$\mathbb{1}${Post-Merger}	0.016	-0.002	-0.044	0.050
	(0.014)	(0.011)	(0.011)	(0.013)

Notes: The dependent variable is log marginal costs from (i) the baseline model, and (ii) an alternative with Bertrand-Nash pricing in all periods. The RCNL-1 specification is used to obtain the implied marginal costs. Observations thus are at the brand-size-region-month-scenario level. The estimation sample excludes observations from June 2008 until May 2009. All regressions include product (brand×size) fixed effects interacted with region fixed effects. Standard errors are clustered at the region level and shown in parentheses.

mately an 13.8 percent $(12.2 + 1.6 = 13.8)$ increase in the costs of Budweiser, relative to the pre-merger costs, to rationalize observed prices under Nash Bertrand competition. By contrast, the baseline model implies that the marginal costs of ABI increase by 1.6 percent with the Miller/Coors merger. Results are similar for the other ABI brands. There is some documentary evidence in the public domain that helps assess the plausibility of these different cost predictions. Inbev motivated its 2009 purchase of Anhueser-Busch on the basis of cost efficiencies. Our understanding is that the merger lowered total costs, but that any

Table 9: Supply-Side Estimates with Alternative Pre-Merger Normalizations

	(i)	(ii)	(iii)	(iv)	(v)
Post-Merger Internalization	0.320	0.382	0.448	0.518	0.593
of Coalition Pricing Externalities	(0.066)	(0.058)	(0.051)	(0.043)	(0.036)
Pre-Merger Internalization	0.10	0.20	0.30	0.40	0.50

Notes: The table shows the baseline supply results, based on method-of-moments estimation. Results generated with the RCNL-1 demand specification. There are 94,656 observations at the brand-size-region-month-year level. The sample excludes the months between June 2008 and May 2009. All regressions include the baseline marginal cost shifters as well as product and period fixed effects. Standard errors clustered by region and shown in parentheses.

marginal cost effects were small because distribution was unaffected for the brands in our sample.[15] Thus, it seems likely that ABI's costs were flat after the Miller/Coors merger. Additionally, the implied ABI cost increases under Bertrand-Nash competition are of similar magnitude to the cost reductions we estimate for MillerCoors, which makes us skeptical that they would pass without notice in the annual reports of ABI and the popular press.

Second, we emphasize that the κ parameter allows us to test for a *change* in the equilibrium concept. Pre-merger competition is normalized to Nash-Bertrand in baseline model, but other normalizations are possible. Table 9 reports estimates obtained under some of these alternatives. Specifically, we impose a nonzero pre-merger κ parameter (0.10, 0.20, ..., 0.50) that governs interactions between the domestic brewers, and we estimate the corresponding post-merger parameters. As shown, the higher the pre-merger parameter, the higher is the post-merger parameter. In each case, the post-merger parameter is statistically different than the pre-merger normalization so the null hypothesis of no change in the equilibrium concept can be rejected. This reinforces that identification hinges on whether observed price changes can be explained by supply-side model, holding fixed the equilibrium concept. The nature of pre-merger competition is not pinned down.

5.5 Interpretation

One limitation of the baseline econometric results is that they are consistent with post-merger coordination, but do not directly inform whether this is actually caused by the merger. This mirrors the documentary record summarized in Section 3.3, which suggests softer price competition after 2008 but does not explain why the shift occurred. One alternative explanation is that coordination becomes easier to sustain when demand is weak (e.g., Rotemberg and

[15] Inbev revised the pay system, ended pension contributions and life insurance for retirees, and transferred the foreign beer operations of Anheuser Busch to InBev (Ascher (2012)).

Saloner (1986)). We explore this possibility in some detail here. The data and documentary record support that the recession had adverse effects on the sales of ABI and MillerCoors products: unit sales decrease both in absolute terms and relative to Modelo/Heineken, and the 2009 ABI Annual Report (p. 17) refers to "an economic environment that was the most difficult our industry has seen in many years."

Nonetheless, some of the empirical evidence cuts against this alternative hypothesis. As shown in Section 3, the prices of ABI/MillerCoors continue to increase relative to Modelo/Heineken over 2009-2011 during a period of macroeconomic recovery, and are positively associated with household earnings. To push this a bit farther, we estimate a supply-side model that allows for a different κ parameter in each post-merger period. Figure 3 plots the point estimates along with a 95% confidence interval. The estimates increase over time. It also is possible to allow κ to vary with mean income, thereby exploiting the cross-sectional variation. To implement, we constrain the values to be between 0 and 1 as follows:

$$\kappa_{rt} = \frac{\exp(a \times \mathbb{1}\{\text{Post-Merger}\}_{rt} + b \text{ Mean Income}_{rt})}{1 + \exp(a \times \mathbb{1}\{\text{Post-Merger}\} + b \text{ Mean Income}_{rt})} \tag{14}$$

where a and b are parameters. We estimate a to be 0.456 (standard error of 0.088) and b to be 0.033 (standard error of 0.039), which again does not provide econometric support for coordination being associated with weak demand. Given the totality of the evidence, we view the Miller/Coors merger as the most likely catalyst for softer price competition.

6 Counterfactual Analysis

We use counterfactual analysis to study the mechanisms through which the Miller/Coors merger affects market outcomes. We focus on (i) "unilateral effects" resulting from the internalization of competition between Miller and Coors; (ii) "coordinated effects" that we capture with the nonzero post-merger κ parameter; and (iii) marginal cost reductions from the merger efficiencies. Our analysis suggests that the raw data reflect coordinated effects and cost reductions. As a result this is our interpretation of the data. Thus, we recompute equilibrium with the RCNL-1 specification under four counterfactual scenarios:

- The merger does not occur.
- The merger occurs with efficiencies and without coordinated effects.
- The merger occurs without efficiencies and without coordinated effects.

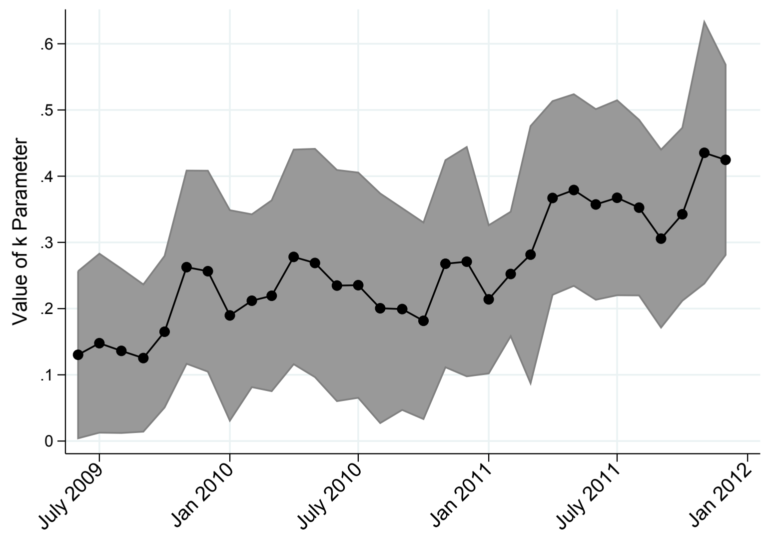

Figure 3: Time-Varying Estimates of the κ Parameter

Notes: The figure plots point estimates of κ for each post-merger month-year combination, along with a 95% confidence interval. Estimation is with the method-of-moments and uses the results generated with the RCNL-1 demand specification. There are 94,656 observations at the brand-size-region-month-year level. The sample excludes the months between June 2008 and May 2009. The marginal cost function includes the baseline cost shifters, and product (i.e., brand$\times$size), period, and region fixed effects. Instruments include an indicator for ABI products in the post-merger periods and mean income. The confidence interval is calculated with standard errors that are clustered by region.

- The merger occurs without efficiencies and with coordinated effects.

Figure 4 shows the prices of Miller Lite 12-packs under each of the five scenarios. Prices in the "No Merger" scenario are substantially lower than the raw data, and appear to roughly track the pre-merger trend. The unilateral effects and marginal cost reductions roughly offset, so that prices in the "No Merger" and "Unilateral, Efficiencies" scenarios are quite similar. The magnitudes of the unilateral effects and the efficiencies both are large, as evidenced by the higher prices that arise in the "Unilateral, No Efficiencies" scenario. Finally, coordinated effects are strong enough that prices increase even given the large marginal cost reductions: the merger increases prices by about fifty cents on average per 12-pack, relative to the "No Merger" scenario.

Figure 5 shows the prices of Bud Light 12-packs. We omit the "Unilateral, Efficiencies" scenario from the graph because they are nearly exactly the same as the "No Merger" scenario. (This is because unilateral effects were almost entirely offset by efficiencies for Miller and Coors, and thus have a very small impact on Bud Light prices.) As shown, observed prices in the post-merger periods exceed substantially the prices that arise in the

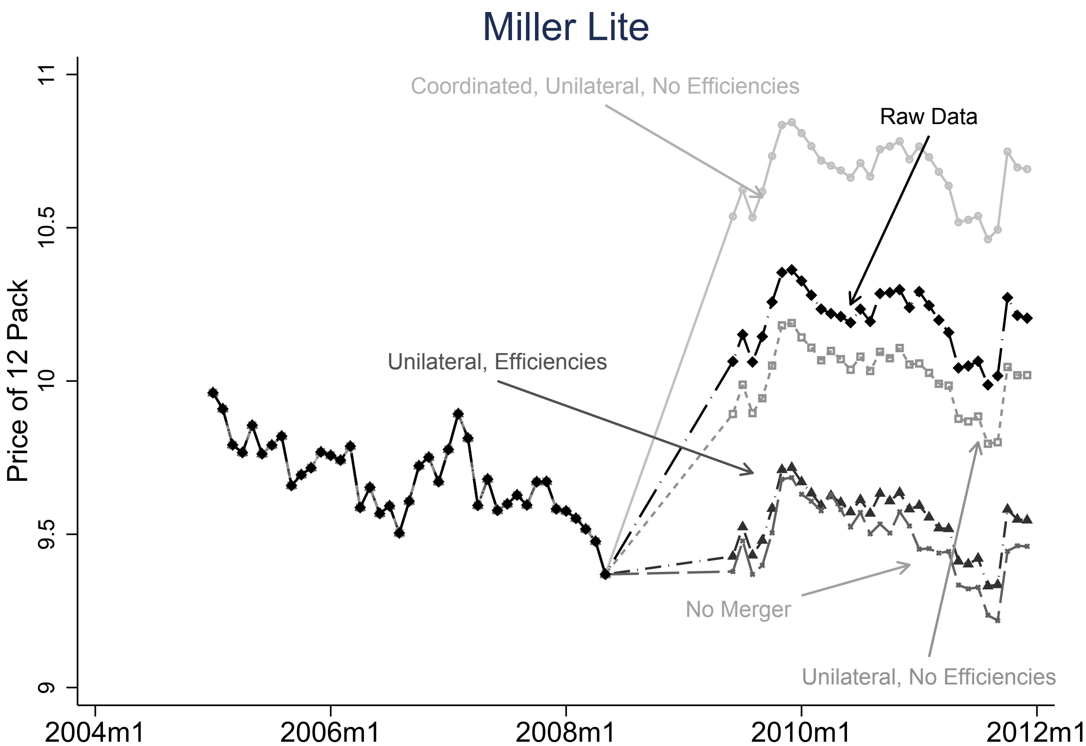

Figure 4: Counterfactuals Prices for Miller Lite

Notes: The figure plots the average retail prices of Miller Lite 12-Packs in the raw data and under four different counterfactual scenarios. Each dot represents the average prices across the 39 regions.

"No Merger" baseline, and the price increases are almost entirely due to coordinated effects.

Table 10 provides the mean retail prices and markups of ABI, Miller, and Coors 12-packs, along with selected welfare statistics based on the complete dataset. All numbers are for 2011, the final year of the sample. A comparison of columns (i) and (v) reveals that the merger increases ABI prices from \$9.43 in the "No Merger" scenario to \$10.03 in the raw data, while Miller prices increase from \$8.19 to \$8.94, and Coors prices increase from \$9.26 to \$10.18. The analogous comparison of markups shows smaller increases for ABI than for Miller and Coors because only the latter brands benefit from marginal cost reductions. A comparison of columns (i) and (iii) reveals that the observed prices of these firms are 6-8 percent higher than they would have been under Nash-Bertrand competition, and markups are 17-18 percent higher.

We now turn to the welfare statistics. All numbers shown are percentage differences relative to the "No Merger" counterfactual in which the Miller/Coors merger does not occur. Column (i) shows that the merger increases producer surplus by 22.1 percent relative to the

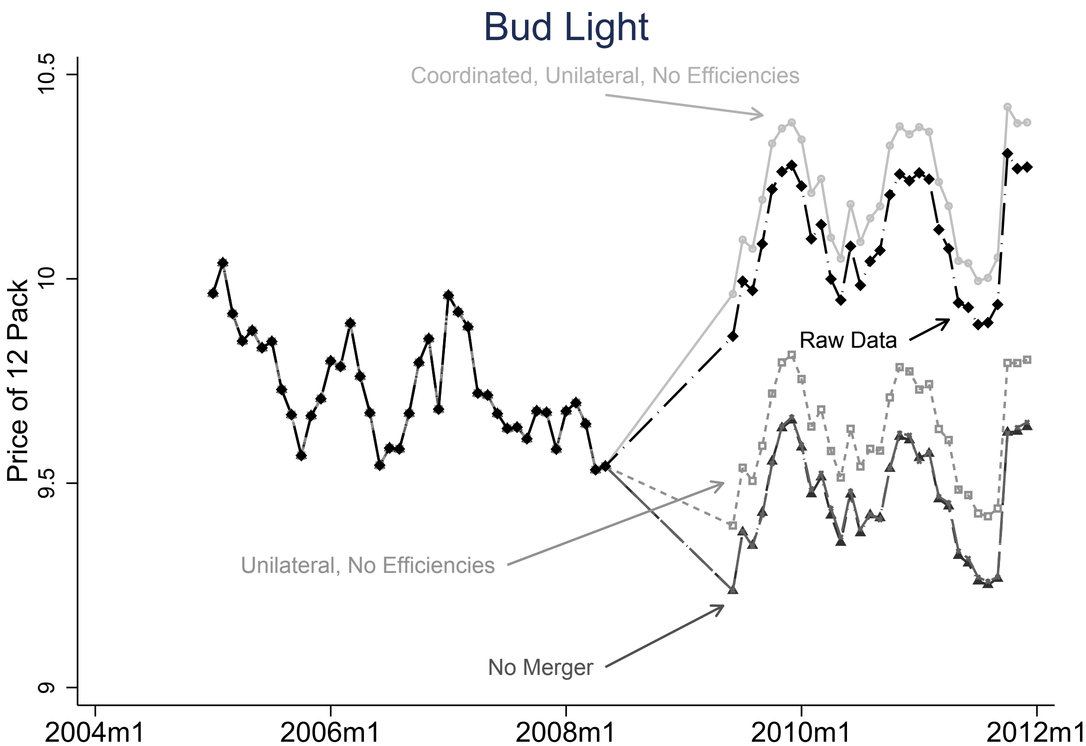

Figure 5: Counterfactuals Prices for Bud Light

Notes: The figure plots the average retail prices of Bud Light 12-Packs in the raw data and under three different counterfactual scenarios. Each dot represents the average prices across the 39 regions.

no merger baseline, and comparing to column (iii) reveals that a little more than half of these gains are due to coordination. Column (i) also shows that the merger reduces consumer surplus by 3.7 percent relative to the no merger baseline. The consumer surplus effects vary substantially and predictably with the roles of coordination and efficiencies. For example, column (iii) shows that without coordinated effects (but with efficiencies) consumer surplus falls by only 0.2 percent due to the merger. This may well have been the scenario deemed most likely by the DOJ in its decision to clear the merger. Lastly, column (i) shows that the merger increases total surplus by 1.3 percent relative to the no merger baseline, so the antitrust clearance of the merger could be justified on a total welfare standard.

7 Conclusion

This article summarizes our empirical investigation into the economic effects of the MillerCoors joint venture in the U.S. brewing industry. That the prices of MillerCoors and ABI

Table 10: Results from Counterfactual Analysis

Coordinated Effects:	yes	yes	no	no	no
Unilateral Effects:	yes	yes	yes	yes	no
Efficiencies:	yes	no	yes	no	no
	(i)	(ii)	(iii)	(iv)	(v)
			Retail Prices		
ABI	10.03	10.14	9.38	9.55	9.43
Miller	8.94	9.37	8.28	8.72	8.19
Coors	10.18	10.85	9.56	10.22	9.26
			Brewer Markups		
ABI	4.45	4.56	3.81	3.97	3.84
Miller	4.52	4.32	3.83	3.63	3.05
Coors	4.25	4.06	3.61	3.41	2.68
			Welfare Statistics		
Producer Surplus	22.1%	19.1%	10.3%	8.2%	·
ABI	10.3%	19.8%	-0.08%	9.3%	·
Miller	37.8%	20.2%	24.6%	9.1%	·
Coors	47.8%	12.9%	34.7%	3.5%	·
Consumer Surplus	-3.7%	-5.3%	-0.2%	-2.1%	·
Total Surplus	1.3%	-0.6%	1.8%	-0.1%	·

Notes: The table provides volume-weighted mean prices and markups, separately for 12-pack flagship brands of ABI, Miller, and Coors, under five different economic scenarios. Also shown are the percentage changes in producer surplus, consumer surplus, and total surplus, relative to the "No Merger" scenario. The welfare statistics are calculated using the complete dataset (i.e., all products). Column (i) is based on raw data and supply-side parameter estimates. Columns (ii)-(v) show results from counterfactual scenarios. The numbers in column (ii) are computed assuming the merger occurs with coordinated and unilateral effects but without efficiencies. The numbers in column (iii) are computed assuming the merger occurs with unilateral effects and efficiencies but no coordinated effects. The numbers in column (iv) are computed assuming the merger occurs with unilateral effects but without efficiencies or coordinated effects. Lastly, the numbers in column (v) are computed assuming that the Miller/Coors merger does not occur. All statistics are for 2011.

increase after the Miller/Coors merger, both in absolute terms and relative to their competitors, is visually evident and confirmed with econometric analysis. The magnitude of the ABI price increase, in particular, is difficult to explain with the standard model of differentiated-products Nash-Bertrand competition. Indeed, if a parameter is added to the standard model that allows MillerCoors and ABI to partially internalize their pricing externality in the post-merger periods, then Nash-Bertrand competition is rejected.

One plausible interpretation of this result is that the merger had coordinated effects.

This would reinforce the expressed view of antitrust agencies that mergers can soften the intensity of price competition between the merging firm and its remaining competitors. It is somewhat novel in the empirical literature, which has focused much more on understanding how mergers affect the unilateral pricing incentive of firms, holding the equilibrium concept fixed (e.g., Deneckere and Davidson (1985); Berry and Pakes (1993); Hausman, Leonard and Zona (1994); Werden and Froeb (1994); Nevo (2000a); Jaffe and Weyl (2013)). We hope that the results developed herein help motivate additional research on the coordinated effect of mergers so as to extend the practical relevance of the literature.

Our empirical analysis does not inform *why* the Miller/Coors merger may have had these coordinated effects. Indeed, one challenge for future research is understanding the conditions under which consolidation either enables collusion or exacerbates the impact of collusion. That said, the U.S. brewing industry does exhibit many of the characteristics that the Merger Guidelines enumerate as contributing to the likelihood of coordinated effects. Retail prices are observable, and ABI and MillerCoors may also gain visibility into wholesale prices through their interactions with wholesalers and retailers. Individual sales are small and frequent, which means that firms may be more easily deterred from making competitive initiatives because the short-term gain is smaller. That market demand is inelastic suggests large gains from coordination. The bargaining power of retailers is limited by the lack of viable private-label store brands and the regulatory prohibition on slotting allowances (which makes it harder for retailers to discipline coordination by auctioning shelf space).

Some business-to-consumer markets are broadly similar to brewing along many of these dimensions; airlines, mobile phone service, and gasoline stations come to mind. Within this class, one possibility is that price leadership (or focal point pricing more generally) became more feasible in the brewing industry after the Miller/Coors merger reduced heterogeneity in distribution costs among the major domestic brewers; such an effect is discussed in the Ivaldi et al (2003) report on coordinated effects to DG Competition. Alternatively, it could be that the increase in concentration alone was sufficient to facilitate coordination. That said, the generality of the results is limited. Beer is distinct among branded consumer products due to the regulatory prohibition on slotting allowances and the lack of viable private-label store brands. Business-to-business markets are more likely to feature privately negotiated contracts, powerful buyers, and other characteristics that make coordination more difficult. While we interpret our results as providing support for the possibility of coordinated effects, prospective analyses of mergers should be grounded in the relevant industry details.

References

Abreu, Dilip, "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 1988, *56* (2), 383–96.

Ascher, Bernard, "Global Beer: The Road to Monopoly," 2012. American Antitrust Institute monograph.

Ashenfelter, Orley, Daniel Hosken, and Matthew C. Weinberg, "Did Robert Bork Understate the Competitive Impact of Mergers? Evidence from Consummated Mergers," *Journal of Law and Economics*, 2014, *57* (S3), S67–S100.

____ , ____ , and ____ , "Efficiencies Brewed: Pricing and Consolidation in U.S. Brewing," *RAND Journal of Economics*, 2015, *46* (2), 328–361. forthcoming.

Asker, John, "Diagnosing Foreclosure due to Exclusive Dealing," *Journal of Industrial Economics*, 2016. Forthcoming.

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

____ and Ariel Pakes, "Some Applications and Limitations of Recent Advances in Empirical Industrial Organization: Merger Analysis," *American Economic Review*, 1993, *83* (2), 247–252.

____ and Philip A. Haile, "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, 2014, *82* (5), 1749–1798.

____ , James Levinsohn, and Ariel Pakes, "Automobile Prices in Market Equilibrium," *Econometrica*, July 1995, *63* (4), 847–890.

Bhattacharya, Debopam, "Asymptotic Inference from Multi-Stage Samples," *Journal of Econometrics*, 2005, *126* (1), 145 – 171.

Bjornerstedt, Jonas and Frank Verboven, "Does Merger Simulation Work? A 'Natural Experiment' in the Swedish Analgesics Market," *American Economic Journal: Applied Economics*, 2015. Forthcoming.

Bresnahan, Timothy, "The Oligopoly Solution Concept is Identified," *Economics Letters*, 1982, *10*, 82–97.

——, "Competition and Collusion in the American Automobile Oligopoly: The 1955 Price War," *Journal of Industrial Economics*, 1987, *35* (4), 457–482.

——, "Empirical Studies of Industries with Market Power," 1989. In *Handbook of Industrial Organization*, R. Schmalensee and R. Willig (eds), Vol. 2. North-Holland, Amsterdam.

Bronnenberg, Bart J., Michael W. Kruger, and Carl F. Mela, "The IRI Marketing Data Set," *Marketing Science*, 2008, *27* (4), 745–748.

Cardell, Scott N., "Variance Components Structures for the Extreme-Value and Logistic Distributions with Applications to Models of Heterogeneity," *Econometric Theory*, 1997, *13*, 185–213.

Ciliberto, Federico and Jonathan W. Williams, "Does Mulitmarket Contact Facilitate Tacit Collusion? Inference on Conduct Parameters in the Airline Industry," *RAND Journal of Economics*, 2014, *45* (4), 764–791.

Conlon, Christopher T. and Nirupama S. Rao, "The Price of Liquor is Too Damn High: The Effects of Post and Hold Pricing," 2016.

Corts, Kenneth S., "Conduct Parameters and the Measurement of Market Power," *Journal of Econometrics*, 1999, *88*, 227–225.

Dafny, Leemore S., Mark Duggan, and Subramaniam Ramanarayanan, "Paying a Premium on Your Premium? Consolidation in the U.S. Health Insurance Industry," *American Economic Review*, 2012, *102* (2), 1161–1185.

Deneckere, Raymond and Carl Davidson, "Incentives to Form Coalitions with Bertrand Competition," *The RAND Journal of Economics*, 1985, *16*, 473–86.

Friedman, James W., "A Non-cooperative Equilibrium for Supergames," *Review of Economic Studies*, 1971, *38* (1), 1–12.

Gilbert, Richard and Hillary Greene, "Merging Innovation into Antitrust Agency Enforcement of the Clayton Act," Presented at the Eigth Annual Searle Center Conference on Antitrust Economics and Competition Policy September 2015.

Goldberg, Pinelopi K. and Rebecca Hellerstein, "A Structural Approach to Identifying the Source of Local-Currency Price Stability," *Review of Economic Studies*, 2013, *80* (1), 175–210.

Grennan, Matthew, "Price Discrimination and Bargaining: Empirical Evidence from Medical Devices," *American Economic Review*, 2013, *103* (1), 145–177.

Grigolon, Laura and Frank Verboven, "Nested Logit or Random Coefficient Logit? A Comparison of Alternative Discrete Choice Models of Product Differentiation," *Review of Economics and Statistics*, 2014, *96* (5), 916–935.

Hansen, Lars P., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 1982, *50*, 1029–1054.

Hausman, Jerry, Gregory K. Leonard, and J. Douglas Zona, "Competitive Analysis with Differentiated Products," *Annales D'Economie et de Statistique*, 1994, *34* (1), 159–180.

Hellerstein, Rebecca, "Who Bears the Cost of a Change in the Exchage Rate? Pass-Through Accounting for the Case of Beer," *Journal of International Economics*, 2008, *76*, 14–32.

Hendel, Igal and Aviv Nevo, "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 2006, *74* (6), 1637–1673.

Houde, Jean-Francois, "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline," *American Economic Review*, 2012, *102* (5), 2147–82.

Igami, Mitsuru, "Market Power in International Commodity Trade: The Case of Coffee," *Journal of Industrial Economics*, 2015, *63* (2), 225–248.

Ivaldi, Marc, Bruno Jullien, Patrick Rey, Paul Seabright, and Jean Tirole, "The Economics of Tacit Collusion," March 2003. Final Report for DG Competition, European Commission.

Jaffe, Sonia and E. Glen Weyl, "The First Order Approach to Merger Analysis," *American Economic Journal: Microeconomics*, 2013, *5* (4), 188–213.

Lau, Lawrence J., "On Identifying the Degree of Competitiveness from Industry Price and Output Data," *Economics Letters*, 1982, *10*, 93–99.

McClain, Joe, "2011-2012 Annual Report," Technical Report, Beer Institute, Chicago Illinois 2012.

McFadden, Daniel, "Modelling of Choice of Residential Location," 1978. In *Spatial Interaction Theory and Residential Location*, F. Snickers, A. Karlquist, L. Lundquist, J. Weibull (eds), Amsterdam: North Holland.

Michel, Christian, "Identification and Estimation of Intra-Firm and Industry Competition via Ownership Change," 2016.

Miller, Nathan H., Marc Remer, Conor Ryan, and Gloria Sheu, "UPP as a Predictor of Merger Price Effects," 2016.

Nevo, Aviv, "Identification of the Oligopoly Solution Concept in a Differentiated Products Industry," *Economics Letters*, 1998, *59* (3), 391–395.

——, "Mergers with Differentiated Products: The Case of the Ready-to-Eat Cereal Industry," *The RAND Journal of Economics*, 2000, *31* (3), pp. 395–421.

——, "A Practitioner's Guide to Estimation of Random Coefficients Logit Models of Demand," *Journal of Economics and Management Strategy*, 2000, *9*, 513–548.

——, "Measuring Market Power in the Ready-to-Eat Cereal Industry," *Econometrica*, 2001, *69* (2), pp. 307–342.

Peters, Craig, "Evaluating the Performance of Merger Simulation: Evidence from the U.S. Airline Industry," *Journal of Law & Economics*, October 2006, *49* (2), 627–49.

Pinske, Joris and Margaret E. Slade, "Mergers, Brand Competition, and the Price of a Pint," *European Economic Review*, 2004, *48*, 617–643.

Porter, Robert, "A Study of Cartel Stability: The Joint Executive Committee, 1880-1886," *Bell Journal of Economics*, 1983, *14* (2), 301–314.

Rojas, Christian, "Price Competition in U.S. Brewing," *Journal of Industrial Economics*, 2008, *66* (1), 1–31.

Romeo, Charles, "Filling Out the Instrument Set in Mixed Logit Demand Systems for Aggregate Data," 2014.

——, "Incorporating Prior Information into a GMM Objective for Mixed Logit Demand Systems," *Journal of Industrial Economics*, 2016, *64*, 336–363.

Rotemberg, Julio J. and Garth Saloner, "A Supergame-Theoretical Model of Price Wars during Booms," *American Economic Review*, 1986, *76* (3), 390–407.

Shapiro, Brad, "Positive Spillovers and Free Riding in Advertising of Prescription Pharmaceuticals: The Case of Antidepressants," *Journal of Political Economy*, forthcoming.

Shapiro, Carl, "The 2010 Horizontal Merger Guidelines: From Hedgehog to Fox in Forty Years," *Antitrust Law Journal*, 2010, *77*, 49–107.

Slade, Margaret E., "Market Power and Joint Dominance in U.K. Brewing," *Journal of Industrial Economics*, 2004, *52* (1), 132–163.

Stock, James A. and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," 2005. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, D.W.K. Andrews and J.H. Stock (eds). New York: Cambridge University Press.

Sweeting, Andrew and Xuezhen Tao, "Dynamic Oligopoly Pricing with Asymmetric Information: Implications for Mergers," 2016.

Tremblay, Victor J. and Carol Horton Tremblay, *The U.S. Brewing Industry*, The MIT Press, 2005.

Weinberg, Matthew C. and Daniel Hosken, "Evidence on the Accuracy of Merger Simulations," *Review of Economics and Statistics*, 2013, *95* (5), 1584–1600.

Werden, Gregory J. and Luke Froeb, "The Effects of Mergers in Differentiated Products Industries: Logit Demand and Merger Policy," *Journal of Law, Economics, and Organization*, 1994, *10*, 407–426.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, 2nd ed., MIT Press, 2010.

Appendix

A Data

Table A.1 provides information on prices and revenue shares for major beer brands, based on the six months over January-June 2008. The brands are listed in the order of their revenue share. Our regression samples include Bud Light, Budweiser, Michelob, Michelob Light, Miller Lite, Miller Genuine Draft, Miller High Life, Coors Light, Coors, Corona Extra, Corona Extra Light, Heineken, and Heineken Light. These included brands account for 68 percent of all unit sales of SAB Miller, Molson Coors, ABI, Modelo, and Heineken. The most popular brands that we omit are regional brands (e.g., Yuengling Lager and Labatt Blue) or subpremium brands that sell at lower price points (e.g., Busch Light, Natural Light, Busch, Keystone, Natural Ice). Many of the subpremium brands are owned by ABI. We also exclude some brands that enter or exit during the sample period (e.g., Budweiser Select, Bud Light Lime). While brands in this last category could be incorporated, it would require brand-specific modifications to the demand system.

We restrict attention to 6-packs, 12-pack, and 24/30 packs. These sizes account for 75 percent of all unit sales among the brands that we consider. Table A.1 also provides the distribution of sales volume across these size categories for each brand listed. For example, 11 percent of Bud Lite is sold at 6-packs, 34 percent is sold as 12-packs, and 55 percent is sold as 24/30 packs. Because these numbers are weighted by volume, it can also be determined that more 12-packs are sold than 24/30 packs (on a unit basis). Domestic beers tend to be sold mostly as 12-packs and 24-packs, while imports tend to be sold mostly as 6-packs and 12-packs. This amplifies the average price differences shown because smaller package sizes tend to be more expensive on a per-volume basis.

We restrict attention to 39 of the 47 geographic regions in the IRI academic database, dropping a handful of regions in which either few supermarkets are licensed to sell beer or supermarkets are restricted to selling low alcohol beer.[16] Table A.2 provides the region-specific HHI in 2011, as well as the pre-merger predicted change in HHI (ΔHHI) as of January-May 2008. There is a fair amount of cross-sectional variation in concentration. 23

[16]The regions included in our sample are: Atlanta, Birmingham/Montgomery, Boston, Buffalo/Rochester, Charlotte, Chicago, Cleveland, Dallas, Des Moines, Detroit, Grand Rapids, Green Bay, Hartford, Houston, Indianapolis, Knoxville, Los Angeles, Milwaukee, Mississippi, New Orleans, New York, Omaha, Peoria/Springfield, Phoenix, Portland OR, Raleigh/Durham, Richmond/Norfolk, Roanoke, Sacramento, San Diego, San Francisco, Seattle/Tacoma, South Carolina, Spokane, St. Louis, Syracuse, Toledo, Washington D.C., and West Texas/New Mexico.

Table A.1: Average Prices and Revenue Shares: January-June 2008

			Fraction of Sales		
Brand	Price	Revenue Share	6-Pack	12-Pack	24/30 Packs
Bud Light	9.45	12.2	10.7	34.0	55.3
Miller Lite	9.42	7.8	8.0	31.1	61.0
Coors Light	9.47	6.0	10.2	33.5	56.3
Budweiser	9.46	5.9	14.4	34.0	60.4
Corona Extra	14.54	5.7	17.7	66.8	15.5
Heineken	14.65	3.3	24.9	70.8	4.3
Busch Light	6.95	3.1	2.1	23.5	74.5
Natural Light	6.48	2.9	5.9	31.5	62.6
Yuengling Lager	9.61	2.7	19.0	60.2	20.8
Corona Light	14.70	2.1	24.4	72.1	3.5
Michelob Ultra	10.91	2.1	24.5	72.5	3.0
Miller High Life	7.23	2.0	7.1	46.8	46.1
Busch	7.00	1.8	3.6	31.9	64.5
Miller Genuine Draft	9.45	1.4	15.5	39.7	44.8
Michelob Light	10.84	1.3	25.8	73.2	1.0
Labatt Blue	9.23	1.3	1.9	35.6	62.5
Keystone Light	6.33	1.2	0.3	19.9	79.8
Blue Moon	14.65	1.1	47.5	52.5	0.0
Budweiser Select	9.47	1.0	12.2	46.7	45.2
Heineken Light	14.87	1.0	24.1	74.1	1.8
Natural Ice	6.45	1.0	6.9	47.1	46.1
Pabst Blue Ribbon	6.99	0.8	3.4	54.2	42.4
Tecate	11.51	0.8	9.3	39.3	51.4
Modelo Especial	14.25	0.7	21.4	76.6	2.1
Coors	9.47	0.6	6.0	39.1	54.5
Bud Light Lime	12.93	0.5	46.8	53.2	0.0

Notes: The table provides summary statistics on the major beer brands. Price is the ratio of revenue to 144oz-equivalent unit sales. Revenue share is total revenue of the brand divided by total revenue in the beer category. The remaining three columns show the fraction of revenues derived from 6, 12, and 24/30 packs, respectively. Calculations are based on the IRI supermarket data from January through May 2008.

of the 39 regions have post-merger HHIs that are above the threshold of 2,500 that the Merger Guidelines recognize as delineating "highly concentrated" markets.

McClain (2012) reports that supermarkets account for 20 percent of off-premise beer sales. The other major sources of off-premise beer sales are liquor stores (38 percent), convenience stores (26 percent), mass retailers (6 percent) and drug stores (3 percent). The IRI Academic Database includes information on sales in drug stores. In the next appendix section, we show that retail price patterns in that channel are similar to those in supermarkets. We do not have data for the other channels.

Table A.2: HHI and Predicted Change in HHI by IRI Region

Region	HHI	ΔHHI	Region	HHI	ΔHHI
Atlanta	2,120	367	Birmingham/Montgomery	2,989	400
Boston	1,925	188	Buffalo/Rochester	1,439	376
Charlotte	2,867	436	Chicago	2,618	484
Cleveland	1,815	400	Dallas	2,860	715
Des Moines	3,171	275	Detroit	2,372	311
Grand Rapids	2,864	311	Green Bay	3,537	448
Hartford	2,717	220	Houston	2,602	295
Indianapolis	3,382	1,022	Knoxville	3,009	371
Los Angeles	1,851	249	Milwaukee	3,718	472
Mississippi	3,647	417	West Texas/New Mexico	2,981	362
New Orleans	2,879	475	New York	1,792	216
Omaha	3,104	318	Peoria/Springfield	3,077	555
Phoenix	2,625	424	Portland, OR	1,551	479
Raleigh/Durham	2,498	265	Richmond/Norfolk	2,599	325
Roanoke	2,929	450	Sacramento	1,672	296
San Diego	1,644	353	San Francisco	1,422	210
Seattle/Tacoma	1,558	370	South Carolina	3,413	368
Spokane	2,528	684	St Louis	3,694	143
Syracuse	1,641	313	Toledo	3,059	396
Washington DC	1,711	289			

Notes: The table shows the post-merger HHI, calculated as the sum of squared market shares in 2011, and the pre-merger predicted change in HHI (ΔHHI) based on market shares in the first five months of 2008. The market shares are calculated based on each brewer's share of total sales in the data. Data are not restricted to the brands/sizes studied in the empirical model, and the market shares do not incorporate the outside good.

B Descriptive Retail Price Regressions

This section addresses some questions may that arise from the descriptive regressions in Section 3, related to the store-level composition of the IRI data, the impact of promotions, and whether results extend beyond the supermarket channel. To do so, we apply the differences-in-differences specification shown in equation (1) to store-level data, replacing product×region fixed effects with product×store fixed effects. For brevity, we consider specification with product-specific trends and no other controls. The dependent variables include the average price, the frequency of promotions, the regular price, and the promotion price. Promotions are not observed, but we follow Hendel and Nevo (2006) and define a price as being promotional if it is less than 50% of the highest price in the preceding month.

Table B.1 provides the results. The basic results that we document in Section 3 hold if store-week data are used rather than region-month data (column 1). There is some evidence that promotions are less frequent after the merger, although this effect is less pronounced

Table B.1: Supplementary Descriptive Price Regressions

Dependent Variable:	Average Price	Promotion Indicator	Regular Price	Promotion Price	Average Price
Unit of Geography:	Store	Store	Store	Store	Region
Periodicity:	Weekly	Weekly	Weekly	Weekly	Monthly
Sector:	Supermarket	Supermarket	Supermarket	Supermarket	Drug Stores
$\mathbb{1}$\{MillerCoors\}	0.047	0.013	0.056	0.051	0.042
$\times\mathbb{1}$\{Post-Merger\}	(0.004)	(0.009)	(0.005)	(0.005)	(0.007)
$\mathbb{1}$\{ABI\}	0.038	0.019	0.046	0.038	0.042
$\times\mathbb{1}$\{Post-Merger\}	(0.005)	(0.011)	(0.005)	(0.004)	(0.006)
$\mathbb{1}$\{Post-Merger\}	-0.008	-0.037	-0.017	-0.020	-0.005
	(0.003)	(0.009)	(0.004)	(0.004)	(0.004)
Observations	15,408,503	15,408,503	12,085,773	3,322,730	100,587

Notes: Estimation is with OLS. Observations in the first four columns are at the brand-size-store-week-year level. Observations in the final column are at the brand-size-region-month-year level. All regressions include product (brand×size) fixed effects interacted with store fixed effects, as well as product-specific linear time trends. Standard errors are clustered at the region level and shown in parentheses.

for MillerCoors and ABI (column 2). The regular and promotion prices seem to change in similar ways over the sample period (column 3 and 4). Taken together, the results indicate that the most important effect of the merger is on the overall price level, rather than on the frequency or magnitude of promotions. Finally, similar average price results are obtained from the drug store sector (column 5). This can be seen graphically in Figure B.1. Average prices are more volatile due to relatively thinner sales, but the same empirical patterns are apparent. Ideally we would also be able to verify that prices increased at convenience stores and liquor stores as well, but we were unable to obtain scanner data for these retailers. That said, we would be surprised if wholesale prices increased very differently across retailers within a region, because they are legally required to buy from the same distributors.

C Numerical Analysis

We provide two numerical exercises in which we perturb the estimated demand derivatives and examine the implications for estimates of the κ parameter. This allows us to explore the robustness of the main econometric result to different demand conditions in a systematic manner, before addressing specific alternatives to the baseline demand specification.

In the first numerical exercise, we assess the extent to which the estimates of κ could be overstated if the baseline RCNL specification does too little to relax the IIA property of the

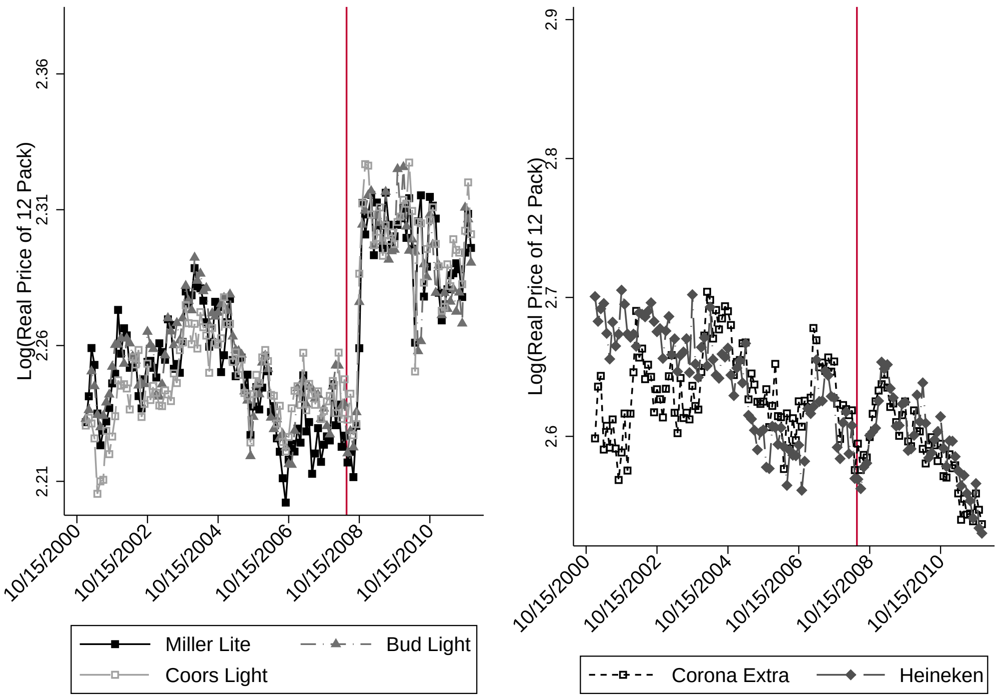

Figure B.1: Average Retail Prices of Flagship Brand 12-Packs: Drug Stores

Notes: The figure plots the national average price of a 12-pack over 2001-2011, separately for Bud Light, Miller Lite, Coors Light, Corona Extra and Heineken. The vertical axis is the natural log of the price in real 2010 dollars. The vertical bar drawn at June 2008 signifies the consummation of the Miller/Coors merger. Horizontal ticks are shown at October of each year due to an industry practice in which brewer prices are adjusted in early autumn.

logit demand system. To clarify this potential source of bias, consider that the presence of consumer heterogeneity likely results in some consumers that prefer domestic beer, and others that prefer imported beer. If this heterogeneity is not fully captured in the demand model then substitution between domestic beer and imports would be overstated, and substitution among domestic beers (e.g., between ABI and MillerCoors) would be understated. The supply-side implication is that the model then would understate the extent to which ABI's prices would increase with the Miller/Coors merger in Nash-Bertrand equilibrium. Because the κ parameter is identified based on whether observed ABI prices increase by more than what is predicted in Nash-Bertrand equilibrium, understating consumer heterogeneity in tastes for imports/exports thus would cause estimates of κ to be too large.

The numerical exercise involves scaling down the estimated demand derivatives between domestic beers and imports (and vice-versa) according to some parameter $\phi \in [0, 1]$. The lost substitution is reassigned to competing brands of the same type. If $\phi = 1$ this produces

markets in which there is zero substitution between domestic beers and imports, and if $\phi = 0$ then the estimated demand derivatives are unaffected. Regardless of ϕ, there is no effect on the diagonal of the demand derivatives matrix, so the own price elasticities are unchanged, and substitution to the outside good also is unchanged.

To be clear about the mathematics of the exercise, we reestimate the supply-side of the model plugging a scaled derivative matrix $\widetilde{\frac{\partial s_t}{\partial p_t}} = \widehat{\frac{\partial s_t}{\partial p_t}} + \Delta(\phi)$ into equation (12), where $\widehat{\frac{\partial s_t}{\partial p_t}}$ is the matrix of estimated derivatives and $\Delta(\phi)$ contains the adjustments. Consider a region-period combination with five products: Bud Light, Coors Light, Miller Lite, Corona Extra and Heineken, respectively. Let the elements of the estimated demand derivative matrix be

$$\widehat{\frac{\partial s_t}{\partial p_t}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{bmatrix}$$

where $a_{21} = \partial s_{2t}/\partial p_{1t}$. The adjustment matrix is given by

$$\Delta(\phi) = \begin{bmatrix} 0 & \frac{\phi(a_{42}+a_{52})a_{12}}{a_{12}+a_{32}} & \frac{\phi(a_{43}+a_{53})a_{13}}{a_{13}+a_{23}} & -\phi a_{14} & \phi a_{15} \\ \frac{\phi(a_{41}+a_{51})a_{21}}{a_{21}+a_{31}} & 0 & \frac{\phi(a_{43}+a_{53})a_{23}}{a_{13}+a_{23}} & -\phi a_{24} & -\phi a_{25} \\ \frac{\phi(a_{41}+a_{51})a_{31}}{a_{21}+a_{31}} & \frac{\phi(a_{42}+a_{52})a_{32}}{a_{12}+a_{32}} & 0 & -\phi a_{34} & -\phi a_{35} \\ -\phi a_{41} & -\phi a_{42} & -\phi a_{43} & 0 & \phi(a_{15}+a_{25}+a_{35}) \\ -\phi a_{51} & -\phi a_{52} & -\phi a_{53} & \phi(a_{14}+a_{24}+a_{34}) & 0 \end{bmatrix}$$

The first column contains adjustments to the share derivatives with respect to the Bud Light price. The diagonal element is zero, ensuring that the own price derivative (and elasticity) is unaffected. The fourth and fifth elements show the reduced substitution to Corona Extra and Heineken. The total lost substitution is $\phi(a_{41} + a_{51})$, and this is reassigned to Coors Light and Miller Lite. Some assumption on the allocation between these domestic brands is required, and we weight by the magnitude of the estimated substitution. This explains the second and third elements. The other columns are analogous. Each of the columns sums to zero, so that substitution to the outside good is unaffected.

Table C.1 provides the results of the first numerical exercise. We show results generated with the estimated demand derivatives of RCNL-1 (panel A) and RCNL-3 (panel B), and $\phi = 1.00, 0.80, \ldots, 0.20$. In each case, the estimate of κ is diminished relative to the baseline estimates of 0.241 (RCNL-1) and 0.291 (RCNL-3). This is because $\phi > 0$ results in greater

Table C.1: Supply-Side Estimates with Adjusted Demand Derivatives (1)

	Panel A: RCNL-1 Specification				
	$\phi = 1.00$	$\phi = 0.80$	$\phi = 0.60$	$\phi = 0.40$	$\phi = 0.20$
Post-Merger Internalization	0.176	0.197	0.214	0.231	0.247
of Coalition Pricing	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)
	Panel B: RCNL-3 Specification				
	$\phi = 1.00$	$\phi = 0.80$	$\phi = 0.60$	$\phi = 0.40$	$\phi = 0.20$
Post-Merger Internalization	0.206	0.223	0.239	0.254	0.270
of Coalition Pricing	(0.019)	(0.019)	(0.020)	(0.021)	(0.022)

Notes: The table shows supply-side results obtained with RCNL demand derivative matrices that are adjusted by $\phi = 1.00, 0.80, 0.60, 0.40, 0.20$ prior to supply-side estimation. This reallocates substitution between domestic and import brands to substitution among brands of the same type; substitution across types is eliminated if $\phi = 0$. There are 94,656 observations at the brand-size-region-month-year level. All regressions incorporate a marginal cost function with the baseline marginal cost shifters and fixed effects. Standard errors clustered by region and shown in parentheses. The standard errors are not adjusted to account for the incorporation of demand-side estimates.

substitution between ABI and MillerCoors, and thus a greater price increase for ABI due to unilateral effects (i.e., in Nash-Bertand equilibrium). If substitution between domestic and import brands is completely eliminated ($\phi = 1.00$), the κ estimates are reduced to 0.176 and 0.206. Thus, if the baseline demand specification does not fully capture consumer heterogeneity in tastes for imports, this explains at most 36% of the baseline κ estimate.

In our second exercise, we scale the entire estimated demand derivative matrix by a single constant, ψ, that we normalize at different levels ($\psi = 0.70, 0.80, ..., 1.20, 1.30$). This makes demand less elastic if $\psi < 1$ and more elastic if $\psi > 1$. Adapting the brewer first order conditions shows that this is equivalent to multiplying brewer markups by $1/\psi$:

$$p_t = mc_t - \left[\Omega_t(\kappa) \circ \left(\psi \frac{\partial s_t(p_t;\theta)}{\partial p_t}\right)^T\right]^{-1} s_t(p_t;\theta) \tag{C.1}$$

$$= mc_t - (1/\psi)\left[\Omega_t(\kappa) \circ \left(\frac{\partial s_t(p_t;\theta)}{\partial p_t}\right)^T\right]^{-1} s_t(p_t;\theta) \tag{C.2}$$

The numerical adjustment does not affect relative substitution patterns between products (including the outside good). Diversion is unchanged. However, it does allows us to investigate the how supply-side inferences are affected by the overall demand elasticity. We estimate the supply-side with the same methods – the demand derivatives from column (i) of Table 4 simply are adjusted before being entered into equation (12).

Table C.2 provides the results of the second exercise based on the derivatives of RCNL-

Table C.2: Supply-Side Estimates with Adjusted Demand Derivatives (2)

	$\psi = 0.70$	$\psi = 0.80$	$\psi = 0.90$	$\psi = 1.10$	$\psi = 1.20$	$\psi = 1.30$
Panel A: RCNL-1 Specification						
Post-Merger Internalization of Coalition Pricing	0.183 (0.022)	0.211 (0.023)	0.238 (0.025)	0.289 (0.027)	0.313 (0.028)	0.336 (0.028)
Panel B: RCNL-3 Specification	$\psi = 0.70$	$\psi = 0.80$	$\psi = 0.90$	$\psi = 1.10$	$\psi = 1.20$	$\psi = 1.30$
Post-Merger Internalization of Coalition Pricing	0.212 (0.020)	0.238 (0.021)	0.262 (0.022)	0.309 (0.024)	0.331 (0.025)	0.352 (0.026)

Notes: The table shows supply-side results obtained with RCNL demand derivative matrices that are multiplied/scaled by some amount $\psi = 0.70, 0.80, ..., 1.30$. This damps or amplifies the magnitude of substitution but maintains relative substitution patterns. There are 94,656 observations at the brand-size-region-month-year level. All regressions incorporate a marginal cost function with the baseline marginal cost shifters and fixed effects. Standard errors clustered by region and shown in parentheses. The standard errors are not adjusted to account for the incorporation of demand-side estimates.

1 (panel A) and RCNL-3 (panel B). We obtain smaller estimates of κ if demand is less elastic (i.e., if brewer markups are larger) and larger estimates of κ if demand is more elastic. The estimates range from 0.152 and 0.216 ($\psi = 0.70$) to 0.225 and 0.357 ($\psi = 1.30$). The null of post-merger Nash-Bertrand pricing is rejected in each instance. Thus, our main econometric finding is robust across a range of elasticities centered around the baseline point estimates. Alternative specifications of demand that result in higher or lower elasticities, but do not affect relative substitution patterns, should not be expected to change the main result.

D Robustness

D.1 Power of the demand-side instruments

In this section, we evaluate the relevance of the excluded instruments in the RCNL-1 and RCNL-3 specifications shown in the baseline demand results. These instruments provide the exogenous variation needed for estimation of the parameters $\theta_0^D = (\alpha, \Pi, \rho)$. We use the first-stages (one per parameter) of the Gauss-Newton regression described in Gandhi and Houde (2015). This provides tests of the null hypotheses that the excluded instruments are irrelevant for explaining variation in the partial derivative of the structural error term with respect to each parameter. Table D.1 reports the resulting partial F-statistics and the "Angrist-Pischke" F-statistics. Most of these test statistics well exceed the rule-of-thumb level of ten. The relevance of the instruments for the nesting parameter is below this threshold, so we explore the robustness of that parameter in greater detail below. We also

Table D.1: First-Stage Diagnostics for RCNL Models

Panel A: RCNL-1 (Column (ii) of Table 4)							
	$\frac{\partial\xi}{\partial\alpha}$	$\frac{\partial\xi}{\partial\Pi_1}$	$\frac{\partial\xi}{\partial\Pi_2}$	$\frac{\partial\xi}{\partial\Pi_3}$	$\frac{\partial\xi}{\partial\Pi_4}$	$\frac{\partial\xi}{\partial\Pi_5}$	$\frac{\partial\xi}{\partial\rho}$
Robust Partial F-statistic	26.78	253.28	154.95	265.42	-	-	21.92
Robust Angrist-Pischke F-statistic	26.78	242.97	126.79	125.69	-	-	5.48
Panel B: RCNL-3 (Column (iv) of Table 4)							
	$\frac{\partial\xi}{\partial\alpha}$	$\frac{\partial\xi}{\partial\Pi_1}$	$\frac{\partial\xi}{\partial\Pi_2}$	$\frac{\partial\xi}{\partial\Pi_3}$	$\frac{\partial\xi}{\partial\Pi_4}$	$\frac{\partial\xi}{\partial\Pi_5}$	$\frac{\partial\xi}{\partial\rho}$
Robust Partial F-statistic	26.78	-	153.35	166.26	236.78	87.44	20.45
Robust Angrist-Pischke F-statistic	14.31	-	123.74	131.14	240.00	58.47	4.46

Notes: F-statistics were calculated while clustering standard errors by city. For the RCNL-1 specification, the p-value for the Kleibergen-Paap test of the null of underidentification is 0.003 and the Cragg-Donald Wald F-statistic is 327.4. For the RCNL-3 specification, the p-value for the Kleibergen-Paap test of the null of underidentification is 0.009 and the Cragg-Donald Wald F-statistic is 187.24.

obtain the Cragg-Donald Wald F-statistics, which are high enough to reject at the .05 level the null hypothesis that the bias in the point estimates is greater than ten percent of the NLS bias, following the testing procedure in Stock and Yogo (2005).

The instruments generally affect the endogenous variables in the expected manner. To illustrate, we regress price and the log conditional share on the excluded instruments (together with product and time fixed effects). We omit some of the interactions to avoid difficulty in interpretation due to multicollinearity. These are the relevant first stages of the nested logit model shown in equation (7). The results are reported in Table D.2. The coefficients mostly take the expected signs. Prices are higher for ABI and MillerCoors in the post-merger periods, if the distance between the brewery and region is greater, and if income is higher. Conditional shares are smaller if distance is greater or there are more products. The conditional shares are larger if the sum of distances is greater (i.e., if competitors have higher costs) and, for imports, if income is higher.

We now examine the nesting parameter in greater detail, motivated in part by the first stage diagnostics discussed above. The parameter plays an important role in the demand system because it helps determine the degree of substitution between the inside goods and the outside good. To develop intuition, consider a standard logit model (i.e., the RCNL with $\rho = \Pi = 0$). The outside good share scales with the market size normalization and, because substitution is proportional to share, so too does diversion to the outside good. The nesting parameter softens this connection and allows for more realistic substitution patterns. Table D.3 provides a number of robustness checks. First, we reestimate demand under alternative market sizes. Recall that the baseline (region-specific) market size equals

Table D.2: Regression of Price and Conditional Shares on IVs

Dependent Variable:	Price	Log Conditional Share
$\mathbb{1}\{\text{ABI or MillerCoors}\} \times \mathbb{1}\{\text{PostMerger}\}$	0.832	-0.048
	(0.063)	(0.045)
Distance	0.082	-0.075
	(0.022)	(0.041)
Sum of Distances	0.003	0.002
	(0.003)	(0.001)
Number of Products	-0.023	-0.059
	(0.041)	(0.014)
Mean Income	0.074	-0.046
	(0.018)	(0.026)
Mean Income$\times\mathbb{1}\{\text{Import}\}$	-0.025	0.064
	(0.013)	(0.001)
Mean Income$\times$Calories	-0.0000	0.0002
	(0.0001)	(0.0001)
Mean Income$\times$Package Size	-0.003	0.001
	(0.001)	(0.001)

Notes: Estimation is with OLS. There are 94,656 observations at the brand-size-region-month-year level. The sample excludes the months between June 2008 and May 2009. All regressions include product (brand$\times$size) and period fixed effects. Standard errors clustered by region and shown in parentheses.

150% of the maximum observed sales. Columns (i) and (ii) show results based on scaling observed sales by 130% and 200%, respectively, to obtain the market size. The magnitude of the ρ estimate changes accordingly such that diversion to the outside good remains relatively constant. Demand is somewhat more elastic with larger market sizes, but we have confirmed that this does not meaningfully affect inferences on the supply-side. Columns (iii) and (iv) are estimated with a normalization that the nesting parameter is 0.70 and 0.60, respectively. We exclude instruments based on the number of products and summed distance, which are no longer needed for identification. As expected, lower values of ρ correspond to greater diversion to the outside good, but otherwise the demand results are mostly unchanged. Considered together, these results indicate that the nesting parameter responds as it should to the market size, and that the results hold over a plausible range of the nesting parameter.

D.2 Validity of the demand-side instruments

We now discuss some possible concerns about the validity of the instruments. We focus particularly on (i) the post-merger indicator for ABI and MillerCoors products, (ii) distance

Table D.3: Performance of the Nested Logit Parameter

ρ Estimated:		Yes	Yes	No	No
Market Size Scaler:		1.30	2.00	1.50	1.50
Variable	Parameter	(i)	(ii)	(iii)	(iv)
		Demand Estimates			
Price	α	-0.1101	-0.0933	-0.1550	-0.1961
		(0.0166)	(0.0129)	(0.0240)	(0.0314)
Nesting Parameter	ρ	0.7752	0.8407	0.70	0.60
		(0.0489)	(0.0358)		
Income$\times$Price	Π_1	0.0010	0.0006	0.0014	0.0019
		(0.0003)	(0.0002)	(0.0002)	(0.0003)
Income$\times$Constant	Π_2	0.0132	0.0129	0.0055	0.0018
		(0.0063)	(0.0041)	(0.0054)	(0.0060)
Income$\times$Calories	Π_3	0.0056	0.0030	0.0058	0.0063
		(0.0018)	(0.0014)	(0.0022)	(0.0026)
		Derived Demand Statistics			
Median Own Price Elasticity		-4.30	-5.35	-4.53	-4.15
Median Market Price Elasticity		-0.68	-0.69	-1.02	-1.25
Median Outside Diversion		16.19%	13.29%	23.24%	31.06%
$J-$Statistic		14.04	14.27	10.27	10.13

Notes: The table shows results for the RCNL-2 specification under various normalizations. There are 31,784 observations at the brand-size-region-quarter-year level. The sample excludes the quarters between June 2008 and May 2009. All regressions include product (brand$\times$size) and period fixed effects. The elasticity and diversion numbers represent medians among all the brand-size-region-month/quarter-year observations. Standard errors are clustered by region and shown in parentheses.

as a cost-shifter, and (iii) the number of products. The first two of these are important for identifying the price coefficient, and the last is important for the nesting parameter.

We start with the post-merger indicator, which implements the assumption that the changes in unobserved demand for ABI/MillerCoors, before versus after the merger, are not systematically different from changes in the unobserved demand for Modelo/Heineken. The validity of the instrument could come into question if the model understates the effect of the recession on preferences for (cheaper) domestic beer. While this is difficult to rule out entirely, the income interactions that we employ in the demand specifications allow the recession to impact demand in a sensible manner, and there is little else to be done empirically. A second possibility is ABI and MillerCoors increased their advertising in the post-merger periods. To investigate, we collected information on advertising spend from Kantar Media's AdSpender database, which also is used in a number of recent articles on advertising (e.g., Shapiro (forthcoming)). We run differences-in-differences regressions based on the specification shown in equation (1). Table D.4 shows the results. The advertising

Table D.4: Changes in Advertising by Parent Company

	(i)	(ii)	(iii)
$\mathbb{1}\{\text{ABI}\}\times\mathbb{1}\{\text{Post-Merger}\}$	2.720	2.252	-1.683
	(3.660)	(1.370)	(2.521)
$\mathbb{1}\{\text{MillerCoors}\}\times\mathbb{1}\{\text{Post-Merger}\}$	-2.142	-1.169	2.515
	(2.941)	(1.584)	(2.773)
1{Post-Merger}	-2.542	-0.119	-0.008
	(2.692)	(1.329)	(0.003)
Pre-Merger Mean Advertising	16.691	13.509	13.509
Product Trends	No	No	Yes
Observations	13,608	32,877	32,877

Notes: Estimation is with OLS. The dependent variable is real advertising spending (in thousands). Observations are at the brand-month-year level. Column (i) contains Bud Light, Coors Light, Miller Lite, Corona Extra, and Heineken. Columns (ii) and (iii) contain also contains Budweiser, Michelob Light, Michelob Ultra, Coors, Miller Genuine Draft, Miller High Life, Corona Light, and Heineken Premium Light. All regressions include brand fixed effects and a linear trend. Brand-specific trends are incorporated in column (iii). Heteroskedasticity robust standard errors are in parentheses.

spend of ABI increases relative to Corona/Heineken in the post-merger periods, while the relative spend of MillerCoors decreases (columns (i) and (ii)). Neither change is statistically significant. Further, the signs flip if product-specific trends are incorporated (column (iii)). Thus, there is little econometric evidence that advertising shifts consumer preferences toward the domestic firms.

The validity of the distance instrument could come into question if consumers that live near an ABI/MillerCoors brewery prefer those products. To investigate, we reestimate the model excluding the regions of St. Louis and Milwaukee, which feature nearby ABI and MillerCoors breweries, respectively. We are skeptical that preferences are affected at greater distances. The results are broadly consistent with those shown in the RCNL-2 specification of the baseline demand table: the median elasticity is -4.13 and diversion to the outside good is 18.06%. Finally, we note the numerical checks in Appendix C indicate that any bias introduced by the invalidity of distance or the post-merger indicator for ABI and MillerCoors would have to be large to materially affect the supply-side inferences.

The number of products primarily serves to identify the nesting parameter. Variation arises because not all products are sold in each region-period. Given the structure of the data (39 regions, 39 products, 72 months) there are 109,512 possible product-region-period observations of which 94,565 (or 86.43%) are realized. Because the more popular products

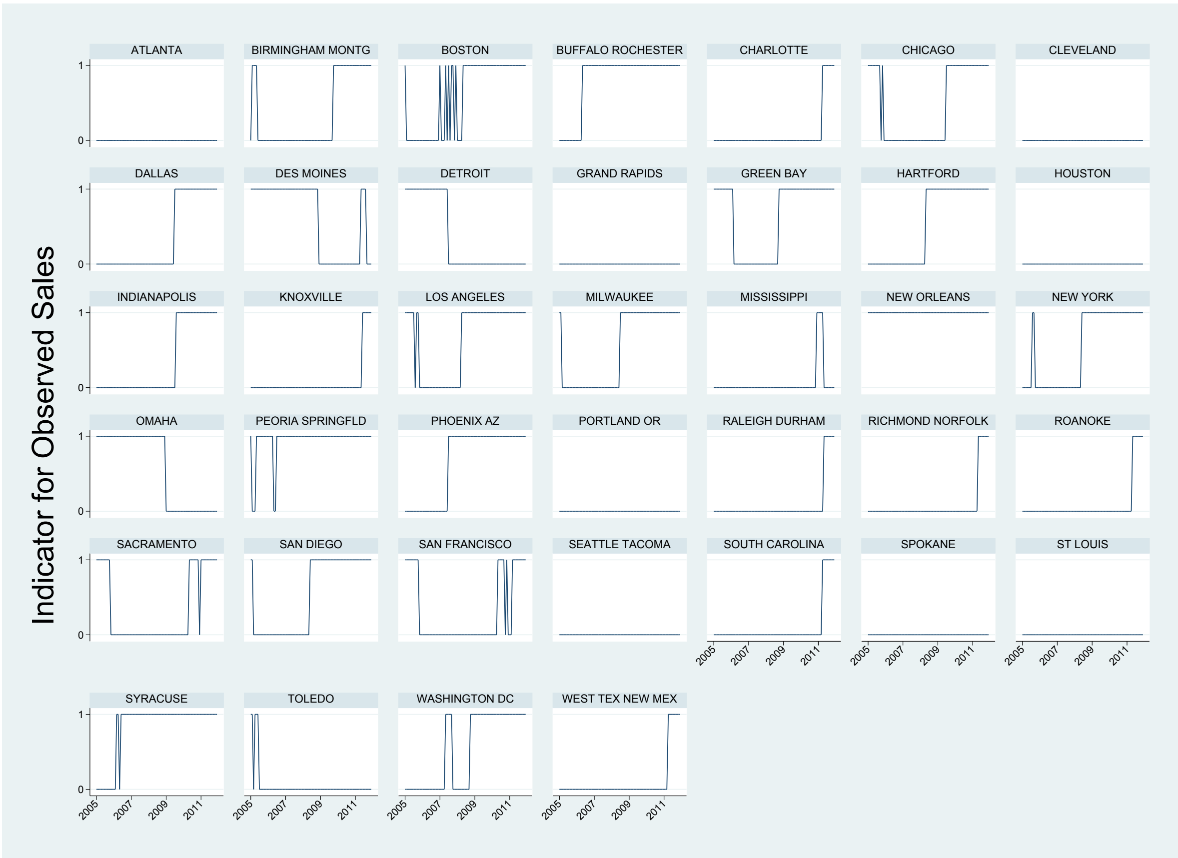

Figure D.1: Region-Periods with Observed Sales of Michelob 24-Packs

are sold in every region-period, variation arises due to the entry and exit of the less popular brands. This could raise concern if sales are sufficiently thin that entry/exit reflects measurement error in the data rather than true underlying variation in conditions. However, this does not seem to be the case. Consider the case of Michelob 24-packs, which are observed in 956 of the 2808 month-year combinations. Figure D.1 plots whether the product was observed in each month, by city. There are some instances in which the empirical patterns could be driven by measurement error (e.g., Boston), but more commonly the existence of sales is persistent. Analyses of other products are similar. Also of note are that results are robust when ρ is normalized and instrument set excludes the number of products.

D.3 Unobserved Trends

The baseline specifications incorporate time fixed effect to account for changes in cost and demand conditions that affect all products. Bias could arise to the extent that unobserved

changes in the market affect products differentially. Of particular concern would be unobserved changes that affect ABI differently than Modelo/Heineken, which would undermine our key identifying restrictions. One possible source of bias is the (gradual) increase in consumption of craft beers that occurs through the sample period. It is possible that craft beer is a closer competitor to imports than to ABI/MillerCoors, on the basis that craft and imported beers tend to have higher prices and sell in smaller package sizes.

As a robustness check, we reestimate the RCNL-2 specification using the shorter time windows of 2008-2009 and 2007-2010. We still exclude a year of data after the consummation of the merger, so the 2008-2009 sample contains four quarters of data and the 2007-2010 sample contains 12 quarters. The influence of unobserved trends should be smaller with these restrictions because the window of time over which the trends operate is abbreviated. The downside is that discarding data could reduce the precision of the estimates.

Table D.5 shows the results of estimation with the shorter samples. The demand-side coefficients are similar to those obtained from full dataset, though the Income×Price parameter is somewhat smaller and no longer statistically significant. The median price elasticities increase from -4.33 with the full dataset to -5.75 and -5.64, and diversion to the outside good also is slightly lower with the shorter time periods. The supply-side estimates of the κ parameter remain positive and statistically different than zero. The somewhat smaller value of κ that is obtained with the 2008-2009 sample is consistent results presented previously that coordination may have strengthened over the post-merger periods (e.g., recall Figure 3). The marginal cost parameters are not shown in the table, but are quite similar to those obtained from the full dataset. Overall, the analysis supports that the main econometric results are not unduly influenced by unobserved trends.

Lastly, Table D.6 provides estimates of an adjusted supply-side model that incorporates bilateral nominal exchange rates (domestic-currency units per unit of foreign currency) into the marginal cost function. This allows us to determine whether exchange rate fluctuations, which affect the relative costs of ABI and Modelo/Heineken, to affect estimates of κ parameter. We show results for each of the demand specifications in Table 4. The exchange rate parameter ranges from -0.041 to 0.760 in the RCNL specifications, and the other estimates are largely unaffected. We exclude the exchange rates from the baseline specifications because structural interpretation is difficult without decomposing the costs of Modelo/Heineken into a local non-traded component due to U.S. distribution and retail and a nonlocal component due to foreign production. Hellerstein (2008) and Goldberg and Hellerstein (2013) examine exchange rate pass-through in substantially more detail.

Table D.5: Estimation with Short Samples

Data Sample:		2008-2009	2007-2010
Variable	Parameter	(i)	(ii)
	Demand Estimates		
Price	α	-0.0837	-0.0946
		(0.0160)	(0.0127)
Nesting Parameter	ρ	0.8536	0.8437
		(0.0373)	(0.0391)
Income×Price	Π_1	0.0003	0.0007
		(0.0003)	(0.0003)
Income×Constant	Π_2	0.0145	0.0114
		(0.0062)	(0.0055)
Income×Calories	Π_3	0.0063	0.0048
		(0.0021)	(0.0017)
	Derived Demand Statistics		
Median Own Price Elasticity		-5.75	-5.64
Median Market Price Elasticity		-0.61	-0.65
Median Outside Diversion		10.91%	11.79%
$J-$Statistic		10.95	13.30
	Selected Supply Estimates		
Post-Merger Internalization		0.1565	0.2449
of Coalition Pricing		(0.0698)	(0.1185)

Notes: The table shows results for the RCNL-2 specification using different sample windows. There are 5,409 observations in the 2008-2009 sample and 16,143 observations in the 2007-2010 sample, at the brand-size-region-quarter-year level. The samples exclude the quarters between June 2008 and May 2009. All regressions include product (brand×size) and period fixed effects. The supply-side also includes the baseline marginal cost shifters and region fixed effects. The elasticity and diversion numbers represent medians among all the brand-size-region-month/quarter-year observations. Standard errors are clustered by region and shown in parentheses.

D.4 Additional results

Figure D.2 plots the marginal cost and demand time fixed effects in the RCNL-1 specification. The demand fixed effects are seasonal, reflecting that demand for beer is strongest in the summer and weakest in the winter. There is a gradual reduction in the willingness-to-pay for beer over the sample period. The marginal cost fixed effects decrease in the pre-merger periods. This is due in part to the price trends shown in Figure 1. The cost fixed effects stabilize over 2009-2010 and then drop again 2011. We suspect that the observed import price decreases over 2009-2010 can be explained from changes in macroeconomic conditions, but that the import price decreases in 2011 load onto the cost fixed effects. As shown above, results are robust to the exclusion of the final year (Table D.5).

Table D.6: Exchange Rates in the Marginal Cost Functions

Demand Model:	NL-1	RCNL-1	RCNL-2	RCNL-3	RCNL-4
Data Frequency:	monthly	monthly	quarterly	monthly	quarterly
	(i)	(ii)	(iii)	(iv)	(v)
Post-Merger Internalization	0.380	0.2654	0.249	0.289	0.345
of Coalition Pricing Externalities	(0.032)	(0.076)	(0.091)	(0.043)	(0.054)
			Marginal Cost Parameters		
MillerCoors×PostMerger	-0.616	-0.654	-0.649	-0.722	-0.526
	(0.040)	(0.050)	(0.056)	(0.042)	(0.041)
Distance	0.143	0.168	0.163	0.169	0.148
	(0.046)	(0.059)	(0.059)	(0.060)	(0.048)
Exchange Rate	0.744	0.389	-0.041	0.760	0.579
	(0.189)	(0.181)	(0.266)	(0.190)	(0.223)

Notes: The table shows the supply results with the exchange rate added to the marginal cost function. Estimation is with the method-of-moments. There are 94,656 observations at the brand-size-region-month-year level in columns (i), (iii), and (iv), and 31,784 observations at the brand-size-region-year-quarter level in column (ii). The sample excludes the months/quarters between June 2008 and May 2009. All regressions include product (brand×size) and period fixed effects. Standard errors are clustered by region and shown in parentheses.

Table D.3 provides a histogram of the marginal cost region fixed effects in the RCNL-1 specification. The fixed effect of Atlanta is normalized to zero. As shown, most region fixed effects range over $(-1, 0)$, with a few regions have lower costs on average. It is possible that these cost fixed effects are picking up some demand-side variability because we do not incorporate region fixed effects on the demand-side (otherwise the RCNL nesting parameter is too difficult to identify with any precision).

Table D.7 shows the mean demand elasticities that correspond to the RCNL-3 specification of Table 4. As shown, the higher-priced imports tend to have more elastic demand. This is mechanical because the specification does not allow income to rotate the demand curve. For example, with logit demand the own price elasticity of product j is simply $\alpha p_j(1 - s_j)$. If income is allowed to rotate demand, as in the RCNL-1 specification, then we find that the higher-priced imports still have more elastic demand, but by a smaller amount. The logit restriction that consumers substitute to other products in proportion to their market shares is somewhat relaxed. Looking along column (5) for Corona Extra, there is considerable heterogeneity in the cross-elasticities, but the same is not true in column (1) for Bud Light. Thus, while the coefficient estimates of RCNL-3 introduce qualitatively similar effects on substitution, they provide less meaningful departures from nested logit demand.

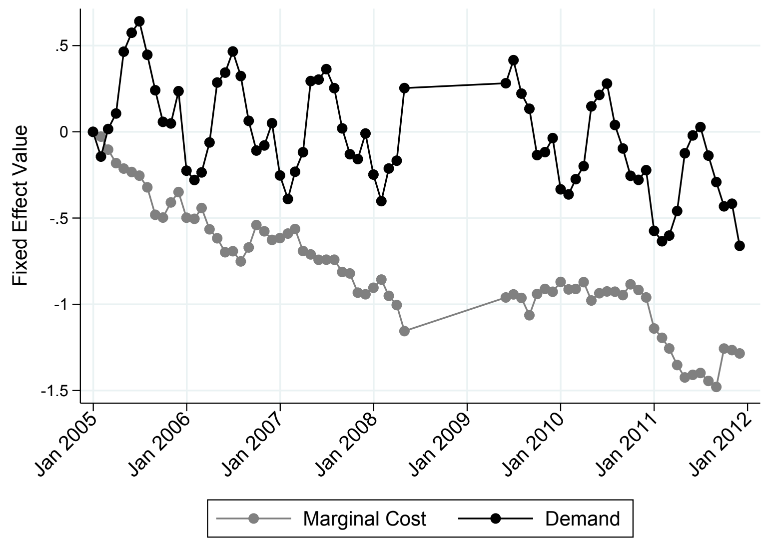

Figure D.2: Marginal Cost and Demand Time Fixed Effects in RCNL-1

Notes: The figure plots the estimated period fixed effects that affect preferences for the inside good (on the demand-side) and the marginal costs of all products (on the supply-side). We divide the demand fixed effects by the absolute value of the price coefficient, prior to plotting, so that the units are in dollars.

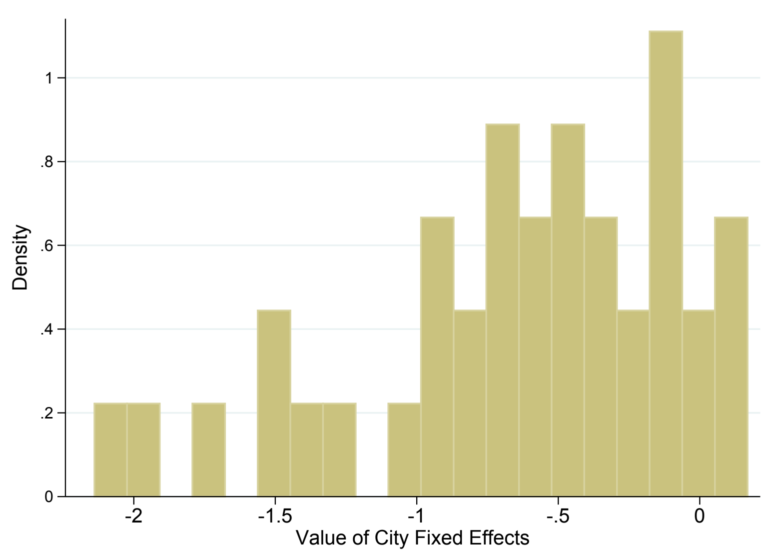

Figure D.3: Region Fixed Effects in the Marginal Cost Function in RCNL-1

Table D.7: Mean Elasticities for 12-Pack Products from RCNL-3

	Brand/Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
		Product-Specific Own and Cross Elasticities												
(1)	Bud Light	-3.819	0.141	0.017	0.150	0.242	0.094	0.158	0.045	0.034	0.108	0.039	0.060	0.164
(2)	Budweiser	0.284	-3.958	0.017	0.146	0.281	0.104	0.187	0.049	0.033	0.105	0.038	0.062	0.162
(3)	Coors	0.281	0.144	-4.094	0.144	0.284	0.103	0.188	0.049	0.033	0.104	0.038	0.062	0.161
(4)	Coors Light	0.288	0.141	0.017	-3.969	0.233	0.092	0.151	0.044	0.034	0.109	0.039	0.059	0.164
(5)	Corona Extra	0.274	0.153	0.019	0.135	-5.349	0.133	0.275	0.062	0.033	0.096	0.036	0.069	0.158
(6)	Corona Light	0.278	0.150	0.018	0.139	0.359	-5.637	0.245	0.057	0.033	0.099	0.037	0.067	0.160
(7)	Heineken	0.272	0.154	0.019	0.133	0.416	0.137	-5.506	0.063	0.033	0.095	0.036	0.070	0.158
(8)	Heineken Light	0.220	0.112	0.014	0.112	0.273	0.100	0.187	-5.724	0.023	0.078	0.027	0.053	0.127
(9)	Michelob	0.260	0.126	0.014	0.125	0.207	0.083	0.139	0.040	-4.452	0.099	0.032	0.053	0.152
(10)	Michelob Light	0.289	0.140	0.017	0.152	0.225	0.090	0.146	0.043	0.034	-4.401	0.039	0.059	0.164
(11)	Miller Gen. Draft	0.287	0.141	0.017	0.150	0.242	0.094	0.158	0.045	0.034	0.108	-4.075	0.060	0.163
(12)	Miller High Life	0.283	0.146	0.018	0.145	0.298	0.108	0.199	0.051	0.033	0.104	0.038	-3.095	0.162
(13)	Miller Lite	0.287	0.142	0.017	0.150	0.245	0.095	0.160	0.045	0.034	0.108	0.039	0.060	-3.949
(14)	Outside Good	0.015	0.007	0.001	0.008	0.012	0.005	0.008	0.002	0.002	0.006	0.002	0.003	0.009
		Total Cross Elasticities by Category												
	6-Packs	0.279	0.147	0.018	0.141	0.336	0.116	0.227	0.054	0.033	0.101	0.037	0.065	0.160
	12-Packs	0.281	0.144	0.017	0.143	0.285	0.107	0.194	0.051	0.033	0.103	0.037	0.062	0.161
	24-Packs	0.288	0.139	0.017	0.152	0.222	0.088	0.143	0.042	0.034	0.110	0.039	0.058	0.164
	Domestic	0.287	0.140	0.017	0.151	0.234	0.092	0.153	0.044	0.034	0.109	0.039	0.059	0.163
	Imported	0.274	0.153	0.019	0.135	0.398	0.134	0.270	0.061	0.033	0.096	0.036	0.069	0.158

Notes: The table provides mean elasticities of demand for 12-packs based on the RCNL-3 specification (column (iv) of Table 4). The cell entry in row i and column j is the percentage change in the quantity of product i with respect to the price of product j. Means are calculated across the year-month-region combinations.

Table D.8: Brewer Markups from RCNL-3

Brand	6-Packs Pre	6-Packs Post	12-Packs Pre	12-Packs Post	24-Packs Pre	24-Packs Post
Bud Light	4.00	4.74	4.03	4.78	4.05	4.81
Budweiser	3.97	4.70	4.00	4.74	4.03	4.78
Coors	2.73	4.56	2.75	4.62	2.77	4.66
Coors Light	2.76	4.63	2.77	4.67	2.78	4.71
Corona Extra	2.84	2.78	2.81	2.76	2.82	2.82
Corona Light	2.80	2.75	2.77	2.73	2.84	2.82
Heineken	2.64	2.63	2.63	2.61	2.74	2.72
Heineken Light	2.62	2.60	2.60	2.58	2.67	2.68
Michelob	4.00	4.85	4.06	4.88	4.04	5.04
Michelob Light	4.01	4.76	4.04	4.80	4.14	4.70
Miller Gen. Draft	3.18	4.64	3.18	4.66	3.19	4.72
Miller High Life	3.20	4.64	3.16	4.60	3.18	4.65
Miller Lite	3.17	4.61	3.18	4.66	3.19	4.70

Notes: The table provides average markups for each brand-size combination, separately for the pre-merger and post-merger periods, based on the RCNL-3 specification shown in column (iv) of Tables 4 and 6.

Table D.8 provides the brewer markups that arise from the RCNL-3 specification. Similar before/after comparisons can be made about the effect of the merger. MillerCoors markup increase due to greater market power and lower costs; ABI markups increase due to greater market power; and Modelo/Heineken markups are relatively unaffected. The markup for products of the same firm are quite similar with the RCNL-3 specification. This is again because the demand estimates to not introduce large departures from nested logit demand (which dictates that multiproduct firms equate markups across products).

D.5 Additional counterfactual results

We develop confidence bounds for selected welfare statistics using a partial bootstrap. We start by drawing 100 values from the estimated distribution of κ. This distribution reflects the impact of demand-side uncertainty on supply-side precision. We calculate impute markups based on the demand estimates and each random value of κ, and reestimate the linear marginal cost parameters. Finally, we recompute equilibrium in the post-merger periods in the counterfactual scenario that the merger does not occur. We use the RCNL-2 specification with quarterly data to reduce the computation burden. The results of RCNL-2 and RCNL-1 are quite similar so the confidence bounds are relevant to the counterfactual analysis presented in the body of the paper.

Figure D.4 plots prices for Miller Light and Bud Light 12-packs. The raw data are

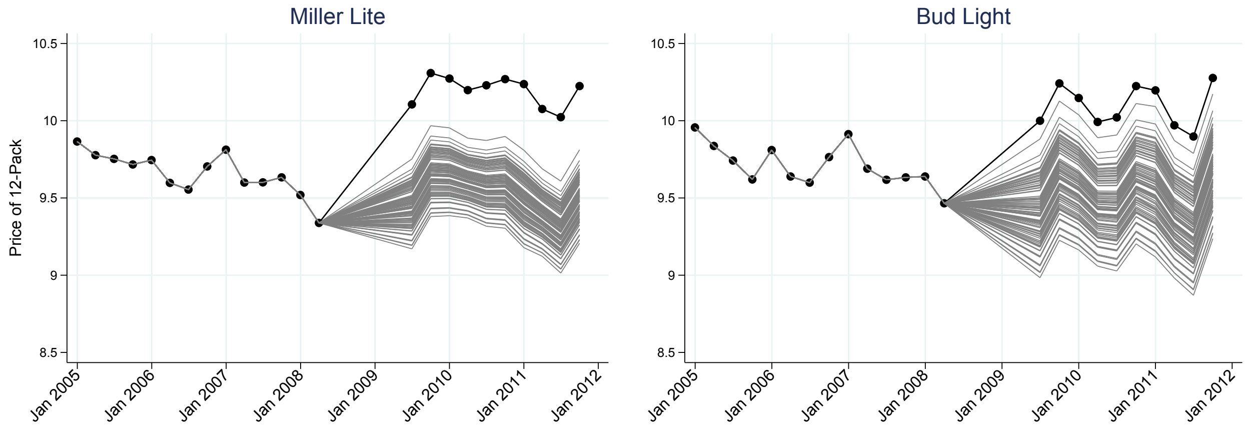

Figure D.4: Bootstrapped Counterfactual Prices

Notes: The figure plots the average prices of Miller Lite and Bud Light 12-packs in the raw data (black) and the average prices in the "no merger" counterfactual scenario under 100 random draws from the estimated distribution of supply-side parameters (gray). Results are based on the RCNL-2 specification with quarterly observations. The averages are across the 39 regions.

shown in black and each of the 100 "no merger" scenarios are shown in gray. The raw data exceed all of the simulated prices. We also calculate the percentage change in producer surplus, consumer surplus, and total surplus due to the merger. We find that producer surplus increases by 19% with a 90% confidence interval of (16%,21%). Consumer surplus decreases by 3.9% with a confidence interval of (−5.6%,−2.5%). Total surplus increases by 1.5% with a confidence interval of (−0.5%,3.0%). A one-sided test rejects the null hypothesis that total surplus decreases at the ten percent level.

An interesting feature of the counterfactual "no merger" scenario is that the Bud Light prices appear to fall somewhat in the post-merger periods relative to the pre-merger periods (e.g., see Figure 5). This happens in part due to the trends down in the demand and marginal cost time fixed effects shown in Figure D.2. To explore this further, we recompute equilibrium prices under all scenarios under the following adjustments:

- Post-merger time fixed effects take the value of the May 2008 fixed effect.
- Diesel prices in the post-merger periods take the May 2008 value.
- The 2008 income draws are used in 2009, 2010, and 2011.
- The structural error terms are zero in all time periods.

Figure D.5 shows the resulting price plots for Miller Lite and Bud Light. The inter-temporal variability in prices that exists post-merger is nearly eliminated (some remains due to product

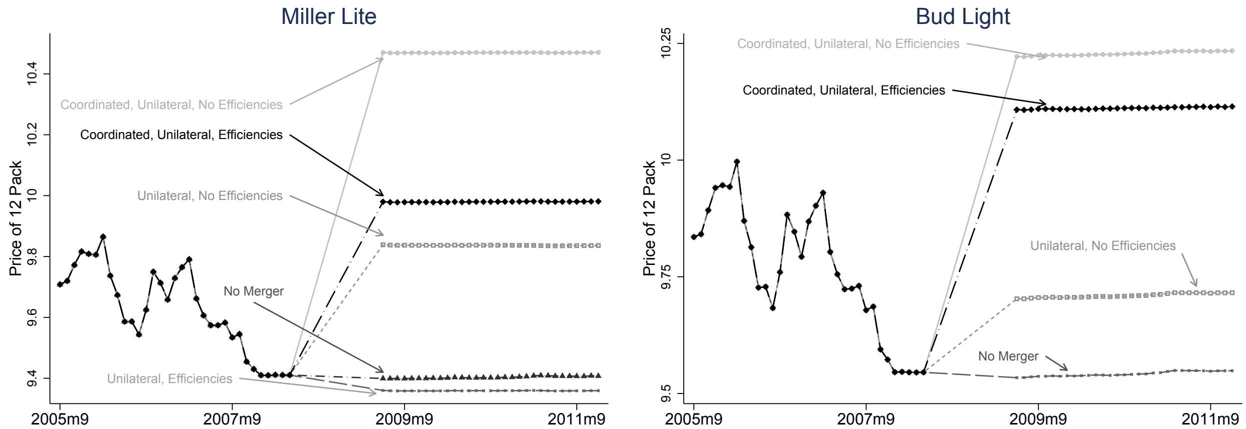

Figure D.5: Alternative Counterfactual Prices for Miller Lite and Bud Light

Notes: The figure plots the average prices of Miller Lite 12-Packs under five different counterfactual scenarios. The post-merger time fixed effects take the value of 2008:Q2 in both the demand and cost functions, demographics and gasoline prices in the post-merger periods take their 2008 values, and unobserved demand and costs are set to zero in all periods. Each dot represents the average prices across the 39 regions.

entry/exit). The price gaps between the different scenarios is similar to what is obtained in the baseline counterfactual simulations. To the extent that the lower Bud Light prices in the baseline simulations reflect prices that are somehow too low, due to estimation error or misspecification, a "conservative" calculation of the merger welfare effects can be obtained by comparing the welfare under the raw data (e.g., as in Figures 4 and 5) to welfare in the revised "no merger" scenario shown in Figure D.5. In that comparison, we find that the merger increases producer surplus by 23%, decreases consumer welfare by 2.0%, and increases total welfare by 2.8%. These numbers are comparable to the welfare statistics shown in Table **??**. Table D.9 shows welfare stats where demand and costs, the income distribution, and diesel prices are all held to May 2008 values. Again, the numbers are very similar to those reported in Table **??**.

E Computational Details

E.1 Demand estimation

Our code is written in Matlab and largely tracks that of Nevo (2000b). The main differences relate to the contraction mapping. Grigolon and Verboven (2014) show that the standard algorithm needs to be adjusted slightly in order to meet the conditions for a contraction mapping if the nesting parameter ρ is sufficiently large. We solve for the mean utility levels,

Table D.9: Welfare Effects with Fixed Demand and Costs

Coordinated Effects:	yes	yes	no	no
Unilateral Effects:	yes	yes	yes	yes
Efficiencies:	yes	no	yes	no
Producer Surplus	19.8%	17.6%	8.8%	7.5%
ABI	9.3%	18.5%	-0.3%	8.9%
Miller	36.5%	16.4%	24.3%	6.8%
Coors	50.6%	13.7%	37.1%	4.2%
Consumer Surplus	-3.1%	-4.7%	0.01%	-1.8%
Total Surplus	1.3%	-0.4%	1.8%	0.0%

Notes: The table provides the percentage changes in producer surplus, consumer surplus, and total surplus, relative to the "No Merger" scenario. All statistics are for the year 2011.

δ_{rt}, in region r and period t by iterating over $i = 1, 2, \ldots$ as follows:

$$\delta_{rt}^{i+1} = \delta_{rt}^{i} + (1-\rho)\ln(s_{rt}) - (1-\rho)\ln(s_{rt}(\delta_{rt}^{i})) \tag{E.1}$$

The presence of $(1-\rho)$ slows the speed of convergence. We compute the contraction mapping in C, separately for each region-period combination, using a tolerance of 1e-14. We minimize the objective function using the Nelder-Mead non-derivative search algorithm with a convergence criterion of 1e-03. The parameter estimates are stable if tighter convergence criteria are used. The Hessian of the objective function at the optimum is positive definite and well-conditioned. We started the RCNL-3 specification using one hundred randomly drawn starting values (constrained within reasonable bounds) to help confirm that the estimation procedure identifies a global minimum of the objective function.

E.2 Standard error adjustment

The supply-side model of price competition is estimated conditional on the demand parameters obtained from the RCNL model. We correct the supply-side standard errors in order to account for the uncertainty present in our demand estimates. The correction is sketched in Wooldridge (2010), although the specific formulation is tailored to our application.

Let $E[g(z_{jm}, \theta_0^S, \theta_0^D)] = 0$ denote a 12 dimensional vector of supply-side moment conditions, where z_{jm} is a vector of instruments for product j in market m, θ_0^S is a 3-dimensional vector of supply-side parameters, and θ_0^D is a 5-dimensional vector of demand-side parame-

ters. The first-order conditions of the supply-side GMM objective function are:

$$0 = \left[J_{\theta^S} g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})\right]^T_{3\times 12} C_{12\times 12}\left[g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})\right]_{12\times 1} \tag{E.2}$$

where C is an estimate of the variance of the supply-side moment conditions, $g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})$ is the sample analogue of the moment orthogonality conditions, and $J_{\theta^S} g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})$ is the 12×3 Jacobian matrix of the sample analog moment conditions with respect to the supply-side parameters.

Taking a mean value expansion of $g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})$ around θ_0^S allows us to rewrite the first-order conditions as:

$$0 = \left[J_{\theta^S} g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})\right]^T_{3\times 12} C_{12\times 12}\left[g(z_{jm}, \theta_0^S, \hat{\theta^D})_{12\times 1} + J_{\theta^S} g(z_{jm}, \bar{\theta^S}, \hat{\theta^D})_{12\times 3}\left(\hat{\theta^S} - \theta_0^S\right)_{3\times 1}\right] \tag{E.3}$$

Solving for $\hat{\theta^S} - \theta_0^S$ and scaling by the square root of the number of markets M gives the following expression for $\sqrt{M}\left(\hat{\theta^S} - \theta_0^S\right)$:

$$-\left[\left[J_{\theta^S} g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})\right]^T C\left[J_{\theta^S} g(z_{jm}, \bar{\theta^S}, \hat{\theta^D})\right]\right]^{-1}\left[\left[J_{\theta^S} g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})\right]^T C\right]\sqrt{M} g(z_{jm}, \theta_0^S, \hat{\theta^D}) \tag{E.4}$$

Now take a mean value expansion of $g(z_{jm}, \theta_0^S, \hat{\theta^D})$ about θ_0^D.

$$g(z_{jm}, \theta_0^S, \hat{\theta^D})_{12\times 1} = g(z_{jm}, \theta_0^S, \theta_0^D)_{12\times 1} + J_{\theta^D} g(z_{jm}, \theta_0^S, \bar{\theta^D})_{12\times 5}(\hat{\theta^D} - \theta_0^D)_{5\times 1} \tag{E.5}$$

where $J_{\theta^D} g(z_{jm}, \hat{\theta^S}, \bar{\theta^D})$ is the 12×5 Jacobian matrix of the sample analog moment conditions with respect to the demand-side parameters. The term $(\hat{\theta_0^D} - \theta_0^D)$ can be rewritten in terms of the sample analog of the demand side moment conditions and the Jacobian of the demand side moment conditions:

$$\left(\hat{\theta_0^D} - \theta_0^D\right) = -\left[\left[J_{\theta^D} h(z_{jm}^D, \hat{\theta^D})\right]^T A\left[J_{\theta^D} h(z_{jm}^D, \bar{\theta^D})\right]\right]^{-1}\left[\left[J_{\theta^D} h(z_{jm}, \hat{\theta_0^D})\right]^T * A\right] h(z_{jm}^D, \hat{\theta_0^D}) \tag{E.6}$$

where $h(z_{jm}^D, \hat{\theta^D})$ is the empirical analog of the vector of demand-side moment conditions and A is an estimate of the variance covariance matrix of the demand-side moment conditions. Plugging this into equation (E.4) gives a first-order representation for $\sqrt{M}(\hat{\theta^S} - \theta_0^S)$:

$$\begin{aligned}\sqrt{M}(\hat{\theta^S} - \theta_0^S) \quad = \quad & \left[\left[J_{\theta^S} g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})\right]^T C^S\left[J_{\theta^S} g(z_{jm}, \bar{\theta^S}, \hat{\theta^D})\right]\right]^{-1}\left[\left[J_{\theta^S} g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})\right]^T C\right] * \\ & \sqrt{M}\left(g(z_{jm}, \theta_0^S, \theta_0^D) + J_{\theta^D} g(z_{jm}, \theta_0^S, \bar{\theta^D}) * (\hat{\theta^D} - \theta_0^D)\right)\end{aligned} \tag{E.7}$$

A consistent estimate of $Var(\hat{\theta}^S)$ is:

$$\Big[G^T C G\Big]^{-1} G^T C \Omega C G \Big[G^T C G\Big]^{-1} \tag{E.8}$$

where

$$\begin{aligned}
G &\equiv \Big[J_{\theta^S} g(z_{jm}, \hat{\theta^S}, \hat{\theta^D})\Big] \\
\Omega &= \sum_{m=1}^{M} (z_m^{S'} \omega_m + F z_m^{D'} \zeta_m)(z_m^{S'} \omega_m + F z_m^{D'} \zeta_m)' \\
F &= J_{\theta^D} g(z_{jm}, \hat{\theta^s}, \hat{\theta^D}) \Bigg[\Big[J_{\theta^D} h(z_{jm}^D, \hat{\theta^D})\Big]^T C^D \Big[J_{\theta^D} h(z_{jm}^D, \hat{\theta^D})\Big]\Bigg]^{-1} \Bigg[\Big[J_{\theta^D} h(z_{jm}, \hat{\theta^D})\Big]^T C^D\Bigg]
\end{aligned}$$

The Jacobians of the supply side moments $J_{\theta^D} g(z_{jm}, \hat{\theta^s}, \hat{\theta^D})$ and $J_{\theta^S} g(z_{jm}, \hat{\theta^s}, \hat{\theta^D})$ were approximated by symmetric two-sided finite differences.

F Retail Sector

F.1 Overview

In this appendix, we extend the supply-side model to incorporate a retail sector. The extension features linear pricing, consistent with industry regulations that prohibit slotting allowances. Brewers set their prices first; the representative retailer observes these prices and sets downstream prices accordingly. Double marginalization arises in equilibrium. The main results regarding brewer competition are largely unaffected. This is because adding retail markups reduces the marginal costs implied by the model by a commensurate amount; the presence of a retail sector is economically similar to a per-unit tax that brewers must pay. One implication is that it is difficult to interpret the magnitude of implied marginal costs unless the model accounts for downstream markups.

The extension features a "retail scaling parameter" that determines both retail pass-through and the magnitude of retail markups. This allows us to scale these objects together between the extremes of retail monopoly power and perfect marginal cost pricing. The model is not a reduced-form in the sense of being linked to fully-specified model of retail competition. Rather, it is a convenient way to think about different degrees of retail competition while avoiding the complications that would arise from modeling the downstream retail pricing game (which is not our focus). The flexibility is desirable in part because individual retailers are unlikely to have monopoly power over entire IRI regions. Previous articles on

the beer industry either assign monopoly power to the retail sector (e.g., Hellerstein (2008); Goldberg and Hellerstein (2013); Asker (2016)) or implicitly assume perfect competition among retailers (e.g., Hausman, Leonard and Zona (1994); Slade (2004); Pinske and Slade (2004); Rojas (2008); Romeo (2016)).

The retail scaling parameter can be normalized or estimated. We provide results over a range of normalizations, but also show that estimation is computationally tractable, leveraging an insight of Jaffe and Weyl (2013). The most likely identification strategy, however, is subject to the Cort's critique and may provide misleading estimates of retail markups if firms do not price according to the model (which is reduced-form). Thus, we are skeptical that retail markups can be recovered from the data in our application.

F.2 Model and identification

Retail prices are set by a representative retailer. Let p_t^R be a vector of retail prices during period t and let p_t^B be a vector of brewer prices (i.e., what retailers pay to the brewer). We suppress region subscripts for brevity. The retail price vector can be decomposed into the brewer prices, a vector of physical/opportunity costs, and a vector of product-specific retail markups:

$$p_t^R = p_t^B + mc_t^R + markup_t^R(\lambda, p_t^B, \theta^D) \tag{F.1}$$

where λ is the retail scaling parameter and θ^D. Brewers set their prices with knowledge of equation (F.4). The resulting first-order conditions are

$$p_t^B = mc_t^B - \left[\Omega_t(\kappa) \circ \left(\frac{\partial p_t^R(p_t^B; mc_t^R, \lambda, \theta^D)}{\partial p_t^B}\right)^T \left(\frac{\partial s_t(p_t^R; \theta^D)}{\partial p_t^R}\right)^T\right]^{-1} s_t(p_t^R; \theta_d) \tag{F.2}$$

where mc_t^B is the vector of brewer marginal costs. Note that brewer markups depend on the retail pass-through matrix $[\partial p_t^R / \partial p_t^B]$ because this determines how brewer prices affect market shares. Plugging back into the retailer pricing equation yields

$$\begin{aligned} p_t^R &= mc_t^R + mc_t^B + markup_t^R(\lambda, p_t^B, \theta^D) \\ &- \left[\Omega_t(\kappa) \circ \left(\frac{\partial p_t^R(p_t^B; mc_t^R, \lambda, \theta^D)}{\partial p_t^B}\right)^T \left(\frac{\partial s_t(p_t^R; \theta^D)}{\partial p_t^R}\right)^T\right]^{-1} s_t(p_t^R; \theta_d) \end{aligned} \tag{F.3}$$

The marginal cost vectors are not separately identifiable, but a composite marginal cost function can be specified along the lines of equation (11). Once the composite cost function

is incorporated, the structural error term can be obtained as in equation (12).

We now examine the retail markup function in greater detail. To build some initial intuition, consider a "fixed markup" system in which retail markups are unaffected by demand conditions and brewer prices (heterogeneity across products or markets may exist). In this specification of the model, retail pass-through is given by the identity matrix and retail markups are not separately identifiable from marginal costs. Thus, the fixed markup system is identical to the baseline supply-side model with no retail sector, with the distinction that implied marginal costs incorporate some unidentifiable retail markup.

An alternative approach is to assume that the representative retailer maximizes profit. The usual derivations show that the vector of product-specific retail markups is given by

$$markup_t^R(\lambda, p_t^B, \theta^D) = \lambda \left[\left(\frac{\partial s_t(p_t^R; \theta^D)}{\partial p_t^R} \right)^T \right]^{-1} s_t(p_t^R; \theta^D) \tag{F.4}$$

This is a standard multi-product monopoly formulation with the simple tweak that $\lambda \in [0, 1]$ scales the retail markups. The retailer sells all products, and internalizes the effects that the retail price of each product has on the sales of other products. For example, the retailer has both Bud Light and Miller Lite on the shelf, and a lower retail price on Bud Light results in some cannibalization of Miller Lite sales. If $\lambda = 1$ then the model corresponds to the representative retailer having monopoly power over each region. If $\lambda = 0$ then the model corresponds to marginal cost pricing; this is observationally equivalent to the constant markup model. If $0 < \lambda < 1$ the model can be interpreted as corresponding to intermediate levels of retail market power, although the mapping to a fully-specified model of retail oligopoly is unclear. We show how λ affects retail pass-through in the next subsection.

If the retail scaling parameter is to be estimated, an additional instrument is required because retail markups are affected by unobserved costs. The obvious candidates are demand-shifters, in which case the logic of identification follows the discussion in Bresnahan (1982). The Corts critique applies: estimates are consistent only if retailers set markups according to equation (F.4). For example, a retailer that does not adjust prices with demand conditions would be inferred to be pricing at marginal cost, even if the retailer has positive markups. Additionally, estimation should be sensitive to demand curvature assumptions, which influence how profit-maximizing prices respond to demand shifters. Because of these concerns, we forgo estimation and instead normalize λ to different levels. We then examine the implications for markups, marginal costs, and brewer competition.[17]

[17] We use a market sizes based on 130% each region's observed sales.

F.3 Retail pass-through

Before proceeding to the results, we develop the connection between λ and retail pass-through, and show how estimation can be made computationally tractable. It is useful to rewrite the retail first order conditions as follows:

$$f(p_t^R) \equiv p_t^R - p_t^B - mc_t^R + \lambda \left[\left(\frac{\partial s_t(p_t^R;\theta^D)}{\partial p_t^R} \right)^T \right]^{-1} s_t(p_t^R;\theta^D) = 0 \tag{F.5}$$

Following Jaffe and Weyl (2013), the implicit function theorem can be applied to obtain the retail pass-through matrix:

$$\frac{\partial p_t^R}{\partial p_t^B} = - \left(\frac{\partial f(p_t^R)}{\partial p_t^R} \right)^{-1} \tag{F.6}$$

The Jacobian matrix on the right-hand-side depends on both the first and second derivatives of demand. For any set of demand parameters, retail pass-through can be calculated by (i) numerically integrating over the consumer draws to obtain the $J \times J$ matrix of first derivatives and the $J \times J \times J$ array of second derivatives; (ii) manipulating these to obtain $\partial f(p_t^R)/\partial p_t^R$; and (iii) obtaining the opposite inverse of the Jacobian. Due to memory constraints, this is most efficiently done separately for each region-period combination in the data.

The computational burden of calculating pass-through for each candidate vector of supply parameters is prohibitive, but this is not necessary in estimation. Instead, one can calculate $\partial f(p_t^R)/\partial p_t^R$ under the assumption $\lambda = 1$, and save this Jacobian. It is then simple to adjust the Jacobian in accordance with any candidate λ. To clarify this procedure, we provide a closed-form expression for column n of the Jabobian. Suppressing period-level subscripts, the column vector is given by

$$\frac{\partial f^R(p^R)}{\partial p_n} = - \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \end{bmatrix} + \lambda \left[\frac{\partial s}{\partial p^R}^T \right]^{-1} \left[\frac{\partial^2 s}{\partial p^R \partial p_n}^T \right] \left[\frac{\partial s}{\partial p^R}^T \right]^{-1} s - \lambda \left[\frac{\partial s}{\partial p^R}^T \right]^{-1} \left[\frac{\partial s}{\partial p_n}^T \right], \tag{F.7}$$

where the 1 in the initial vector is in the n^{th} position. In estimation, start with the Jacobian obtained under $\lambda = 1$ and then, for each vector of candidate supply-side parameters, (i) subtract the identity matrix from the initial Jacobian, (ii) scale the remainder by λ, (iii) add back the identity matrix; and (iv) take the opposite inverse to obtain a retail pass-through

Table F.1: Supply-Side Estimates with Retail Market Power

		(i)	(ii)	(iii)	(iv)	(v)	(vi)
Post-Merger Internalization of Coalition Pricing	κ	0.291 (0.047)	0.292 (0.045)	0.294 (0.045)	0.297 (0.046)	0.300 (0.047)	0.303 (0.048)
Retail Scaling Parameter	λ	0.00	0.025	0.05	0.10	0.15	0.20
Derived Statistics							
Marginal Costs < 0		0.00%	0.25%	0.49%	1.56%	4.90%	10.78%

Notes: The table shows supply-side results based on a model that incorporates retail market power. The demand derivatives from the RCNL-3 specification are used. There are 94,656 observations at the brand-size-region-month-year level. The sample excludes the months between June 2008 and May 2009. The marginal cost functions incorporate the baseline cost-shifters as well as product (brand×size) and period fixed effects. Standard errors are clustered by region and shown in parentheses.

matrix that is fully consistent with the candidate parameter vector.

F.4 Results

Table F.1 provides the results of supply-side estimation for different normalizations of the retail scaling parameters. Two main patterns are relevant. First, estimates of κ are largely unaffected by the magnitude of the retail scaling parameter. Second, the number of products for which implied marginal costs are negative increases as the retail scaling parameter grows larger: under the baseline specification ($\lambda = 0$) there are no negative marginal costs, but at the highest level shown ($\lambda = 0.20$) more than 10% of the marginal costs are negative.

Both of these patterns arise because retail markups are economically similar to a per-unit tax paid by the brewers. (This is exact with fixed retail markups and approximate with profit-maximizing retailers.) In Table F.2 we provide the average pre-merger markups and marginal costs that arise under each normalization of λ. Retail markups increases monotonically with λ. Under the baseline specification ($\lambda = 0$) there are no retail markups, but at the highest level shown ($\lambda = 0.20$) the average retail markup is \$4.04. Brewer markups change very little over different levels of λ, but the implied composite marginal costs decrease nearly one-for-one as retail markups increase.

This is why it is difficult to interpret the magnitude of implied marginal costs without outside information about retail markups. Tremblay and Tremblay (2005, p. 162) indicate that, in the year 1996, retail and distributor markups accounted for about 35% of the retail price, in which case the normalization $\lambda = 0.15$ seems like a decent fit for the industry. An obvious caveat is that the Tremblay and Tremblay information substantially predates the regression sample. For our purposes, the link between marginal costs and retail markups is helpful because inferences about brewers are unaffected by degree of retail market power.

Table F.2: Markups and Marginal Costs with Retail Market Power

	$\lambda = 0.00$ (i)	$\lambda = 0.025$ (ii)	$\lambda = 0.05$ (iii)	$\lambda = 0.10$ (iv)	$\lambda = 0.15$ (v)	$\lambda = 0.20$ (vi)
			Average Markups			
Retail	0.00	0.51	1.01	2.02	3.03	4.04
Bud Light	3.85	3.86	3.87	3.89	3.90	3.91
Budweiser	3.82	3.81	3.79	3.76	3.72	3.69
Coors	2.60	2.56	2.52	2.44	2.36	2.28
Coors Light	2.63	2.61	2.58	2.53	2.48	2.43
Corona Extra	2.67	2.68	2.68	2.70	2.71	2.71
Corona Light	2.62	2.59	2.56	2.50	2.45	2.39
Heineken	2.50	2.49	2.48	2.46	2.45	2.43
Heineken Light	2.46	2.42	2.39	2.32	2.25	2.18
Michelob	3.88	3.84	3.80	3.71	3.64	3.56
Michelob Light	3.87	3.84	3.81	3.75	3.69	3.63
Miller Gen. Draft	3.03	2.99	2.95	2.86	2.78	2.70
Miller High Life	3.00	2.97	2.94	2.87	2.81	2.74
Miller Lite	3.02	3.01	2.98	2.94	2.90	2.85
			Average Composite Marginal Costs			
Bud Light	5.88	5.37	4.86	3.84	2.83	1.82
Budweiser	5.90	5.41	4.92	3.94	2.97	1.99
Coors	7.10	6.63	6.17	5.24	4.31	3.38
Coors Light	7.10	6.62	6.15	5.20	4.25	3.30
Corona Extra	11.62	11.10	10.58	9.54	8.50	7.46
Corona Light	11.69	11.21	10.73	9.76	8.80	7.84
Heineken	11.71	11.20	10.70	9.68	8.67	7.65
Heineken Light	11.74	11.26	10.78	9.83	8.87	7.92
Michelob	6.77	6.31	5.85	4.93	4.00	3.08
Michelob Light	6.94	6.47	6.00	5.07	4.12	3.18
Miller Gen. Draft	6.69	6.23	5.77	4.85	3.93	3.01
Miller High Life	4.40	3.93	3.45	2.50	1.56	0.61
Miller Lite	6.66	6.18	5.70	4.74	3.78	2.82

Notes: The table provides average markups and marginal costs for each 12-pack product over the pre-merger periods, based on the RCNL-3 specification shown in column (iv) of Table 4 and the supply-side results of Table F.1. Composite marginal costs are the sum of brewer and retail marginal costs.