Vertical Relationship and Merger Effects in the U.S. Beer Industry

Bo Bian*
Pennsylvania State University

July 7, 2018

Abstract

In this paper, I study the MillerCoors joint venture of 2008 in the U.S. beer industry. In particular, I focus on impacts of the cost efficiency (in terms of shipping distance and production cost) and increased market power of this merger and importantly how they are affected by the vertical market structure. With vertical relationship, the upstream shock does not fully pass through to downstream retail price because both upstream and downstream will adjust markups to the shock. Thus, merger analysis in the beer industry depends on concentration of both upstream and downstream markets. I use random coefficient model to estimate demand for beer and price elasticities. In the supply side, I model the double marginalization problem of beer retailers and brewers by assuming linear pricing contracts between upstream and downstream firms. Cost saving of the merger is estimated by comparing pre- and post-merger implicit marginal costs. I simulate the two markups in the post-merger period without joint venture to quantify and disentangle the merger effects. I find that average cost saving of producing a 12 oz serving is 2 cents for Coors light and 1.6 cents for Miller lite. Cost saving through shipping distance is at most 7.4 cents for Coors and 2.2 cents for Miller. Market power of MillerCoors increases brewers' markups which dominate the cost saving. However, retailers' markups decreases to mitigate the impact on retail prices especially for more concentrated downstream markets. Social welfare increases after the joint venture.

Key Words: double marginalization, merger, cost saving, market power. (*JEL: L4, L13, L66*)

*Department of Economics, The Pennsylvania State University. Email: **bub186@psu.edu**. The conclusions drawn from the Nielsen data provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business are those of the researchers and do not reflect the views of Nielsen Company (US), LLC. Nielsen Company holds the copyright © 2018, however is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

This paper studies the mega merger between the second and the third largest beer firms, SABMiller and Molson Coors, in the U.S. beer industry. This merger was completed in summer 2008. At that time, the market share of Miller is 18% and of Coors is 11% which makes the joint firm MillerCoors have 29% market share in comparison to 49% market share of the largest firm Anheuser-Busch. To evaluate the merger effect, it is standard to analyze and compare the two opposite effects, cost saving and increased market power. On the one hand, cost saving will decrease the post-merger price and on the other hand consolidation will increase price.

In the literature, there are mainly two types of merger study. First, a merger can be studied in a retrospective way with both pre- and post-merger data. Usually a retrospective study uses a reduced form analysis. Or, a structural model approach is applied to predict a merger effects with only pre-merger data. However, the prediction assumes fixed market environment or fixed unobserved shocks after the merger. My paper is a mixture of both types of studies. I build a structural model to analyze the MillerCoors merger with both pre- and post-merger data. The benefits are twofold. First, with both pre- and post-merger data, I can estimate and capture changes of unobserved demand and supply shocks which can not be obtained with only pre-merger data. Second, with structural model approach, I can quantify and disentangle the welfare changes of this merger in terms of implicit marginal costs, markups and consumer welfare in comparison to a reduced form analysis. Most importantly, I can study the impact of vertical relationship and market structure on the merger effects which can be extended to understand merger in other industries.

Specifically, I build a model with both demand and supply side decisions. In the demand side, I model consumers' discrete choices among differentiated beer brands. In the supply side, I model two stage pricing of upstream brewers and downstream retailers. In the second stage, downstream retailers set optimal retail prices given the whole sale prices from upstream. In the first stage, upstream firms set optimal wholesale prices anticipating the best response of retailer prices. Given the demand estimates, I can calculate double marginalizations and joint implicit marginal costs of retailers and brewers without observing wholesale prices using the approach of Villas-Boas (2007). Then I can estimate the cost saving by comparing pre- and post-merger implicit marginal costs. Cost saving through shipping is estimated via the variation of

distance between breweries and markets. Cost saving through production is estimated by adding interaction of merger dummy and brand dummies in the linear equation of marginal costs. Given the model estimates, I simulate the scenario of no merger for the post- merger period to understand: 1, cost saving effects; 2, market power effects; and how different downstream concentration transfers upstream shocks differently[1].

The contribution of this paper to literature is threefold. First, it contributes to the reduced form analysis of this merger by Ashenfelter, Hosken and Weinberg (2015). In their paper, they also study the MillerCoors merger and they use reduced form analysis to study how increased market power (measured by change of Herfindahl-Hischman index (HHI)) and cost saving (measured by reduced shipping distance) affect final retail prices of beer. However, change of retail prices does not fully reflect the change of marginal costs and change of markups. For example, given the double marginalization model, one dollar decrease of marginal cost will not decrease the final price by one dollar because firms will adjust markups (Hellerstein 2008, Goldberg and Hellerstein (2013)). With structural model, I can quantify the cost saving and markups and their pass-through to final prices, in other words, the underlying mechanism of price change. Moreover, I can disentangle the cost saving and market power effects on price.

Second, I contribute to the structural merger analysis with both pre- and post-merger data. As pointed out by Nevo and Whinston (2010), the limitation of merger prediction with pre-merger data is that it relies on assumption of unchanged market environment after the merger. For instance, if the estimated demand shocks or supply shocks change after the merger, the structural model can not account for them. There are several studies about the accuracy of merger simulation such as Peter (2006), Houde (2012), Weinberg (2011) and Hosken and Weinberg (2013). With a long sample period covering this merger, I can account for the change of unobserved shocks and simulate the no merger scenario for the post merger period such that the merger analysis does not suffer from the limitation above. Moreover, this paper contributes to the merger analysis by considering vertical relationship in the supply side rather than one stage supply decision (Nevo (2000, 2001), Fan (2013)).

Third, the most important contribution of this paper is to study how downstream market concentration affects the transition of upstream shocks, which is merger in this paper, in a vertical relationship. There are many empirical literature on vertical

[1]The reason of simulating no merger scenario for the post-merger period to compare with post-merger data is similar to the difference in differences idea.

relationship of supply side. Chen (2014), Asker (2016) and Lee (2013) study the exclusive contracts between upstream and downstream firms. Villas-Boas (2007) develops a model with double marginalization to study different vertial relationships between manufactureres and retailers. Murry (2017) study the advertising efforts between dealers and manufacturers in automobile industry. Yang (2017) studies product innovation in vertical relationship of smart phone industry.

There are several works close to this paper which also study beer industry including Hellerstein (2008), Goldberg and Hellerstein (2008), Dearing (2016), Miller and Weinberg (2017), Sweeting and Tao (2017). Hellerstein (2008), Goldberg and Hellerstein (2008) study the pass-through of cost shocks to retailer prices in the beer industry. Similar to their idea, my paper studies the pass-through of cost saving and market power due to upstream merger to retail prices. Moreover, my paper emphasizes the heterogeneity of the pass-through due to different downstream concentration. Dearing (2016) studies how upstream affects downstream chain's setting uniform retail price across stores within the chain. In comparison, Villas-Boas (2009) studies welfare effects of uniform wholesale pricing. Miller and Weinberg (2017), Sweeting and Tao (2017) study the merger of MillerCoors in the aspect of collusion or incomplete information on marginal costs. However, all these works model at most one firm in the downstream market such that they do not study the downstream market concentration effects.

One work quite close to this paper is by Manuszak (2010) which studies the impact of upstream mergers on downstream market in gasoline industry. However in his paper, downstream retailers are affiliated with upstream suppliers unlike beer industry in which a retailer chain sells brands from all upstream brewers. Due to the exclusive relationship between downstream stations and upstream suppliers, upstream merger mainly affect participating firms and associated downstream stations in Manuszak (2010). In my paper, upstream merger affects all downstream retailers.

The rest of this paper is structured as follows. Section 2 introduces the background of the U.S. beer industry. Section 3 introduces the data used in this paper and section 4 shows preliminary analysis. The empirical model is discussed in section 5 with estimation results in section 6. In section 7, I simulate counterfactuals and analyze welfare change of the merger. Section 8 concludes the paper.

2 Background of the U.S. Beer Industry

The U.S. beer industry has a long history and is quite mature in modern years. Unlike other industry, the beer industry is highly regulated by government. After the repeal of Prohibition (1919-1933), when individual state was given the right to regulate its beer sales, the policies on beer consumption differ across states. Even so, there are laws in common for almost all states. The most important one is the "three-tier distribution system" feature of beer distribution.

In the three-tier system, beer manufacturers are not allowed to sell beer directly to consumers, retailers, restaurants or bars. Instead, they must sell their beer through state licensed beer wholesalers who thereafter sell beer to retailers, restaurants or bars. Exception adopted by some states allows small craft brewers to sell beer directly to retailers, given that their annual output does not exceed certain limits[2]. Almost all the U.S. states adopt this three-tier system. The main intent of this system is to avoid over-consumption and alcohol abuse, which led to the Prohibition (1919 to 1933). In principle, brewers are free to choose the beer distributor/wholesaler and distributor is free to choose the brands portfolio to carry. Finally, it is up to the retailers who decide which brands to put on the shelves. Within the three-tier system, any kind of vertical integration is discouraged. However, some manufacturers still try to build special relationship with their distributors. For example, Anheuser-Busch has some contracted exclusive wholesalers who can only sell Anheuser-Busch brands. Chen (2014) studies the foreclosure effect of the Anheuser-Busch exclusive wholesalers on entry costs of rival brewers. The three-tier distribution system which is in common across states justifies why I model the vertical relationship in the supply side.

At retail level, given each state has its own regulation on alcohol consumption, they can be categorized into control or non-control states. In control states, wine and especially spirits are not allowed to sell in grocery stores. Instead, they can be sold only in liquor stores (some are state-owned). Though regulation on beer sales is less restricted, some states do not allow grocery stores to sell beer (i.e. Delaware, New Jersey, and North Dakota) or only allow beer with less alcohol (ABW<3.2%) to be sold in groceries (i.e. Colorado, Kansas, Minnesota, Oklahoma and Utah.). Beer sales in gas station, convenience store, or pharmacies also vary across states. At the wholesale

[2]In this paper, local craft beer is not included in the sample. Thus, all brands considered are sold through the vertical framework.

level, there are many wholesalers serving each state. Each wholesaler has exclusive territory to distribute beer. In most states, wholesalers form alliance or association. Uniform wholesale prices to retailers within states are encouraged by state or wholesaler association. In this paper, since wholesalers data is not available, for simplification I integrate the manufacturers and wholesalers as one layer such that manufacturers set wholesale price to retailers.

As for the style of beer sold in the U.S., they can be categorized into larger, light, ales, porter, and stout. The distinction is how each style is brewed. Usually the ale, porter and stout have dark color, bitter taste and high alcohol by volume(abv). Light and lager are quite similar except that light has lower calories. Among these styles, light and lager beer account for most of the beer sales by volume, approximately 92.7%. Most of the national brands brewed by large manufacturers are lager or light beer. Ale, porter and stout are mostly brewed by craft breweries and have quite small market share. During the data period 2007-2011, the U.S. beer industry is highly concentrated. It is dominated by three large domestic breweries before the merger including Anheuser-Busch, Miller and Coors followed by two large imported beer companies, Heineken and Grupo Modelo. In June 2008, the second and third largest brewers Miller and Coors created a joint venture named MillerCoors, in which Miller owns 58% and Coors owns 42% of the joint firm. After this merger, Anheuser-Busch has 49% market share, while MillerCoors has 30%. Since this merger almost turns U.S. beer industry from oligopolies into duopoly, it is very interesting and important to evaluate this merger approved by Department of Justice (DOJ). As the DOJ stated in its closing statement, the Division verified that the joint venture is likely to produce substantial and credible savings that will significantly reduce the companies costs of producing and distributing beer. One goal of this paper is to quantify the cost savings and markup change due to this mega-merger.

3 Data

In this paper, I choose sample period from 2007 to 2011 which covers 6 quarters of pre-merger and 14 quarters of post-merger. The data come from several sources. The beer prices and sales data come from Nielsen retail scanner data, which records the weekly sales of all beer in more than half U.S. retail stores (in sales volume) across the country. The demographics data is from American Community Survey. These two datasets

are used for demand estimation. The third dataset comes from Quarterly Census of Employment and Wages (QCEW) of Bureau of Labor Statistics. The QCEW dataset includes average weekly wages in each geographic area. I also add the median gross rent (from ACS) of each geographic area. I use local wages and gross rent to control for retailer costs of each geographic market. In addition, I collect beer characteristics from brewers' websites for demand estimation and collect hop, malt prices to control for brewer's cost. Shipping distance between breweries to markets are calculated using arcGIS. The details of each dataset and how I construct the sample is described below.

3.1 Nielsen Retail Data

The Nielsen retail scanner dataset is at weekly level which records beer sales of participating stores. In other words, one observation in the data is for weekly sales of a Universal Product Code (UPC) in a store. The product comes at the UPC level, which differs in pack size, container and volume per container for the same brand. A store is uniquely identified by store id and 3-digit zip code, county, state. Each store also has a parent code to identify the ownership or chain it belongs to such that I can tell the number stores of a chain in a geographic area. The chains are categorized as different channels including food store, drug store, mass merchandiser, liquor store and convenience store. The coverage of Nielsen across channel is different. Since the coverage rate for liquor and convenience store is quite low, I only consider food chains in the paper.

For each observation, I know the price, quantity sold in a week, UPC information and brand information. Brand information includes the type of the beer (i.e. lager, light, ale, stout or porter) and the brand name. I supplement the product characteristics by collecting abv, carb, calorie, whether domestic or imported and firm it belongs to. I aggregate the UPC into brand-package size level regardless of the container or volume per serving. The reason to distinguish package size is that price per 12 oz serving is quite different between large pack and small pack[3]. Given the variety of pack size, I only distinguish large size (> 12 packs) and small size (≤ 12). Furthermore, I aggregate sales of a given brand-size across stores into chain level. It implicitly assumes that chain sets uniform retail prices across stores which is studied by Dearing (2016). If I allow store to set individual price, the number of "products" will be quite

[3]In general, large pack has smaller price per serving than small pack. By this definition, I have two products for a brand.

large and the market becomes more competitive than price setting by chains. Thus, a well defined product in my paper is a brand-size-chain combination in a geographic markets. I divide the area of the market by number of stores of a chain as proxy for travelling cost to buy a product or how easily to access to a chain.

Finally, I aggregate weekly level into quarterly level to avoid effects of temporary store discounts or household storage behavior (Hendel and Nevo (2016)) on demand estimation. In sum, one observation in the sample is quarterly sale of a brand-size-chain during 2007-2011.

3.2 ACS&QCEW

The American Community Survey data from U.S. census are used to simulate households' demographics in each geographic market. For every quarter of a year, I randomly draw demographics of residents in a geographic market including ratio of income to poverty level, age, education and race based on the distribution of demographics provided in ACS. The benefit of using income to poverty level is that it measures "richness" per capita rather than household income which is not discounted by household size in ACS. Usually, researchers use Current Population Survey(CPS) to generate demographics. However, CPS is not appropriate to analyze geography smaller than a state[4].

I use ACS and QCEW to collect data on local retailer costs. From ACS, I collect gross rent in the market as proxy for commercial rent of retail chains. I also collect average wage in supermarket industry from QCEW to control for retailer costs. I also collect malt and hop prices to control for production costs of beer.

3.3 Shipping distance

One important cost factor is the shipping distance from breweries to the geographic markets which is an argument of cost saving for MillerCoors merger. Ashenfelter, Hosken and Weinberg (2015) finds the reduced shipping distance significantly accounts for post-merger price change. I use the same method of them to calculate the shipping distance. First, I locate all plants of Anheuser-Busch, Miller, Coors and other domestic

[4]The comparison between ACS and CPS is listed in https://www.census.gov/topics/income-poverty/poverty/guidance/data-sources/acs-vs-cps.html

brewers[5]. Then, I use arcGIS to compute the distance between geographic market and closest plant of a brewer as shipping distance. It is implicitly assumed that a plant of a brewer produces the whole product line of brands. Given the fact that I include selected nation wide brands in my sample, this assumption is plausible. For small brewers in the sample, such as Yuengling, Sierra Nevada and New Belgium brewery, though they are not sold nation wide, they only have few plants and the assumption still holds. As for imported brands, I follow Miller and Weinberg (2017) to calculate distance of markets to the ports. But in the estimation, I simply use fixed effect for imported brands to control for shipping costs. In the post-merger periods, I combine Coors and Miller's plants as MillerCoors plants to calculate the post-merger shipping distance and therefore the reduced distance.

3.4 Market Definition

Information about store location in Nielsen includes 3-digit zipcode, county and state. I define a market as a metropolitan statistics area(MSA) which comprises several central/outlying counties. The reason of using MSA as a market is that residents in each MSA rarely travel outside to purchase beer and MSA seems to have the proper area size for retailers and brewers to make strategic pricing decision. Ashenfelter, Hosken and Weinberg (2015) also uses MSAs as separate geographic markets. I aggregate quarterly beer sales of Nielsen stores in a MSA as market size of the MSA in the quarter. This definition of market size includes all the observed beer sales in Nielsen dataset including all channels such as food store and drug store. The big concern of this market size definition is the coverage rate of Nielsen. Though Nielsen covers, on average, 50% food stores in the U.S., the coverage rate varies a lot across locations[6]. If the missing data comes from stores of the same chains in Nielsen, I could probably use in-sample share as mirror projection to the market share. However, if the missing data comes from chains that do not cooperate with Nielsen and if those chains are big players in the market, my definition of the market size could overestimate the market power of chains in the sample. This issue also occurs to Miller and Weinberg (2017) when they use IRI data. They scale the observed beer sales in data by 1.5 as the market size which is equivalent to normalizing outside market share. An alternative way of defining market size is Hellerstein (2008) which scales the population by beer consumption per capita.

[5]Anheuser-Busch has 12 plants over the country. Miller has 6 plants and Coors has 2 plants.

[6]Nielsen provides the coverage rate by channel at Scantrack markets level.

However, the problem of applying this method is that I do not know the beer consumption per capita in food store channel which varies across MSAs. Second, it would generate a quite large outside market share which could underestimate concentration of retailers.

Instead, I borrow this idea of Hellerstein (2008) to select MSAs included in my final sample. The selection criteria is that the per capital consumption of beer (calculated by dividing Nielsen beer sales by MSA population) is greater than 2 servings per month. Furthermore, I select the MSAs with market share of beer sold through food channel greater than 70% and population larger than 0.2 million. According to these criteria, only MSAs with high food store coverage and large market share of food channel are left. These MSAs are idea for the analysis of this paper because firstly I need high coverage rate of my sample in order to measure the downstream concentration and secondly I do not model competition of food chains with drug stores or convenience stores. In the end, I have 50 MSAs over 20 quarters from 2007-2011 in my sample. Table 1 shows the 50 MSAs and corresponding market information including the number of chains, the number of products (brand-size-chain), the total market share of inside products, market size (sales observed in Nielsen) and food channel coverage of the DMA[7].

4 Preliminary analysis

In this section, I start the analysis by showing some key variables in the data and preliminary regression results on retail prices. First of all, figure 1 shows the dynamics of retail price per 12 oz serving for selected brands by package size(large or small). For a given time period, the price is averaged over MSAs and package sizes and therefore the change is not quite obvious. However, it still illustrates that on average the price of Bud light, Budweiser, Coors light and Miller Lite increases by 10 cents (12.5%) per serving for small packs. The increase is smaller for large packs. After the merger point, both MillerCoors and Anheuser-Busch brand increase prices in comparison to imported beer such as Heineken and Corana. The increase of prices could result from increased market power after merger or increased production cost of domestic brewing. The two key factors that change after the merger are upstream concentration and shipping

[7]DMA is Nielsen Designated Marketing Areas. The document provided by Nielsen does not show the coverage for all DMAs. Usually DMA is larger than MSA which means that coverage rate of a MSA could be larger or smaller than the coverage of a DMA it belongs to.

distance. Figure 2 and 3 demonstrate these two variables for 50 markets.

To obtain 2, I calculate the proxy for HHI change such that $\Delta HHI_{brewer} \approx (s_{miller} + s_{coors})^2 - s_{miller}^2 - s_{coors}^2 = 2 * s_{miller} * s_{coors}$ which is used in Ashenfelter, Hosken and Weinberg (2015). For each MSA-quarter, I calculate the HHI change and the histogram of all MSA-quarter in post-merger periods are shown in figure 2. The variation of ΔHHI_{brewer} is large across markets ranging from 0.01 to 0.07. Given the HHI of national market share $0.49^2 + 0.18^2 + 0.11^2 = 0.28$, after merger brewers' HHI increases more than 10% for most markets. Figure 3 illustrates the reduction of shipping distance between 50 MSAs to MillerCoors breweries. The horizontal axis is shipping distance before merger and vertical axis is shipping distance after the merger. A 45 degree line is used as reference such that the vertical distance from spot to 45 degree line is the reduced shipping distance. As it shows, the merger primarily reduce the shipping distance for Coors than Miller. The reason is that Coors only have 2 plants before the merger and Miller has 6 over the country. Out of the 50 MSAs, only 5 markets have slightly shorter distance after the merger.

To understand the merger effects on retailer prices and vertical relationship in the supply side, I run two specifications on logarithm of price per serving.

$$\begin{aligned}\log(p_{jcmt}) = {} & \alpha_1 HHI_{mt}^{brewer} + \alpha_2 HHI_{mt}^{retailer} + \alpha_3 HHI_{mt}^{brewer} \times HHI_{mt}^{retailer} \\ & + postmerger \times (\beta_1 HHI_{mt}^{brewer} + \beta_2 HHI_{mt}^{retailer} + \beta_3 HHI_{mt}^{brewer} \times HHI_{mt}^{retailer}) \\ & + d_{large} + \gamma \log(distance) + d_{jmt} + \varepsilon_{jcmt} \end{aligned} \tag{4.1}$$

and the second specification

$$\begin{aligned}\log(p_{jcmt}) = {} & \alpha \Delta HHI_{mt}^{brewer} + \beta HHI_{mt}^{retailer} \times \Delta HHI_{mt}^{brewer} \\ & + d_{large} + \gamma \log(distance) + d_{jmt} + \varepsilon_{jcmt} \end{aligned} \tag{4.2}$$

where the subscript j stands for brand j; c stands for chain c; m is geographic market; t is time period. On the left hand side of both equations 4.1 and 4.2 is retailer price of brand j sold in chain c in market m at time t. On the right hand side of 4.1, I use HHI_{mt}^{brewer} to control for brewer markup and $HHI_{mt}^{retailer}$ to control for retailer markup.

Importantly, I add interaction term of two HHI to estimate the vertical relationship between upstream and downstream markets. In other words, the interaction term measures how downstream market concentration affects the transition(slope) of upstream HHI to retail price. For flexibility, I also interact the HHIs with post-merger dummy. Furthermore, dummy for large pack, logarithm of shipping distance, local wage, rent and dummies for market, time and brand are included. The second specification uses change of HHI instead of HHI level similar to Ashenfelter, Hosken and Weinberg (2015). However, I add interaction of $HHI^{retailer}$ with ΔHHI^{brewer} in difference to Ashenfelter, Hosken and Weinberg (2015) in order to study how downstream HHI affects the transition of market power to price.

The regression results are given in table 2. The first column is regression result of equation 4.1 without interaction terms of post-merger dummy. The second column is full specification of 4.1. First, the result shows that both HHI^{brewer} and $HHI^{retailer}$ positively affect retail prices. Coefficient on brewer HHI means increase of HHI by 0.01 points will increase retail price per 12 oz by 0.43%. According to the histogram of figure 2, increasing HHI by 0.03 after the merger will increase retail price by 1.23%. The estimate of HHI interaction term is negative for both column 1 and 2. To interpret, increase of upstream concentration will raise the final retail price less for market with more concentrated downstream market. In other words, market with a severe competition in the downstream market has less power to dampen the upstream shocks. Estimates of other coefficient are negative for large packs and positive for shipping distance. As distance is a key factor in merger, the coefficient means that reduction of shipping distance by 1% will decrease the price by 0.031%.

In column 3 of table 2, it shows regression result of 4.2. Similar to the first specification, increase of upstream concentration will positively affect retail price. However, the effect is mitigated by the downstream concentration according to the negative coefficient -1.232 of interaction term. The interpret is the same to the first specification. In order to quantify changes of marginal cost, markups and vertical relationship in the merger rather than their reduced reflection in price, I build a demand and supply side model in the next sections.

5 Structural Model

5.1 Beer Demand

I model a consumer's decision for purchasing one 12 oz standard serving of beer using a discrete choice model following Berry (1994), Berry et al. (1995), and Nevo (2001). In each MSA market of year-quarter, a consumer's choice set includes all brands sold in all in-sample chains of the market and outside option. A product is defined as a brand-size-chain combination. Stores within the same chain set uniform price for the same brand-size. I do not distinguish stores and products sold by stores within the same chain[8]. The utility function of a consumer i, in market and period mt, of choosing brand-size j in chain c is:

$$u_{ijcmt} = \delta_{jcmt} + \varepsilon_{ijcmt} \tag{5.1}$$

with:

$$\delta_{jcmt} = \alpha p_{jcmt} + \beta x_{jcmt} + \lambda_{jcmt} + \xi_{jcmt} \tag{5.2}$$

where the first term δ_{jcmt} is mean utility of product which comprises the following variables: p_{jcmt} is price per 12 oz of brand j sold in chain c; x_{jcmt} includes product characteristics such as logarithm of the radius ($mile^2$) per store of the chain as measure of travelling distance and dummy for package size; λ_{jcmt} is full set of fixed effects including brand dummies λ_j, market-chain dummies λ_{cm}, year and season dummies λ_t. Brand characteristics such as abv, calorie, carb, dummy for light beer and dummy for domestic beer are not shown up in δ_{jct}, because they are fixed for any given brand and therefore are fully accounted in the brand dummies. ξ_{jcmt} is unobserved demand shock. In Nevo(2001), parameters in δ_{jct} are referred as *linear* parameters $\Theta_1 \equiv \{\alpha, \beta, \lambda\}$. The last term ε_{ijcmt} of utility captures the idiosyncratic preference shock, which is assumed to follow Type 1 extreme-value distribution. This is the standard simple logit demand model.

The random coefficient discrete choice model adds another term to equation 5.1 such that:

[8] As mentioned before, the uniform retail price within chain is studied by Dearing (2016). He finds that upstream adjustment will significantly dampen the profit gains of retailer by deviating from uniform pricing to store-level pricing.

$$u_{ijcmt} = \delta_{jcmt} + \mu_{ijcmt} + \varepsilon_{ijcmt} \tag{5.3}$$

with:

$$\mu_{ijct} = [p_{jcmt}, \tilde{x}_{jcmt}](\Pi D_i + \Sigma v_i) \tag{5.4}$$

This additional term μ_{ijct} includes consumer demographics to make the substitution among products more flexible than simple logit model to solve I.I.A. problem. In μ_{ijct}, together with price $\tilde{x}_{jcmt}$ are product characteristics that consumers with different demographics have heterogeneous tastes to. In the estimation, I include ABV, dummy for light beer, and dummy for domestic beer in $\tilde{x}_{jmct}$; D_i are consumer demographics which captures consumers' heterogeneous preference over product characteristics. I use income and age as demographics variables in the estimation; v_i are consumer i's idiosyncratic preference, which is assumed to follow standard normal distribution in the estimation. The matrix $\Theta_2 \equiv \{\Pi, \Sigma\}$ are *nonlinear* parameters, which measures the different preferences of consumers.

The mean utility of choosing outside option is:

$$u_{i0mt} = \delta_{0mt} + \mu_{i0mt} + \varepsilon_{i0mt} \tag{5.5}$$

the utility of choosing outside good is normalized to be $u_{i0t} = \varepsilon_{i0t}$ for both simple and mixed logit model.

Consumer with demographics $\{D_i, v_i, \boldsymbol{\varepsilon}_i\}$ chooses the option which gives her the highest utility such that $jc^* = \text{argmax}_{jc}\, u_{ijcmt}$. Denote the set of consumers choosing product (j, c) as $A_{jct} = \{D_i, v_i, \boldsymbol{\varepsilon_i} | u_{ijct} > u_{ij'c't}, \forall j', c'\}$. Then, the market share for product (j, c) is:

$$s_{jcmt} = \int_{A_{jcmt}} dP_m^*(D_i, v_i, \boldsymbol{\varepsilon}_i) \tag{5.6}$$

where P_m^* is the joint probability distribution function of $\{D_i, v_i, \boldsymbol{\varepsilon}_i\}$. Given the T1EV distribution assumption on $\boldsymbol{\varepsilon}_i$, this formula can be rewritten as:

$$s_{jcmt} = \int_{D_i, v_i} \frac{\exp(\delta_{jcmt}(\Theta_1) + \mu_{ijcmt}(D_i, v_i; \Theta_2))}{1 + \sum_{j'} \exp(\delta_{j'cmt}(\Theta_1) + \mu_{ij'cmt}(D_i, v_i; \Theta_2))} dP_m(D_i, v_i) \tag{5.7}$$

In the special case of $\Theta_2 = 0$, it becomes simple logit model and the market share is:

$$s_{jcmt} = \frac{\exp(\delta_{jcmt}(\Theta_1))}{1 + \sum_{j'} \exp(\delta_{j'cmt}(\Theta_1))} \tag{5.8}$$

to back out δ_{jcmt} of simple logit model easy such that:

$$\delta_{jcmt} = \log(s_{jcmt}) - \log(s_{0mt}) = \alpha p_{jcmt} + \beta x_{jcmt} + \lambda_{jcmt} + \xi_{jcmt} \tag{5.9}$$

Since market share and outside share are observed in data, I can simply construct the left hand side of equation 5.9 and regress on the variables on the right hand side.

Back to the full random coefficient model, BLP proves that the contraction mapping can help recover mean utility δ from equation 5.7). To be more specific, given each pair of parameter values, I randomly draw ns persons' (D_i, v_i) from "known" distribution of P_m, such that I can calculate the monte carlo simulated share as:

$$s_{jcmt} = \frac{1}{ns} \sum_i \frac{\exp(\delta_{jcmt} + \mu_{ijcmt})}{1 + \sum_{j'} \exp(\delta_{j'cmt} + \mu_{ij'ct})} \tag{5.10}$$

BLP proves that there is a unique vector of δ that can match the simulated market share and observed market share in data. The iteration process for δ is:

$$\boldsymbol{\delta}_{mt}^{h+1} = \boldsymbol{\delta}_{mt}^{h} + \log \boldsymbol{s}_{mt} - \log \boldsymbol{s}(p_{mt}, x_{mt}, \boldsymbol{\delta}_{mt}^{h}, P_m; \Theta_2) \tag{5.11}$$

where the first vector $\boldsymbol{s}_{mt}$ is data and the second vector $\boldsymbol{s}$ is calculated based on simulation. After backing out δ, I can simple run linear regression on mean utility to estimate Θ_1. One problem occurs to both simple and mixed logit model that p_{jcmt} is endogenously correlated with unobserved demand shock ξ_{jcmt}. It is because that when manufacturers and retailers set optimal strategic price, they observe unobserved product characteristics and account for it in price setting. In the estimation, I instrument prices with cost side variables such as local wage, rent, hop and malt prices.

5.2 Supply

The three-tier system of beer distribution includes beer manufacturers, distributors and retailers. To decompose the price setting, it is ideal if I could model all three tiers in the supply side. However, there are more than 3300 licensed beer wholesalers through-

out the U.S. with each of them carrying multiple brands, sharing different territories, having different storage and shipping capacities. To collect detailed information on those wholesalers is hard and the data is not available. Since I don't have data on beer distributors or even the number of distributors in a given MSA, I simply follow the literature (Hellerstein (2008), Goldberg and Hellerstein (2013)) by integrating manufacturers and distributors into one price setting stage which sets whole sale price, and retailer is the second stage setting retail price. In the following discussion, I use brewer or manufacturer to represent brewer-distribution integration. Brewers sets wholesale prices, while retailers set retail prices. As shown in the preliminary regression, the retail price consists of brewers' markup, retailers' markup and joint marginal cost of firms. The supply model solves double markups given estimates of price elasticities in the demand side. Then total marginal cost of brewer and retailer is recovered by subtracting markups from retail price.

In order to calculate the double markups, the price setting is assumed to follow Bertrand-Nash linear pricing game between upstream and downstream firms. Linear pricing contract between upstream and downstream is assumed in Hellerstein (2008), Goldberg and Hellerstein (2013) and Dearing (2016). Most recently, there is a paper by Faheem and Gayle (2017) studying the nonlinear pricing contract between brewers and retailers and its fitness to the data. Sine my paper focuses on the downstream market concentration on upstream merger shocks, I simply assume linear pricing contract.

At first stage, beer brewer sets wholesale price for its differentiated brands accounting for the downstream retailers' best response retail prices. At the second stage, after observing the wholesale prices of all brands by all brewers, the retailers set optimal retail prices for the brands they carry. I do not model retailers' choosing product portfolio in this paper which could be an interesting extension. Instead, I take the brands sold by each chain retailer as exogenously given and retailers only set endogenous retail prices. This assumption is reasonable since I only include "popular" brands in the sample. I use backward induction to solve this two stage game.

5.2.1 Retail Price

The retailer c which is a food chain in a market m at period t chooses retail price p_{jcmt} to maximize its profit[9]:

$$\Pi_t^c = \Sigma_{jc \in J_{ct}}[p_{jct} - p_{jt}^{\omega} - mc_{jct}^r]s_{jct}(\boldsymbol{p}_t) \tag{5.12}$$

where J_{ct} is the set of products sold by chain c; p_{jt}^{ω} is wholesale price of brand j; mc_{jct}^r is marginal cost of chain c for selling product (j, c). To avoid the duplication of script, I use r instead of c and ω to distinguish retailer and manufacturer; s_{jct} is market share of product (j, c). The wholesale price doest not have c in the subscript, because wholesalers are assumed to charge uniform wholesale price to different retailers due to state law or distributor association. Retailer c chooses optimal price vector $\boldsymbol{p}_{ct}$ to satisfy first order conditions:

$$s_{jct} + \Sigma_{j',c \in J_{ct}}[p_{j'ct} - p_{j't}^{\omega} - mc_{j'ct}^r]\frac{\partial s_{j'ct}}{\partial p_{jct}} = 0 \tag{5.13}$$

for all $(j, c) \in J_{ct}$. If I write all products (j, c) of market t into vector, the first order conditions can be rewritten as:

$$\boldsymbol{p}_t - \boldsymbol{p}_t^{\omega} - \boldsymbol{mc}_t^r = -(T_r * \Delta_{rt})^{-1}s_t(\boldsymbol{p}_t) \equiv mkup_t^r \tag{5.14}$$

with T_r as the retailer's ownership matrix. The dimension of matrix T_r equals the number of products in the market, say N_t. The element $T_r(i, j)$ equal to 1 when both product i and j are sold by the same chain and 0 otherwise. $\Delta_{rt}(i, j)$ is N_t by N_t matrix containing the first derivatives of all the shares with respect to all retail prices, with element $\Delta_{rt}(i, j) = \partial s_{jt}/\partial p_{it}$. The values in 5.14 are the retailer markups.

5.2.2 Manufacturers

The manufacturer sets optimal wholesale price p_{jt}^{ω} for brand j taking retailers' optimal pricing strategy of 5.14 into account. manufacturer ω maximizes profit function:

$$\Pi_t^{\omega} = \Sigma_{j \in J_{\omega t}}[p_{jt}^{\omega} - mc_{jt}^{\omega}]s_{jt}(\boldsymbol{p}^*(\boldsymbol{p}^{\omega})) \tag{5.15}$$

[9]To simplify the notation, I drop m and use t to represent market.

where $J_{\omega t}$ is the set of brands owned by manufacturer ω; $\boldsymbol{p}^*(\boldsymbol{p}^\omega)$ is best response function of retailers on wholesale prices $\boldsymbol{p}^\omega$; s_{jt} is the total market share of brand j sold by all chains, which equals $\Sigma_c s_{jct}$[10]. The first order condition of manufacturer's profit w.r.t. p^ω_{jt} become:

$$s_{jt} + \Sigma_{j' \in J_{\omega t}} [p^\omega_{j't} - mc^\omega_{j't}] \frac{\partial s_{j't}}{\partial p^\omega_{jt}} = 0 \tag{5.16}$$

Similarly, let T_ω be a matrix of ownership for the manufacturers. The dimension of matrix T_ω is different to the dimension of T_r because number of brands is less than number of products. Manufacturers do not distinguish brand sold in different chains. Thus, the dimension of matrix T_ω equals the number of brands available in the market, say N^U_t with element $T_\omega(i,j) = 1$ if brand i and j belongs to the same manufacturer.

Let $\Delta_{\omega t}$ be the manufacturer's response matrix with element $\Delta_{\omega t}(i,j) = \partial s_{jt}/\partial p^\omega_{it}$ which has the dimension of N^U_t by N^U_t. To obtain the matrix $\Delta_{\omega t}$, I need to calculate the derivatives of optimal retail prices with respect to wholesale prices, because

$$\frac{\partial s_{jt}}{\partial p^\omega_{it}} = \Sigma_c \Sigma_k \frac{\partial s_{jct}}{\partial p_{kt}} \frac{\partial p_{kt}}{\partial p^\omega_{it}} \tag{5.17}$$

Define matrix Δ_{pt} with element $(i,j) = \partial p_{jt}/\partial p^\omega_{it}$ such that dimension of Δ_{pt} is N^U_t by N_t matrix. Once I get Δ_{pt}, I can construct matrix with derivatives of all product market shares with respect to all wholesaler prices $\Delta_{pt}\Delta_{rt}$, which is a N^U_t by N_t matrix. To simplify notation, I drop the subscript t and c in the following derivation. To get the expression for Δ_p, I totally differentiate first order condition of optimal retail price 5.13 for a given product j with respect to all retail prices $(\mathrm{d}p_k, \forall k = 1, \ldots, N)$ and wholesale price p^ω_F such that:

$$\Sigma^N_{k=1} \underbrace{[\frac{\partial s_j}{\partial p_k} + \Sigma^N_{i=1}(T_r(i,j)\frac{\partial^2 s_i}{\partial p_j \partial p_k}(p_i - p^\omega_i - mc^r_i)) + T_r(k,j)\frac{\partial s_k}{\partial p_j}]}_{\text{g(j,k)}} dp_k - \underbrace{\Sigma_{f \in F} T_r(f,j)\frac{\partial s_f}{\partial p_j}}_{\text{h(j,F)}} dp^\omega_F = 0 \tag{5.18}$$

putting all products j together, let G be the matrix with element $g(j,k)$ and let H_F be the N_t-dimensional vector with element $h(j,F)$. I can rewrite 5.18 in matrix form

[10]The market share of brands and market share of products(brand-chain) use the same notation s which is confusing. But I distinguish them by adding c to the subscript or not.

$G\mathrm{d}\boldsymbol{p} - H_F\ \mathrm{d}p_F^\omega = 0$. Then I get $\Delta_p' = G^{-1} * [H_1 H_2 \dots H_{N^U}]$ with dimension N by N^U. Finally, according to 5.17 $\Delta_{\omega t} = \Delta_{pt}\Delta_{rt}U$, where matrix U is a N_t by N_t^U matrix with element $U(i,j) = 1$ if product i is brand j. Matrix U is used to aggregate product market share s_{jcmt} into brand market share s_{jmt}. Similar to retailer markup, wholesaler markup is :

$$\boldsymbol{p}_t^\omega - \boldsymbol{mc}_t^\omega = -(T_\omega * \Delta_{\omega t})^{-1}\boldsymbol{s}_t(\boldsymbol{p}) \equiv mkup_t^\omega \tag{5.19}$$

Note that the vector of brand market share in 5.19 has N_t^U element in difference to the product market share vector in 5.14.

Calculating the two markups in 5.14 and 5.19 only requires simulated market share and first/second order derivatives of market share with respect to retail prices. The derivatives of market share w.r.t retail prices can be calculated once price elasticities are estimated in the demand side. Since I do not observe wholesale prices $\boldsymbol{p}_t^\omega$, it is impossible to calculate marginal costs respectively using 5.14 and 5.19. However, by combing these two equations, I can recover the joint costs of retailer and manufacturer such that:

$$\boldsymbol{p}_t - \boldsymbol{mkup}_t^r - \boldsymbol{mkup}_t^\omega = \boldsymbol{mc}_t^r + \boldsymbol{mc}_t^\omega \tag{5.20}$$

By moving markups to the right hand side, this equation is analogous to specification 4.1 in the preliminary analysis. The difference between 5.20 and 4.1 is that HHI is a measure for markups but fixing HHI in 4.1 is not equivalent to fixing markups of 5.20. Therefore, the coefficients of cost variables in 4.1 are not the same to those in regressing $\boldsymbol{mc}_t^r + \boldsymbol{mc}_t^\omega$ on cost variables. Once I back out marginal costs for both pre- and post-merger periods, I can estimate cost saving through reduced shipping distance and production rather than their effects on prices.

6 Estimation

6.1 Demand estimation

Estimation of the model has two steps. First, I estimate the random coefficient demand model in the way of past literature. For any given values of Θ_2, I use contraction mapping to solve for fixed point of mean utility $\boldsymbol{\delta}(\Theta_2)$ such that model predicted

market share equals observed share in data. Then, I use the mean utility equation 5.2 to back out unobserved demand shock $\boldsymbol{\xi}(\Theta_1, \Theta_2)$. Due to price endogeneity, I use instruments $Z = [z_1, \ldots, z_L]$ and GMM estimation to estimate parameters $\{\Theta_1, \Theta_2\}$. The moment conditions are:

$$E[z_l \xi] = E[z_l(\delta(\Theta_2) - \alpha p - \beta x - \lambda)] = 0, \quad l = 1, \ldots, L \tag{6.1}$$

with the GMM estimator being:

$$\hat{\Theta} = \text{argmin}\, \boldsymbol{\xi}(\Theta)' Z W Z' \boldsymbol{\xi}(\Theta) \tag{6.2}$$

where W is weight matrix. Following Nevo (2001), the estimation can be simplified by substituting the estimator $\hat{\Theta}_1$ given guessed parameter values Θ_2 :

$$\hat{\Theta}_1 = (X' Z W Z' X)^{-1} (X' Z W X' \boldsymbol{\delta}(\Theta_2)) \tag{6.3}$$

into the GMM estimation such that the estimation algorithm only search over Θ_2 rather than $\{\Theta_1, \Theta_2\}$ to minimize the objective function. For simple logit model, $\boldsymbol{\delta}$ can be calculated using market shares and the estimator for Θ_1 above is equivalent to 2sls iv estimator. Matrix X is product characteristics including retail price, logarithm of the radius per store, dummy for package size, and full set of dummies for brand, market-chain, year and season.

As for the instruments, I primarily use cost shifters and market demographics. To be specific, I use local retailer's costs such as local average wage in supermarket industry and local gross rent. I also use manufacturers' costs including shipping distance between brewery and market, malt and hop prices. I interact malt and hop prices with firm dummies to allow heterogeneous production costs across manufacturers. Hellerstein (2008) and Goldberg and Hellerstein (2008) also use input prices as instrument when they study the pass through rate of cost in the U.S. beer industry. Following Miller and Weinberg (2017), I also use mean demographics interacting with exogenous product characteristics in X as instrument for estimating parameters Π in the random coefficients.

6.2 Supply estimation

Once the demand side is estimated, I can calculate the partial derivative of market shares to retail prices. Based on 5.14 and 5.19, I can calculate markups for retailers and manufacturers. Then the joint marginal cost can be recovered by subtracting markups from retail price based on 5.20. Then I use OLS regression to estimate cost function:

$$mc^r + mc^w = \alpha_1 \log(distance) + \alpha_2 \log(rent) + \alpha_3 \log(wage) + \lambda_{brand} + \lambda_{merge} \times \lambda_{brand} \times \lambda_{millercoors} + \lambda_{mt} + \nu \quad (6.4)$$

where the coefficient α_1 measures shipping cost and interaction terms of dummies $\lambda_{merge} \times \lambda_{brand} \times \lambda_{millercoors}$ (MillerCoors brand dummies after merger) measure the average cost saving after the merger other than shipping cost. For example, if the production cost of MillerCoors brand decreases after the merger, I will have negative coefficients on this set of dummies. One thing to note is that I do not add brand dummies interacting with merger dummy for other brewers. As discussed earlier in the paper, estimation of the cost function and cost saving of the merger can only be achieved when I back out marginal costs in the left hand side of 6.4 and the sample period covers both pre- and post-merger periods. In the counterfactual of this paper, I simulate the scenario of no merger for post-merger period to disentangle cost saving and increased market power of the merger. The cost without merger is calculated using the estimates of cost function 6.4 above.

6.3 Results

6.3.1 Simple logit demand

I start with the simple logit demand model and estimate the mean utilities on product characteristics using OLS and 2sls estimation. The regression results are listed in 3. The price coefficients are negative which means utility of choosing a product decreases in the price. Since price is endogenous and positively correlated with demand shock, OLS estimation will underestimate the price elasticity. Comparing the estimates of OLS and IV regression, the absolute value of price coefficient is higher in IV regression

which means that the cost shifters address the price endogeneity problem. The second product charateristics is the logarithm of radius per store. This variable is calculated by dividing MSA area by the number of stores in the chain that product brand-chain belongs to. If a chain has fewer stores than the other in the same market, it will have a larger value of radius such that the travelling cost of consumers to purchase the brand-chain is higher. The coefficient on this variable has negative sign as expected which implies that 1% increase of radius will decrease utility by 0.0168. The third variable is dummy for large package size. I define package size less or equal to 12 (regardless of the volume per serving in the pack) as small pack such that dummy equals 1 if pack size is greater than 12. The estimate of large size on utility is -0.25 which implies that controlling for other characteristics on average consumers are less likely to purchase large size. It makes sense since the moving cost, storage cost of large size is higher for consumers not to mention about that beer is perishable goods. Finally, the estimates of selected brand fixed effects are listed. Flagship brands such as bud light, coors light and miller lite have higher utility of cheap brand such as Busch light. Due to the limited substitution effects and I.I.A problem of simple logit demand model, I estimate the random coefficient model with market demographics in the next section.

6.3.2 Random Coefficient Model

In the random coefficient model, I add the interaction terms of consumer's demographics and product characteristics according to 5.3. For each market, I randomly draw 300 consumers and their demographics from known joint distribution of income to poverty ratio and age. The product characteristics which interact with demographics and i.i.d standard normal variable v include retail price, dummy for light beer, dummy for domestic beer and ABV. The purpose of adding these interaction terms is to account for the heterogeneous tastes of consumers with different demographics on product characteristics so that substitution effects among products with similar characteristics are stronger. Identification of coefficients on these interaction terms comes from the different consumption patterns for markets with different demographics. For example, if consumers in market A has higher income than consumers in market B, and we observe increase of one brand's price in both markets does not affect the market share of that brand in market A as much as in market B. That implies that consumers with higher income level are less sensitive to price so that coefficient on income interacting price is positive.

The estimates of random coefficient model is shown in table 4. I use the same instruments in 2slsiv regression. The first column of table 4 is estimates of Θ_1 in comparison to table 3. Estimates of Θ_2 are provided in the last 3 columns. Since I add brand dummies in product charateristics which are linearly correlated with dummy for light beer, dummy for domestic beer and ABV, I apply the minimum-distance estimation following Chamberlain (1982) and Nevo (2000). To interpret the results, fist of all, the price coefficient is -11.308 which means that without considering demographics in the random coefficient, one dollar increase of a brand's price will decrease utility of choosing the brand by 11.308[11]. The interaction term of income and price has estimated coefficient 1.034 which means that the price coefficient of consumer with higher income level is smaller in absolute value. The interaction of income with light, domestic dummies and ABV are alll negative such that high income consumers are less likely to buy light beer, domestic beer and high alcoholic beer. As for elderly consumers, they are more likely to buy light beer and domestic beer but less likely to buy high alcoholic beer. The estimates of demographics in random coefficient provides more flexible substitution among products.

The estimates of Θ_1 are similar to 2slsiv. The coefficient on radius of chain stores is -1.688 similar to -1.684 in 2slsiv. The coefficient on size dummy is still negative. The coefficient on light dummy, domestic dummy and ABV are retrieved from minimum-distance estimation. In the bottom of table 4 I also report the statistics of estimated own price elasticity. The concern is that random coefficient model may have positive price coefficient due to the interaction terms of demographics and price.

The average own price elasticity and cross price elasticity by brand are provided in table 5. These numbers are obtained firstly by aggregating market share of a given brand over size and chain in a market, then calculating partial derivatives of brand shares to brand prices from a representative chain, and finally averaging over MSAs and time periods. The first panel shows own and cross price elasticities for flagship brands in the industry. For example, the first rows shows percentage change of demand for Bud light to price increase of other brands. To interpret, if the price of Bud light increase by 1%, its demand will drop by 4.92%. If the price of Coors light increase by 1%, demand for Bud light will increase by 0.75%. The elasticities are quite reasonable such that substitution among light beer such as Bud light, Coors light and Miller Lite

[11] In the estimation, I do not demean the demographics. Otherwise, the estimates of coefficient on price represents the average effect of price on indirect utility over consumers.

is stronger than substitution between light and lager. Moreover, substitution among domestic brands are stronger than imported brands. Finally, the second panel of table 5 lists almost all the 50 brands in my sample including large brewers and top craft brewers such as New Belgium, Yuengling and Boston brewing.

6.3.3 Supply estimates

With demand estimates, I calculate the markups of retailer and manufacturer, and implicit marginal costs for all products (j, c, m, t) according to the optimal pricing strategy 5.14, 5.19 and 5.20. The average statistics of markups and costs by brewers for pre- and post-merger periods are given in 6. The number is obtained by unweighted averaging key variables over all brands, chains, markets and periods of each firm for pre- and post-merger respectively. For example, in the pre-merger periods (6 quarters) the average marginal cost of one 12 oz serving of Anheuser-Busch beer (regardless of brands) is 10 cents. On average, a retailer's profit of selling one serving of Anheuser-Busch product is 36 cents. Anheuser-Busch's profit per serving is 26 cents.

In order to understand the merger effect, first I compare the marginal costs. Comparing the pre- and post-merger marginal costs for all brewers in the first column, it is clear that in the post-merger periods marginal cost of selling beer brands increases in general. For example, after MillerCoors merger, the cost of Anheuser-Busch increases from 10 cents to 19 cents. This indicates a national level cost shock in the post-merger period. However, the amount of increased marginal costs is smaller for Coors which can be explained by the cost saving of the merger especially the significant reduced shipping cost of Coors. Without estimating cost function, it is hard to tell cost saving due to merger from the common supply shock.

As for the effect of increased market power, I compare the markups of retailers and brewers for Miller and Coors before and after the merger. I find that on average, retailer's profit of selling one serving of Miller or Coors beer decreases by 2 cents. The markup of Coors increases from 15 cents to 19 cents and Miller's increases from 16 cents to 19 cents. The conclusions are twofold. First, upstream brewers' profits increase in their market power. Second, downstream retailers may sacrifice their profits as a buffer to partially offset the positive shock of upstream on retail price. It is very important to note that the decrease of retailer markup is reduced response to both changes of brewer's markup and marginal cost. For example, even without significant change of brewer's markup, retailer's markup of selling Anheuser-Busch, Heineken,

Modelo, Boston and Sierra Nevada decreases. In the counterfactual analysis, I disentangle effects of cost and market power by releasing one effect while controlling for the other. The last column of table 6 provides average quarterly profits of brewers. Miller and Coors profits increase by 4.3 million and 2.2 million after the merger. Profits of Anheuser-Busch also increases by 0.8 million which is mainly due to the shift of demand to Anheuser-Busch due to higher MillerCoors prices.

Given the recovered implicit cost, the OLS regression result of cost equation 6.4 is shown in table 7. The coefficient on distance is 0.013 which measures how cost per serving is correlated with shipping distance. In the bottom of table 7 I calculate the maximum cost saving per serving across MSAs due to reduced shipping distance of the merger. For Miller, merger reduces shipping cost per serving by at most 2.2 cents and for Coors it reduces shipping cost by 7.4 cents. The results corresponds to 3 that Coors benefits more than Miller in terms of shipping distance. The estimates of changes of brand dummies after merger are listed in 7 for selected brands. For example, Coors light has 2 cents cost reduction after merger and Miller Lite has 1.6 cents reduction. These cost reductions other than shipping cost may come from synergy of production after the merger. Most importantly, they can not be estimated without backed out cost level and sample covering both pre- and post-merger. Without information about these estimates (i.e. with only pre-merger data), merger simulation may be less accurate under the fixed environment assumption. Another factor that could affect merger simulation is the residual of the cost regression. The residual that measures unobserved cost is not constant before and after merger which also affects merger simulation. With all these precise estimates of unobserved demand and supply shocks, shipping cost and cost synergy, I can simulate counterfactual without merger in the post-merger period similar to a retrospective analysis in order to, firstly disentangling merger effects and secondly understanding vertical relationship in upstream merger.

7 Counterfactuals

In this section, I simulate several counterfactual scenarios and solve double marginalizations in each counterfactual. In the first counterfactual, I calculate marginal cost in the post-merger period without MillerCoors merger. To do that, I subtract cost saving through shipping distance and production from the backed out marginal cost. I treat this scenario as benchmark. In the second counterfactual, I analyze the cost

saving of merger without consolidation by changing the marginal cost but fixed the ownership matrix of brands. And the third scenario is what I observed and estimate in the sample that both cost saving and upstream consolidation occur. Moreover, I simulate two scenarios without vertical relationship such that brewer sets retail price. These two simulations help to compare merger effects with or without vertical relationship. Moreover, it shows that welfare analysis would be inaccurate if supply side is improperly specified.

The simulation results are listed in table 8. Each column represents a scenario with the first column as benchmark without consolidation and the third column as "observed" merger in the sample. The values in the table for key variables (i.e. cost, markups) are calculated by firm and by concentration of downstream retailers. For instance, Coors(low) means (unweighted) average statistics over all Coors products sold in markets with HHI_{chain} less than 0.28 where 0.28 is the median value of HHI_{chain} across markets. Comparing brewer(high) and brewer(low) illustrates the heterogeneous responses to upstream shocks for markets with different downstream concentration.

First of all, comparing column (1) and (2) shows the change of key variables after cost saving. In the first panel, merger saves the marginal cost of Coors beer by 2.8 cents per 12 oz serving. It also saves the marginal cost of Miller by 1.3 cents in markets with high concentrated downstream and 0.9 cents in less concentrated ones. The differences in costs are mainly due to location of markets rather than vertical relationships. The second panel shows changes of retailer markups with cost saving. The changes are not quite obvious probably due to the averaging. In the third panel, brewer's markups increase after the cost saving. For instance, Coors increases markups by 0.3 cents and 0.2 cents in response to 2.8 cents cost saving per serving. This finding means that cost saving is significant for Coors after the merger, while it does not fully pass through to retail price due to the slight increase of Coors markup. Interestingly, Anheuser-Busch also decreases its markup in order to compete with lower price of MillerCoors.

Column (2) and (3) demonstrate the changes due to increased market power. Marginal costs are the same in the third column to the second column. After increasing market power, retailer markups in low concentrated market on average increases by 0.1 cents while in high concentrated market decrease by 0.1 cents for Coors and decrease by 0.3 cents for Miller[12]. Comparing "low" and "high" markets for other brewers, I

[12]The change of average retailer markups seems tiny even in percentage. One way to improve the comparison could be calculating the change of each market and showing percentile rather than mean.

find that retailers of markets with high concentrated downstream have more power to adjust retail markups than those in "low" markets. As for brewer markups, Miller and Coors markups increase further with more market power. By comparing (1) and (3), it is obvious that increased brewer markups of Miller and Coors dominate the cost saving of the merger.

In the two panels at the bottom of table 8, I calculate the changes of brewers' and retailers' total profits of selling MillerCoors brands and all brands. The second column (2)-(1) shows that with cost saving the total profits of MillerCoors increases by 22 million dollars and retailers' profits incresse by 73 million dollars. The third column shows profits' changes when MillerCoors maximizes joint profits. MillerCoors' profit increases further, whereas retailers' profits decrease significantly. One reason is that the total consumption of Miller and Coors decreases (by $3.27 * 10^8$ servings) due to higher prices and the other reason could be decreased retailers' markups. In the last panel, it shows the changes of surplus and social welfare. Column (2) shows that total profits of all brewers' decrease (by $1.81 * 10^7$) with cost saving of MillerCoors though MillerCoors' profits increase (by $2.2 * 10^7$). Beer consumption shifts from other brands to Miller and Coors due to cost saving and consumer welfare increases. After considering change of market power as in column (3), all brewers' profits increase and consumption shifts from MillerCoors to other brands. Retailers' profits and consumer welfare decreases. By summing up column (2) and (3), it implies the joint effects of cost savings and market power. Consumers and retailers are worse off in this merger which is dominated by the increase of brewers' profits. The social welfare increases.

Finally, the last two columns show brewer markups with only one stage price setting. The findings are three-fold. First, the markups in one stage price setting are much less than those in two stages which corresponds to the reason of building three-tier beer distribution system to discourage beer consumption. Second, brewers' markups are higher in one stage price setting. The reason is that retailers' markups decrease the marginal revenue of oligopolistic prices and therefore brewers charge small markups in a two stage price setting system. Third, similar to the second finding, the effect of increased market power on retail price is smaller with vertical relationship which indicates that downstream market restricts the increase of brewers' markups after the merger.

8 Conclusion

In this paper, I study and quantify the impacts of cost synergy and increased market power of upstream consolidation in the U.S. beer industry. I use Nielsen retail data of beer sales in food stores from 2007-2011 to estimate demand for beer in 50 selected MSA markets. With the estimates of demand, I model vertical relationship in the supply side and assume a Bertrand-Nash linear pricing game between upstream brewers and downstream retailers to estimate double marginalizations. Implicit costs for both pre- and post-merger periods are backed out by subtracting markups from retail price. By regressing recovered costs on supply side shifters such as distance and interaction of brand dummies with merger, I can estimate cost saving through reduced shipping distance and production cost. I find that on average MillerCoors joint venture reduces production cost of Coors light by 2 cents per 12 oz serving, and cost of Miller lite by 1.6 cents. The shipping cost of Coors decreases by 7.4 cents per serving at maximum and shipping cost of Miller decreases by 2.2 cents at maximum.

In order to disentangle cost saving and market power effects of the merger, I simulate several counterfactual scenarios. I find that brewer will increase markups in both scenarios of cost saving and increased upstream market power. Retailers in markets with high concentrated downstream are more likely to adjust retail markups to dampen the shocks from upstream. In the simulation, I find that MillerCoors increase markups after the merger which dominates the cost saving. In terms of welfare, the mega-merger increases MillerCoors profits but hurt retailers' profits markedly. The total consumption of MillerCoors beer for all 50 markets from mid-2008 to 2011 decrease by $1.78 * 10^8$ servings. As for change of total surplus, consumers and retailers are worse off due to the merger which is offset by the increased brewers' profits. The social welfare increases due to MillerCoors joint venture.

References

[1] Ashenfelter, O, Hosken, D. and Weinberg, M.(2015),"Efficiencies Brewed: Pricing and Consolidation in the US Beer Industry", *RAND Journal of Economics*, 46, 328-361.

[2] Asker, J.(2016),"Diagnosing Foreclosure due to Exclusive Dealing",*Journal of Industrial Economics*,64(3), 375-410.

[3] Berry, S.(1994),"Estimating Discrete-choice Models of Product Differentiation", *Rand Journal of Economics*, 25(2), 242-262.

[4] Berry, S., Levinsohn, J. and Pakes, A.(1995),"Automobile Prices in Market Equilibrium",*Econometrica*, 63,841-890.

[5] Chamberlain, G.(1982),"Multivariate Regression Models for Panel Data", *Journal of Econometrics*, 18(1), 5-46.

[6] Chandra, A. and M. Weinberg(2017),"How Does Advertising Depend on Competition? Evidence from U.S. brewing", *Management Science*.

[7] Chen, C.(2014),"Estimating the Foreclosure Effect of Exclusive Dealing: Evidence from the Entry of Specialty Beer Producers", *International Journal of Industrial Organization*, 37, 47-64.

[8] Dearing, A. (2016),"Pricing Policy and Welfare in Retail Chains: A Case for Vertical Incentives", *working paper*.

[9] Faheem, A. and P. Gayle (2017),"Mergers, Competitive Conduct, and Vertical Relationships between Firms: Evidence from US Beer Industry", *working paper*.

[10] Fan, Y.(2013),"Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market", *American Economic Review*, 103, 1598-1628.

[11] Goldberg, L. and Hellerstein, R.(2013),"A Structural Approach to Identifying the Sources of Local Currency Price Stability", *Review of Economic Studies*, 80, 175-210.

[12] Houde, J.(2012), "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline",*American Economic Review*, 102(5),2147-2182.

[13] Hellerstein, R.(2008), "Who Bears the Cost of a Change in the Exchange Rate?Pass-through Accounting for the Case of Beer", *Journal of International Economics*, 76, 14-32.

[14] Lee, R. S.(2013):" Vertical Integration and Exclusivity in Platform and Two-sided Markets",*American Economic Review*, 103(7),2960-3000.

[15] Manuszak, M.(2010),"Predicting the Impact of Upstream Mergers on Downstream Markets with an Application to the Retail Gasoline Industry", *International Journal of Industrial Organization*, 28, 99-111.

[16] Miller, N. and M. Weinberg (2017),"Understanding the Price Effects of the Miller/Coors Joint Venture", *Econometrica*, 85, 1763-1791.

[17] Murry, C.(2017),"Advertising in Vertical Relationships: An equilibrium Model of the Automobile Industry", *working paper*.

[18] Nevo, A.(2000), "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand", *Journal of Economics and Management Strategy*, 9, 513-558.

[19] Nevo, A.(2001),"Measuring Market Power in the Ready-To-Eat Cereal Industry", *Econometrica*, 69, 307-342.

[20] Nevo, A. and Whinston, M.(2010),"Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference", *Journal of Economic Perspectives*, 24(2), 69-82.

[21] Peters, C.(2006), "Evaluating the Performace of Merger Simulation: Evidence from the U.S. Airline Industry", *Journal of Law and Economics*, 49(2), 627-649.

[22] Sweeting, A., and X. Tao (2016),"Dynamic Oligopoly Pricing with Asymmetric Information: Implications for Mergers.",*working paper*.

[23] Villas-Boas, S. B.(2007), "Vertical Relationships Between Manufacturers and Retailers: Inference with Limited Data", *Review of Economic Studies*, 74, 625-652.

[24] Villas-Boas, S. B.(2009), "An Empirical investigation of the Welfare Effects of Banning Wholesale Price Discrimination", *RAND Journal of Economics*, 40(1), 20-45.

[25] Weinberg, M.(2011),"More Evidence on the Performance of Merger Simulations", *American Economic Review: Papers & Proceedings*, 101(3),51-55.

[26] Weinberg, M. and D. Hosken (2013),"Evidence on the Accuracy of Merger Simulations", *Review of Economics and Statistics*, 95(5),1584-1600.

[27] Yang, C. (2017), "Could Vertical Integration Increase Innovation", *working paper*.

Appendix:

Figure 1: average price by brand over 50 markets

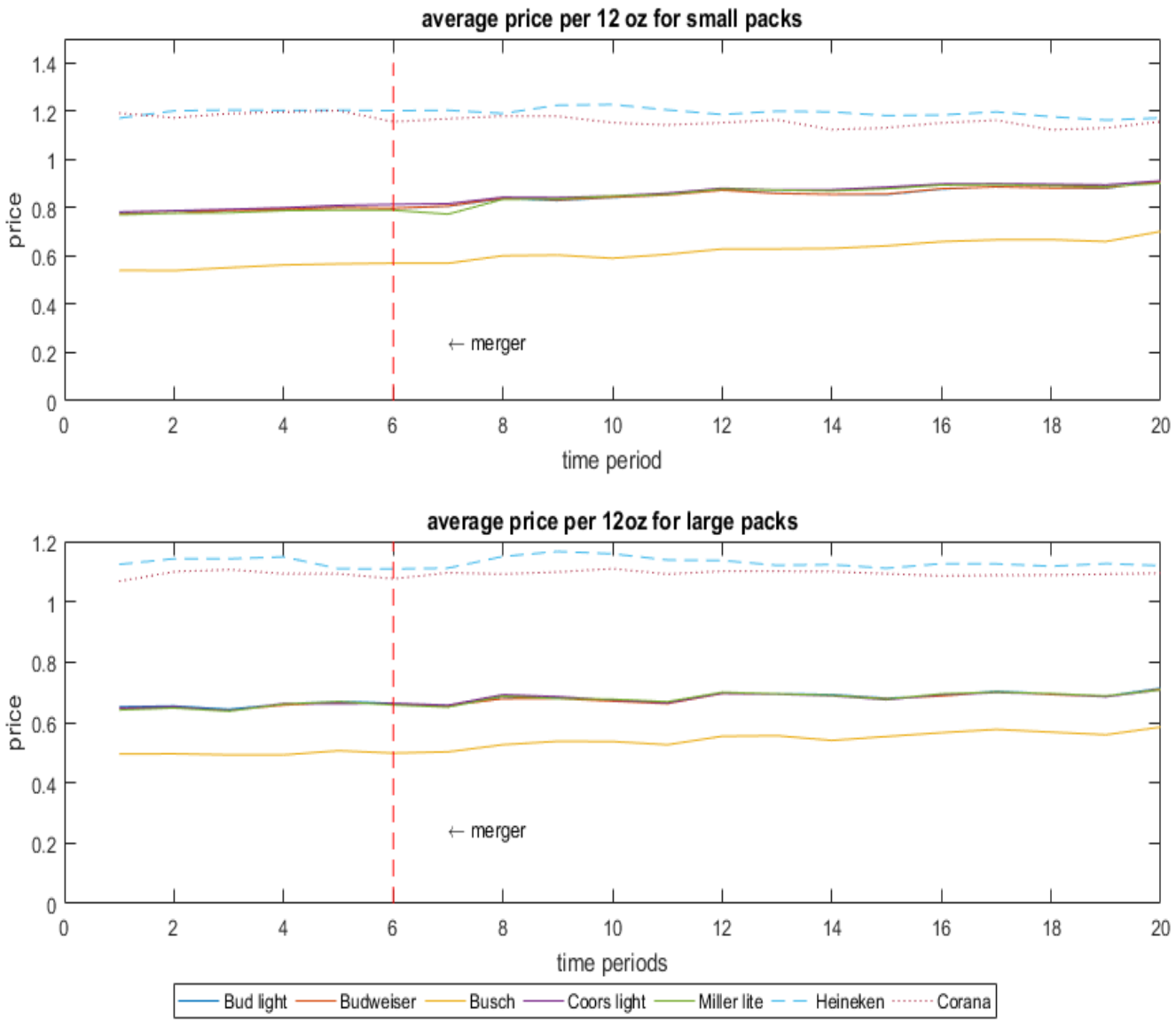


Figure 2: Distribution of HHI increases for brewers after merger

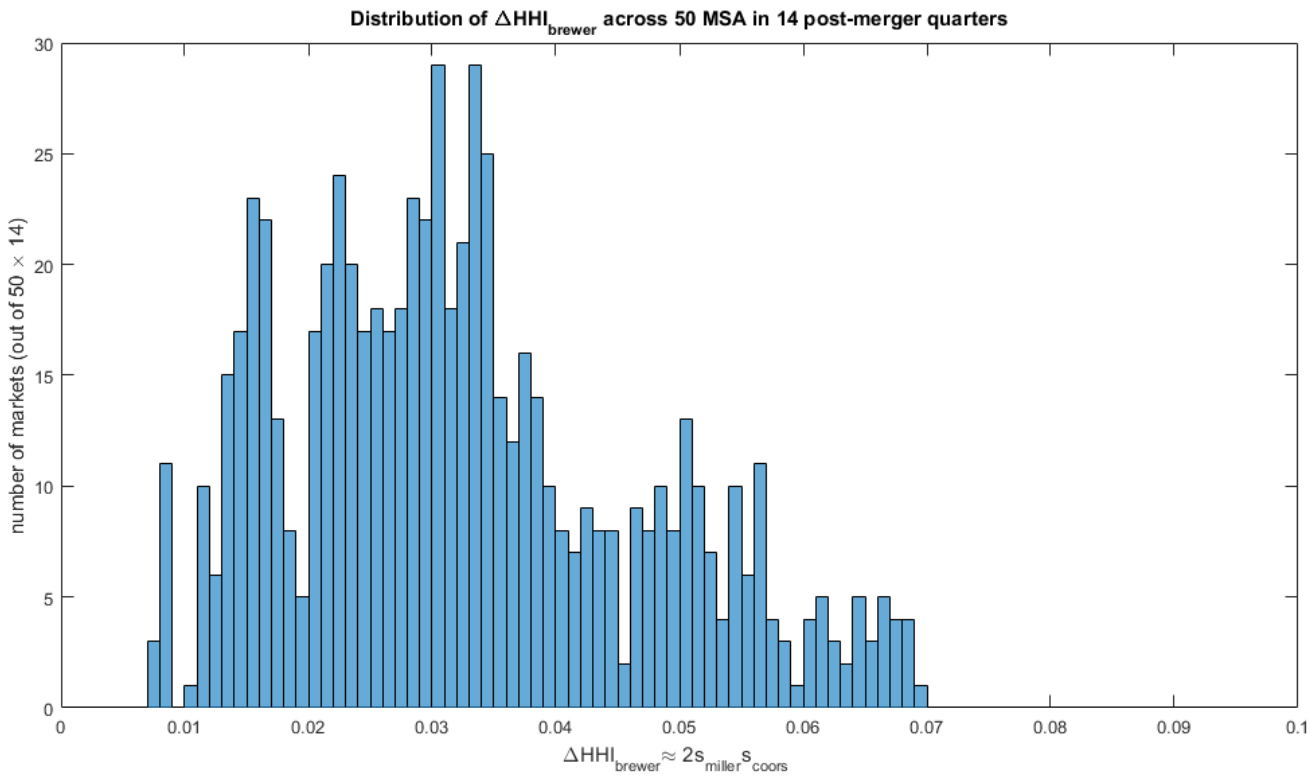
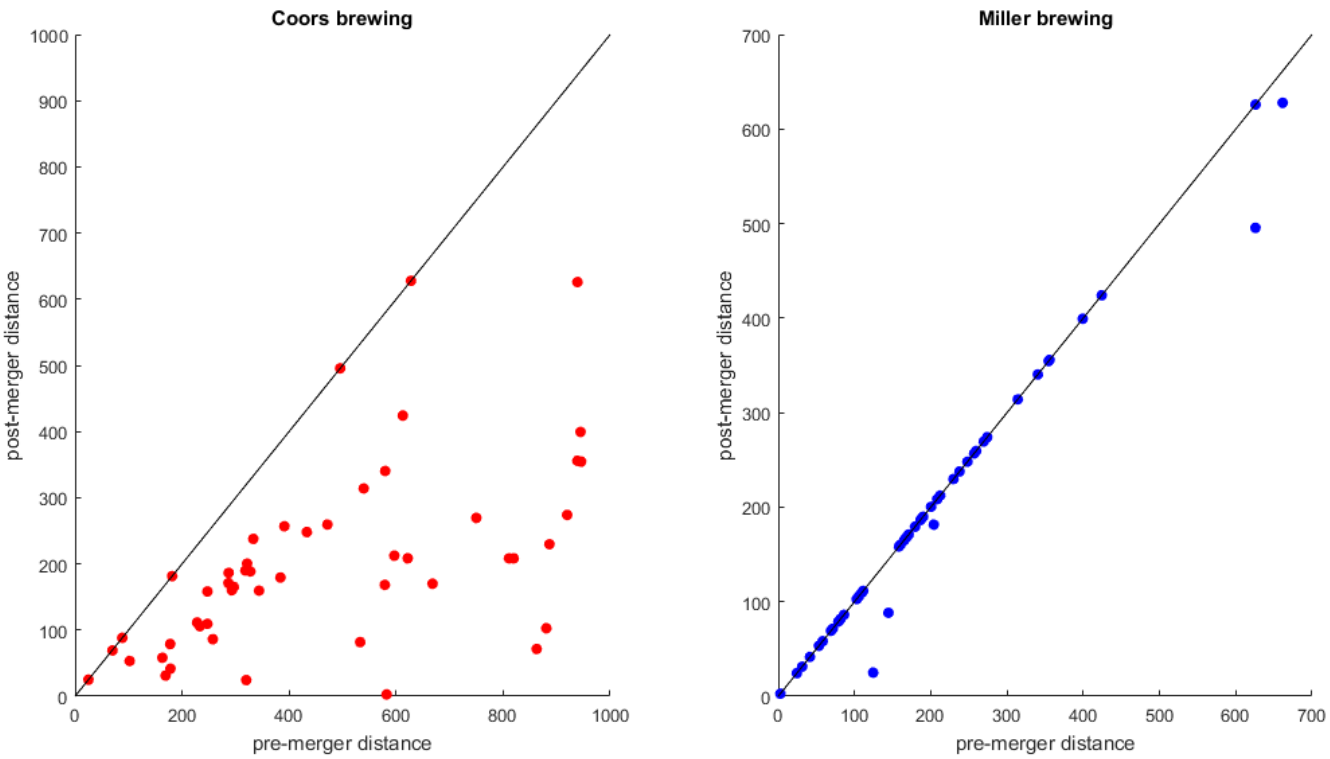


Figure 3: distance(miles) between brewers and 50 markets for pre/post-merger



Note: distance is calculated as shortest distance of MSA to breweries.

Table 1: The 50 MSAs with average statistics over 20 quarters

Market	No. chains	No. products	inside share	market size(10^7 oz)	DMA food coverage
Asheville	1	77	0.63	9.56	0.77
Augusta	3	202	0.88	5.50	
Boise City	2	135	0.71	6.21	
Charleston	3	198	0.87	9.84	
Charlotte	3	215	0.85	33.4	0.86
Charlottesville	3	210	0.77	4.06	
Chattanooga	2	134	0.76	3.96	
Chicago	2	131	0.66	86.7	0.65
Cincinnati	1	73	0.86	21.8	0.64
Columbia	2	137	0.84	9.32	
Columbus	2	139	0.81	18.5	0.67
Davenport	1	72	0.81	5.18	
Durham	3	200	0.79	8.16	0.77
Fayetteville	1	68	0.79	3.86	0.77
Florence	3	176	0.87	1.94	
Fresno	4	220	0.61	8.61	
Greensboro	2	142	0.88	10.9	
Greenville	3	191	0.86	9.23	0.77
Hickory	2	98	0.71	3.54	0.86
Houston	3	153	0.71	47.2	0.50
Jacksonville	2	137	0.62	12.3	0.47
Kingsport	3	204	0.83	3.38	
Knoxville	3	173	0.71	9.82	
Lafayette	3	179	0.79	2.53	
Lake Havasu	3	204	0.82	6	0.84
Las Vegas	3	198	0.77	21.7	0.76
Lynchburg	2	137	0.87	4	
Manchester	3	197	0.67	12.1	0.83
Medford	3	179	0.68	1.88	
Milwaukee	2	105	0.81	20.9	0.72
Myrtle Beach	3	199	0.90	7.74	

continued from previous page

Market	No. chains	No. products	inside share	market size(10^7 oz)	DMA food coverage
Nashville	2	131	0.76	12.1	0.60
Oxnard	3	180	0.60	12	0.52
Phoenix	3	207	0.79	58.2	0.84
Prescott	3	201	0.81	3.59	0.84
Raleigh	2	146	0.74	21.1	0.77
Richmond	2	138	0.80	19.5	0.81
Roanoke	2	142	0.86	5.81	
Salinas	3	160	0.64	3.94	
San Francisco	2	110	0.57	35.5	0.48
Santa Barbara	3	173	0.54	6.61	
Santa Rosa	3	146	0.65	6.09	0.48
Shreveport	2	120	0.67	3.38	
Spartanburg	3	200	0.88	3.99	0.77
Tampa	2	143	0.72	26.3	0.29
Toledo	1	72	0.76	5.09	
Tucson	3	197	0.72	14.1	
Virginia Beach	2	146	0.72	33.3	
Wilmington	2	142	0.89	9.77	
Winston	2	141	0.88	6.44	

Table 2: OLS regression of retail price(12oz) on HHI

		log(price)	
HHI-brewer	0.463	0.244	
	(0.144)**	(0.148)	
HHI-retailer	0.279	0.062	
	(0.149)	(0.152)	
HHI-brewer ×HHI-retail	-1.47	-0.871	
	(0.413)**	(0.418)*	
post-merger×			
HHI-brewer		0.193	
		(0.038)**	
HHI-retailer		0.208	
		(0.043)**	
HHI-brewer ×HHI-retailer		-0.675	
		(0.147)**	
Δ HHI-brewer			1.762
			(0.26)**
HHI-retailer × Δ HHI-brewer			-1.232
			(0.60)*
large pack	-0.659	-0.67	-0.659
	(0.002)**	(0.147)**	(0.002)**
log(Distance)	0.031	0.031	0.031
	(0.001)**	(0.001)**	(0.001)**
Year dummies	X	X	X
Season dummies	X	X	X
Brand dummies	X	X	X
Market dummies	X	X	X
R-square	0.92		
Observation	155,973		

** 1-percent or * 5-percent level significant

Table 3: Demand estimates from simple logit model

variable	OLS	IV
price	-3.941	-5.79
	(0.025)**	(0.172)**
log(radius per store)	-1.716	-1.684
	(0.004)**	(0.005)**
large size	-0.043	-0.253
	(0.005)**	(0.020)**
Bud light	1.652	0.991
	(0.021)**	(0.064)**
Budweiser	0.868	0.204
	(0.021)**	(0.064)**
Natural light	-0.315	-1.315
	(0.023)**	(0.094)**
Busch light	-0.833	-1.858
	(0.024)**	(0.097)**
Miller lite	0.753	0.090
	(0.021)**	(0.064)
Miller high life	-0.820	-1.814
	(0.023)**	(0.094)**
Coors light	0.993	0.341
	(0.021)**	(0.063)**
Heineken	1.014	1.099
	(0.019)**	(0.021)**
Corona extra	1.448	1.451
	(0.019)**	(0.020)**
constant	7.451	9.335
	(0.099)**	(0.200)**
Year dummies	X	X
Season dummies	X	X
Brand dummies	X	X
Market-Chain dummies	X	X
min.brand dummy	-1.97	-3.10
max.brand dummy	1.65	1.45

** 1-percent or * 5-percent level significant

Table 4: Demand estimates from random coefficient model

		Interaction with:		
variable	mean in population	unobservable	income	age
Price	-11.308	0.943	1.034	0.793
	(1.237)**	(0.650)	(0.301)**	(0.638)
Large size	-0.360	0.898		
	(0.058)**	(0.249)**		
Light	4.175	0.863	-0.578	0.625
	(0.129)**	(0.419)*	(0.125)**	(0.161)**
Domestic	5.221	0.753	-0.140	0.722
	(0.120)**	(0.167)**	(0.105)	(0.249)**
ABV	-0.148	1.030	-0.107	-0.179
	(0.015)**	(0.055)**	(0.055)*	(0.136)
log(radius per store)	-1.688			
	(0.011)**			
Bud light	1.414			
	(0.429)**			
Budweiser	-0.128			
	(0.497)			
Miller lite	0.513			
	(0.430)			
Coors light	0.770			
	(0.429)			
Heineken	1.124			
	(0.026)**			
Corona extra	1.432			
	(0.091)**			
constant	14.583			
	(1.175)**			
Year dummies	X			
Season dummies	X			
Brand dummies	X			
Market-Chain dummies	X			
own price elasticity> 0	0%			
own price elasticity> -1	0.0064%			
Obervations	155,973			

** 1-percent or * 5-percent level significant

Table 5: Own and Cross Price Elasticity (average over markets)

Cross-price elasticity

	Bud light	Budweiser	Coors light	Corona extra	Heineken	Miller Lite
Bud light	-4.924	0.364	0.757	0.096	0.066	0.644
Budweiser	0.832	-5.071	0.455	0.225	0.179	0.390
Coors light	1.362	0.363	-5.593	0.096	0.066	0.644
Corona extra	0.479	0.495	0.255	-6.084	0.361	0.227
Heineken	0.468	0.549	0.250	0.488	-6.465	0.221
Miller Lite	1.366	0.364	0.758	0.096	0.066	-5.644

Own-price elasticity

Anheuser-Busch			
Natural light	-4.602	Michelob light	-6.059
Natural ice	-4.563	Stella Artois	-7.118
Busch light	-4.506	Rolling rock	-5.387
Michelob amber bock	-5.717	Beck's	-6.643
Michelob ultra light	-6.091	Bud ice	-5.291
Budweiser select light	-5.601	Bud light lime	-6.512
Budweiser select	-5.216	Busch	-4.298
Coors			
Keystone light	-5.593	George killians	-6.086
Coors banquet	-5.549	Blue moon	-6.434
Heineken			
Tecate	-5.900	Dos equis especial	-6.306
Newcastle brown ale	-6.748		
Miller			
Miller genuine draft	-5.455	Milwaukee's best	-3.879
Miller genuine draft light	-5.573	Miller chill light	-6.459
Miller high life light	-4.722	Steel reserve 211	-4.011
Milwaukee's best light	-4.155	Icehouse	-4.909
Milwaukee's best ice	-4.052		
Modelo			
Corona light	-6.814	Pacifico	-6.670
Modelo especial	-6.287		
New belgium			
Fat tire amber ale	-6.879		
Pabst brewing			
Pabst blue ribbon	-4.354		
Sierra nevada			
Sierra nevada pale ale	-6.703		
Yuengling			
Yuengling	-5.060		
Boston brewing			
Samuel adams	-6.366		

Table 6: Statistics on estimated markups and costs(average over products)

	marginal cost		retailer markup		brewer markup		qtrly profit
Firm Name	mean	sd	mean	sd	mean	sd	(in \$ 10^7)
Anheuser-Busch(pre)	0.10	0.51	0.36	0.46	0.26	0.044	5.35
Anheuser-Busch(post)	0.19	0.44	0.34	0.38	0.25	0.043	5.43
Coors(pre)	0.25	0.49	0.35	0.44	0.15	0.018	0.70
Coors(post)	0.27	0.43	0.33	0.38	0.19	0.033	1.13
Miller(pre)	0.08	0.52	0.35	0.47	0.16	0.033	2.32
Miller(post)	0.14	0.43	0.33	0.38	0.19	0.036	2.54
Heineken(pre)	0.56	0.44	0.37	0.42	0.17	0.016	0.30
Heineken(post)	0.58	0.37	0.35	0.35	0.17	0.016	0.31
Modelo(pre)	0.59	0.46	0.38	0.45	0.18	0.015	0.43
Modelo(post)	0.60	0.37	0.35	0.36	0.17	0.015	0.42
Boston(pre)	0.59	0.48	0.41	0.46	0.17	0.011	0.06
Boston(post)	0.66	0.39	0.38	0.38	0.17	0.012	0.07
New belgium(pre)	0.74	0.27	0.31	0.29	0.16	0.016	0.01
New belgium(post)	0.78	0.37	0.37	0.36	0.17	0.018	0.02
Pabst(pre)	0.05	0.49	0.35	0.48	0.12	0.006	0.05
Pabst(post)	0.14	0.40	0.33	0.38	0.12	0.005	0.08
Yuengling(pre)	0.25	0.29	0.32	0.28	0.14	0.006	0.06
Yuengling(post)	0.29	0.28	0.32	0.27	0.14	0.007	0.07
Sierra Nevada(pre)	0.66	0.49	0.41	0.48	0.17	0.013	0.04
Sierra Nevada(post)	0.73	0.39	0.39	0.38	0.17	0.012	0.05
$markup^r < 0$	0%						
$markup^w < 0$	0%						
$mc < 0$	13%						
obs	155,973						

Table 7: OLS regression on marginal cost(12oz)

	$mc^r_{jcmt} + mc^\omega_{jmt}$
log(Distance)	0.013
	(0.0006)**
log(Gross rent)	-0.04
	(0.002)**
log(Wage)	0.02
	(0.015)
post-merger×	
Coors light	-0.02
	(0.005)**
Coors Banquet	-0.02
	(0.005)**
Blue moon	-0.019
	(0.007)**
George killian	-0.017
	(0.008)*
Miller lite	-0.016
	(0.005)**
Miller high life	-0.008
	(0.005)
Miller chill light	-0.13
	(0.008)
Miller genuine draft	-0.003
	(0.005)
Year dummies	X
Season dummies	X
Brand dummies	X
Market dummies	X
R-square	0.90
Observation	155,973
cost saving through $\Delta log(Distance)$	max($ per 12oz)
Coors brand(12oz)	-0.074
Miller brand(12oz)	-0.022

** 1-percent or * 5-percent level significant

Table 8: Couterfactuals: average cost and markups by firm and HHI_{chain}

	two tiers			one tier	
Brewer name	no merger	costsaving	costsaving+power	costsaving	costsaving+power
marginal cost					
coors(low)	0.440	0.412		0.440	0.412
coors(high)	0.158	0.130		0.158	0.130
miller(low)	0.318	0.305		0.318	0.305
miller(high)	0.003	-0.006		0.003	-0.006
AB(low)	0.341				
AB(high)	0.055				
heineken(low)	0.669				
heineken(high)	0.493				
retailer markups					
coors(low)	0.225	0.224	0.225		
coors(high)	0.457	0.457	0.456		
miller(low)	0.213	0.213	0.214		
miller(high)	0.446	0.447	0.444		
AB(low)	0.224	0.225	0.224		
AB(high)	0.455	0.456	0.452		
heineken(low)	0.250	0.251	0.249		
heineken(high)	0.481	0.483	0.478		
brewer markups					
coors(low)	0.159	0.162	0.194	0.177	0.212
coors(high)	0.150	0.152	0.203	0.171	0.228
miller(low)	0.156	0.156	0.189	0.170	0.206
miller(high)	0.173	0.172	0.199	0.192	0.222
AB(low)	0.249	0.243	0.253	0.263	0.273
AB(high)	0.260	0.254	0.266	0.284	0.296
heineken(low)	0.171	0.171	0.171	0.189	0.188
heineken(high)	0.174	0.175	0.173	0.197	0.196
Only for MillerCoors		(2)-(1)	(3)-(2)		(5)-(4)
total brewer profits$		$2.2 * 10^7$	$4.5 * 10^7$		$5.6 * 10^7$
total retailer profits$		$7.3 * 10^7$	$-1.78 * 10^8$		
total cost saving$		$-3.53 * 10^7$	$2.34 * 10^7$		
total servings		$1.49 * 10^8$	$-3.27 * 10^8$		$-3.8 * 10^8$
All firm		(2)-(1)	(3)-(2)		(5)-(4)
total brewer profits$		$-1.81 * 10^7$	$1.28 * 10^8$		$1.68 * 10^8$
total retailer profits$		$2.56 * 10^7$	$-5.40 * 10^7$		
total cost saving$		$-4.46 * 10^7$	$2.86 * 10^7$		
total servings		$3.99 * 10^7$	$-7.86 * 10^7$		$-5.52 * 10^7$
total consumer welfare$		$5.7 * 10^7$	$-1.17 * 10^8$		$-1.78 * 10^8$
total welfare$		$6.45 * 10^7$	$-4.3 * 10^7$		$-1.00 * 10^7$

Note: "high" indicates $HHI_{chain} > 0.28$ where 0.28 is the median HHI over markets in post-merger