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

Bo Bian*
Shanghai University of Finance and Economics

September 21, 2022

Abstract

This paper studies the 2008 MillerCoors joint venture in the U.S. beer industry through the framework of a vertically related market. The cost efficiency and increased upstream market power impacts of this merger are quantitively measured and, more importantly, the effect of downstream concentration on pass-through of upstream merger in a vertical relationship is studied. In a vertical relationship, the upstream shock does not fully pass through to retail price because both upstream and downstream firms adjust their pricing post-merger. Downstream concentration not only determines the markups charged by retailers but also affects the capability of upstream firms to exercise their market power. Estimating demand side and supply side in a linear pricing model with double marginalization uncovers the changes of costs and markups using pre/post-merger retail data. The results show that average cost saving of producing a 12 oz serving is 9.27% for Coors and 7.23% for Miller. Brewers' markups increase, while retailers' markups decrease to mitigate the merger impact on retail prices especially for more concentrated downstream markets. The effect of market power is greater than cost saving in this merger. The brewers profit gain dominates the welfare losses of consumers and retailers thereby increasing social welfare in aggregate.

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

---

*Assistant Professor, College of Business, Shanghai University of Finance and Economics. Mailing address: Office 322, College of Business, 100 Wudong St., Shanghai, China 200433. Email: **bian.bo@mail.shufe.edu.cn**. I thank Mark Roberts, Paul Grieco, Charles Murry and Peter Newberry for their comments and discussions on this paper. I am solely responsible for any errors.

1 Introduction

This paper studies the mega merger between the second and the third largest beer manufacturers, SABMiller and Molson Coors, in the U.S. beer industry in 2008. At that time, Miller's market share was 18% and Coors' 11% which gave the joint firm MillerCoors a 29% market share in comparison to the 49% market share of the largest firm Anheuser-Busch. It is standard to evaluate the merger by analyzing the effects, cost savings and increased market power. On the one hand, cost savings will decrease the post-merger retail prices and on the other hand consolidation will increase firm's market power to charge higher prices.

In the literature, there are mainly two types of merger studies. First, a merger can be studied retrospectively with pre- and post-merger data. Usually, a retrospective study uses a reduced form analysis. A predictive study applies a structural model to simulate a merger's effects with only pre-merger data and assumes a fixed market environment or fixed unobserved shocks after the merger. My paper uses both types of studies. A structural model is built to analyze the MillerCoors merger with both pre- and post-merger data. The benefits are twofold. First, with both pre- and post-merger data, changes of unobserved demand and supply shocks are estimated which cannot be obtained with only pre-merger data. Second, with structural model approach, I quantify and disentangle the welfare changes of this merger in terms of implicit marginal costs, markups and consumer welfares in comparison to a reduced form analysis. In addition, markups are further divided in a vertical structure of beer distribution so that the role of downstream market structure on the pass-through of upstream merger can be evaluated. Ignoring the vertical relationship in a horizontal merger may lead to incorrect welfare conclusion. The findings of this paper can be applied to merger evaluation when upstream and downstream markets have imperfect competition.

This paper models both demand and supply-side decisions. The demand side estimates consumers' discrete choices among differentiated beer brands. In the supply side, the three-tier distribution system is modelled by a two-stage linear price setting game between upstream brewers and downstream retailers and resolved using backward induction[1]. In the second stage, downstream retailers set optimal retail prices given the wholesale prices set by brewers. In the first stage, brewers set optimal wholesale

[1]The distributors in the middle tier is not explictly modelled due to data availability. The upstream player can be deemed as a joint decision maker of brewer and distributor as in Hellerstein (2008), Goldberg and Hellerstein (2013).

prices anticipating the best response of retailers. Given the demand estimates, double markups can be computed and joint implicit marginal costs of retailers and brewers are uncovered without observing wholesale prices following the approach of Villas-Boas (2007). To evaluate this merger, I simulate the counterfactuals of no merger and merger with only cost saving but not exercising market power (by maximizing joint profit) in the post-merger period to disentangle cost saving effects and market power effects. Moreover, the cross-sectional variation of downstream concentration and retailer markups helps to understand the role of downstream concentration in upstream merger evaluation.

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). They study the MillerCoors merger using reduced form analysis to examine how increased market power (measured by change of Herfindahl-Hirschman index (HHI)) and cost saving (measured by reduced shipping distance) affect the final retail prices of beer. However, the change of retail prices does not fully reflect the magnitude change of marginal cost or change of markup. The pass-through rate in a vertically related market is less than one where one dollar cost saving does not lead to one dollar price decrease (Hellerstein (2008), Goldberg and Hellerstein (2013)). With a structural model, the cost saving is quantified by separating from adjustment of markups to cost changes. Moreover, changes of markups can be disentangled into two parts, adjustment to cost saving and exercise of market power.

Second, this paper contributes to the structural merger analysis with both pre- and post-merger data. As described 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, the structural model cannot account for changes in estimated demand or supply shocks after the merger. There are several studies about the accuracy of merger simulation such as Peter (2006), Houde (2012), Weinberg (2011) and Weinberg and Hosken (2013). With a long sample period covering this merger, this study accounts for the change of unobserved shocks and simulates the no merger scenario for the post-merger period such that the merger analysis does not suffer from the limitation noted 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 in a vertical relationship which could be a concern in merger evaluation. Vertical relationship is crucial to affect the results of a horizontal merger. Are retailers in a concentrated market more likely to dampen/amplify the upstream shocks than in a comeptitive market? This question is related to a large stream of literature on countervailing power. The countervailing power hypothesis proposed by Galbraith (1952) suggests that powerful retailers can lower the wholesale prices against the wholesalers. Thus, the relative power of downstream versus upstream firms matters for final price and consumer welfare. Theoretical papers usually focus on either upstream or downstream competition and how it affects prices (Katz (1987), Horn and Wolinsky (1988), Snyder (1996),von Ungern-Sternberg (1996), Dobson and Waterson (1997), Tyagi (2001), Chen (2003), Dana (2012), Loertscher and Marx (2019)). Empirical studies of countervailing power include Melnik et. al. (1992), Sorensen (2003), Ellison and Snyder (2010) in the healthcare industry, which also holds one side of the vertical contract fixed and studies how changing competition of the other side affects prices. This paper differs from the literature in that it compares the countervailing powers that markets with different downstream concentrations exhibit to upstream merger rather than the differential power of individual buyers.

There are several works related to this paper which also study beer industry including Hellerstein (2008), Goldberg and Hellerstein (2013), Dearing (2016), Miller and Weinberg (2017), Sweeting et. al. (2021). Hellerstein (2008), Goldberg and Hellerstein (2013) that study the pass-through of cost shocks to retailer prices in the beer industry. Similarly, this paper studies the pass-through of cost saving and increased market power to retail prices due to upstream merger. Dearing (2016) studies how upstream market power affects the downstream chain's choice of uniform pricing policy across stores. In comparison, Villas-Boas (2009) studies the welfare effects of uniform wholesale pricing. Miller and Weinberg (2017), Sweeting et. al. (2021) study the merger of MillerCoors from the aspect of collusion or incomplete information on marginal costs. However, these papers do not study the downstream market concentrations. One work quite close to this paper is by Manuszak (2010) which studies the impact of upstream mergers on downstream markets in the gasoline industry. However, in his paper, the downstream retailer is affiliated with a single upstream supplier unlike beer industry in which a retailer carries multiple suppliers' products. Thus, the upstream merger has a broader impact on downstream retailers in the beer industry than gasoline industry with single-supplier stations.

There are many empirical literatures on vertical relationship of various topics. 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 vertical relationships between manufacturers and retailers. Murry (2017) studies the advertising incentives between dealers and manufacturers in the automobile industry. Yang (2020) studies product innovation in vertical relationship of the smart phone industry. This paper focuses on downstream structure on the pass-through of an upstream merger. The rest of this paper is structured as follows. Section 2 and 3 introduces the background of the U.S. beer industry and data. Section 4 shows preliminary regression results. The empirical model is presented in section 5 with estimation results in section 6. Section 7 simulates counterfactuals and analyzes welfare change of the merger. The paper is concluded in section 8.

2 Background of the U.S. Beer Industry

The U.S. beer industry has a long history and has matured in recent years. Unlike other industries, the beer industry is highly regulated by the government. After the repeal of Prohibition (1919–1933), when individual states were 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 common one is the "three-tier distribution system" feature of beer distribution.

In the three-tier system, beer brewers 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. Some states have adopted an exception allowing small craft brewers to sell beer directly to retailers as long as 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 is what led to the Prohibition in the first place. In principle, brewers are free to choose the beer distributor and distributor is free to choose the portfolio of brands to carry. Finally, it is up to the

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

retailers who decide which brands to put on the shelves. Within the three-tier system, any kind of vertical integration is discouraged. However, some brewers still try to build special relationship with their distributors. For example, Anheuser-Busch has exclusive contracts with some distributors who can only sell Anheuser-Busch brands. The three-tier distribution system which is in common across states justifies the vertical relationship model built for the supply side in my paper.

At the wholesale level, there are many distributors serving each state. Each distributor has exclusive distribution territory. In most states, distributors form alliances or associations. Since data on distributors is not available, I integrate the brewers and distributors as one layer assuming that they jointly set wholesale prices to retailers. Therefore, the three-tier system is reduced into a double marginalization model. Moreover, uniform wholesale prices charged to retailers within states are encouraged by many states or wholesaler associations. Thus, discriminatory prices are excluded from the model. That being said, an individual buyer does not exhibit any bargaining power to countervail the wholesale price. Instead, the downstream market structure impacts the wholesale price setting strategy.

Since each state has its own regulation on alcohol consumption, they can be categorized into control or non-control states at the retail level. In control states, wine and spirits are not allowed to be sold at 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 selling beer with small alcohol volume (ABV) (i.e., Colorado, Kansas, Minnesota, Oklahoma, and Utah.). Beer sales in gas stations, convenience stores, or pharmacies also vary across states.

As for the style of beer sold in the U.S., they can be categorized into lager, 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. Light and lager beers are quite similar except that light beer 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 a 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 and approved by Department of Justice (DOJ), 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 important to evaluate this merger and related questions. As the DOJ stated in its closing statement, 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 changes of this mega merger.

3 Data

This paper uses quarterly sample data from 2007 to 2011 which covers 6 pre-merger quarters and 14 post-merger quarters. Data come from several sources. The beer prices and sales data are collected from Nielsen retail scanner data, which has records of the weekly sales of all beer products in more than half U.S. retail stores across the country. The demographics data are collected from American Community Survey. Beer characteristics are also collected from brewers' websites. These data are used for demand estimation. In the supply side, some cost control variables are collected including average weekly wages in each geographic area provided by Quarterly Census of Employment and Wages (QCEW) of Bureau of Labor Statistics, gross rent from ACS, prices of hop and malt, and more importantly the shipping distance between breweries and markets. The distance is calculated for any pair of location spots using ArcGIS. Details of the main datasets and sample construction are described below.

3.1 Nielsen Retail Data

The Nielsen retail scanner dataset is at store-product-week level which records weekly sales of all beer products sold via participating stores. The product comes at the Universal Product Code (UPC) level which could vary in pack size, container, and volume per container of 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 the number of stores in a chain of a geographic area can be counted. The chains are categorized into different channels including food store, drug store, mass merchandiser, liquor store, and convenience store. The coverage of Nielsen across channels is different. As the coverage rate of liquor and convenience store is quite low in Nielsen data, only beer sales in food chains are considered in the paper.

Each observation includes the price, quantity, 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. Other than the beer type, characteristics such as ABV, carb, calorie, whether it is domestic or imported brand and its brewer are collected as supplement. To reduce the number of products, UPCs are aggregated into brand-package size level by summing over different containers and volumes per serving. The reason to distinguish package size is that price per 12 oz serving is quite different between large and small packs[3]. Large pack (small pack) includes UPCs with more than (less than) 12 packs. The number of total products is further reduced by aggregating sales across stores of the same chain. This aggregating implicitly assumes that chain sets uniform retail prices across stores as studied by Dearing (2016). If stores are allowed to set individual price, the number of "products" will be quite large and the market becomes more competitive than price setting by few chains. Thus, the sample consists of weekly sales of a brand-size-chain products in a geographic market which is defined at the end of this section. Product characteristics are captured by chain dummy and a constructed variable, the ratio between the market area and number of stores in the chain, as proxy for traveling cost of store visit.

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

[3]In general, large pack has smaller price per serving than small pack. Under this definition, I have at most two products for the same brand.

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. Usually, researchers use Current Population Survey(CPS) to generate demographics. However, CPS is not appropriate to analyze geography smaller than a state[4]. The demographics are mainly used in the demand estimation.

ACS and QCEW also provide data related to local retailer costs. From ACS, gross rent is collected in each market as proxy for commercial rent of retail chains. Average wage in the supermarket industry is provided by QCEW as another control for retailer costs. I also collect malt and hop prices as control for brewers' production costs. More cost controls can be obtained by searching for cost related data of where the breweries are located.

3.3 Shipping distance

One important cost factor is the shipping distance from breweries to the geographic markets which is an important argument of cost saving for MillerCoors merger. Ashenfelter, Hosken and Weinberg (2015) finds the reduced shipping distance significantly accounts for post-merger price changes. I use the same method to calculate the shipping distance. First, I locate all plants of Anheuser-Busch, Miller, Coors, and other domestic brewers[5]. Then, using ArcGIS, I compute the distance between the geographic market and closest plant of a brewer as shipping distance. It implicitly assumes that a brewers' plant produces the whole product line of brands. Given the fact that only selected nationwide brands are included in my sample, this assumption seems plausible. Small brewers such as Yuengling, Sierra Nevada, and New Belgium brewery only have few plants and mainly serve nearby markets which still satisfies the assumption. As for imported brands, I follow Miller and Weinberg (2017) approach to calculate distance of markets to the ports. In the post-merger periods, shipping distance of MillerCoors products is calculated as distance between the market and the closest plant of either Miller or Coors. Its difference to pre-merger distance reflects cost saving.

[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.

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

3.4 Market Definition

The detail of store location is private. Instead, Nielsen only reveals information on store's 3-digit zip code, its county, and state. A market is defined 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 has proper area size for retailers and wholesalers to compete. This market definition is simmilar to Ashenfelter, Hosken and Weinberg (2015). Stores are matched to markets according to their counties.

The market size is defined as aggregate quarterly beer sales through all Nielsen participating stores in an MSA. It sums the observed beer sales of two main channels: food store and drug store. Sales through other channels are omitted due to the low coverage rate of stores. The coverage rate is one limitation to properly define the market size and furthermore the market shares. Though Nielsen covers 50% food stores in the U.S., the coverage rate varies much across locations[6]. If the uncovered stores belong proportinally to chains in Nielsen, the in-sample market shares approximate the true shares. However, there are chains not covered by Nielsen dataset which results in overestimation of the market share and biased measure of concentration rate in the market. Miller and Weinberg (2017) also encounter such issues using IRI data. They scale the observed sales by 1.5 as the market size. An alternative way of defining market size is proposed in Hellerstein (2008) which scales the population by beer consumption per capita. The problem of applying her method is that the beer consumption per capita in food stores are not available at MSA level. Moreover, it makes the outside market share very large and leads to underestimation of concentration rate of chains covered in sample.

Given the various coverage rates, the sample only includes selected MSAs. The selection criteria are per capita consumption of beer (calculated by dividing Nielsen beer sales by MSA population) greater than two servings per month. I further select the MSAs with beer sales through food store channels greater than 70% and population larger than 0.2 million. Thus, only MSAs with high food store coverage and beer sales are left. These MSAs are appropriate for the empirical analysis because measure of downstream concentration requires high coverage rate of stores in data. Due to the data limitation, drug stores or convenience stores are not modelled. Finally, there are

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

50 MSAs over 20 quarters in sample. Appendix 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

This section starts the analysis by showing some key variables in the data and preliminary regression results on retail prices. First, figure 1 shows the dynamics of average retail price per 12 oz serving for selected brands over 20 quarters. Though the variation of average prices is not quantitatively large, the price trend is obvious especially for small packs. For instance, the average retail price of Bud Light, Budweiser, Coors Light, and Millter Lite increase by 10 cents (12.5%) per serving for small packs. The increase is less obvious for large packs since large packs often have discounted prices. One interesting finding is that both MillerCoors and Anheuser-Busch increase prices in comparison to imported beer such as Heineken and Corona which tend to decrease their prices. The intuition is that MillerCoors raises their prices as an exercise of their increased market power. Since competing brands have higher prices, Anheuser-Busch is also able to raise its prices. The slight change of imported brands may indicate weak substitution between domestic and imported beer. The two key driving forces on prices after the merger are upstream concentration and cost saving. Figure 2 shows the changes of upstream HHIs of MSAs and figure 3 shows the shipping distances for Coors and Miller.

Figure 2 is obtained by calculating proxy for HHI change by formula $\Delta HHI_{brewer} \approx (s_{miller} + s_{coors})^2 - s^2_{miller} - s^2_{coors} = 2 * s_{miller} * s_{coors}$ similar to Ashenfelter, Hosken and Weinberg (2015). One caveat is that overestimated market share may lead to overestimated change of HHI. The histogramm of figure 2 shows frequencies of ΔHHI_{brewer} for 50 MSAs in 14 post-merger quarters. The variation of ΔHHI_{brewer} is large across markets ranging from 0.01 to 0.07 (or 100 to 700 if multiplied by 100^2). Given the HHI level of the whole country which is $0.49^2 + 0.18^2 + 0.11^2 = 0.28$, this merger increases brewers' HHI by more than 10% for most local markets. Figure 3 illustrates the re-

[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.

duction of shipping distance between 50 MSAs to MillerCoors breweries respectively. 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 reduces the shipping distance for Coors than Miller. The reason is that Coors only have two plants before the merger and Miller has six plants across the country. For Miller, only five markets have slight distance reduction after the merger.

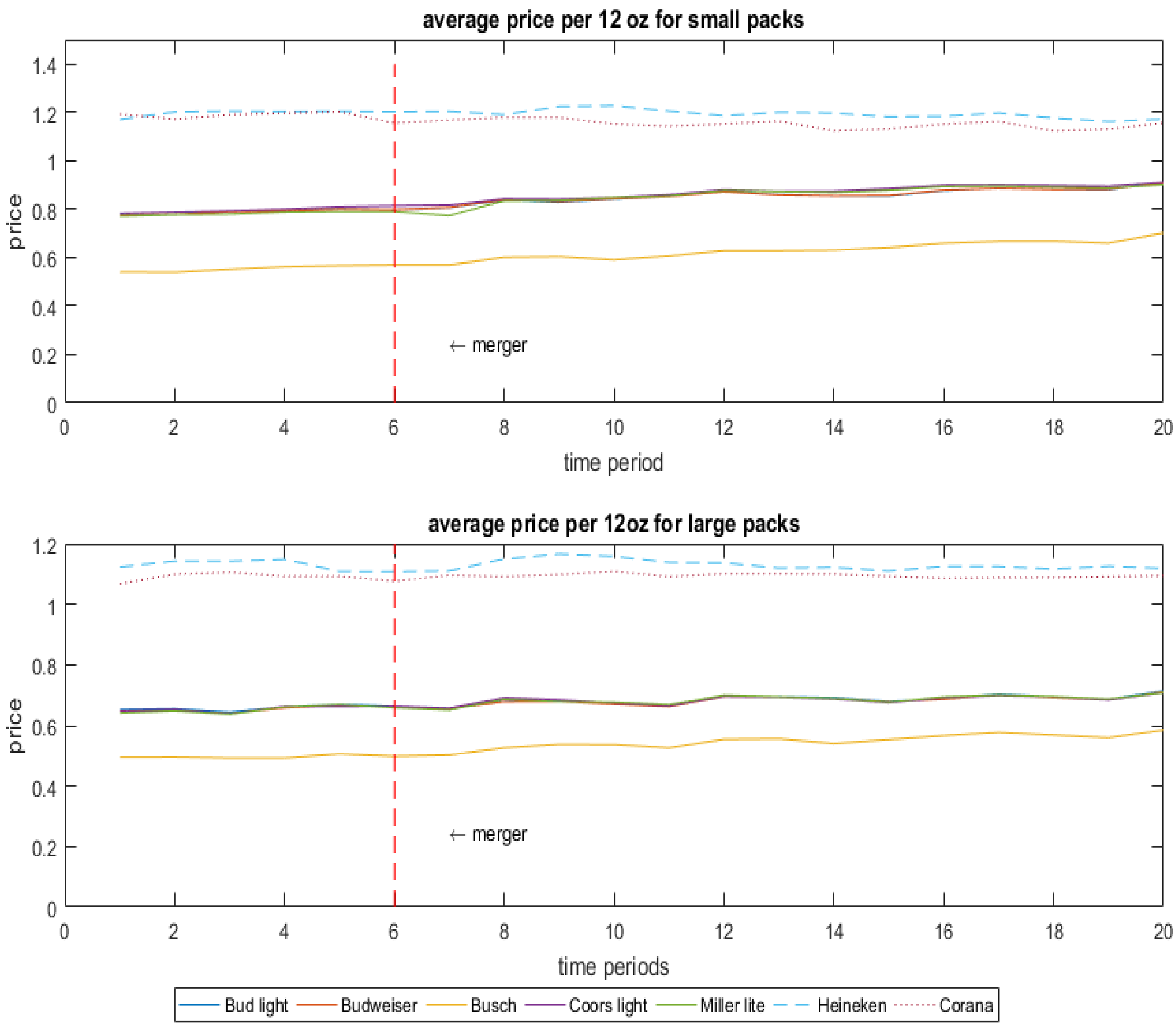


Figure 1: Average price per serving by brand over 50 markets

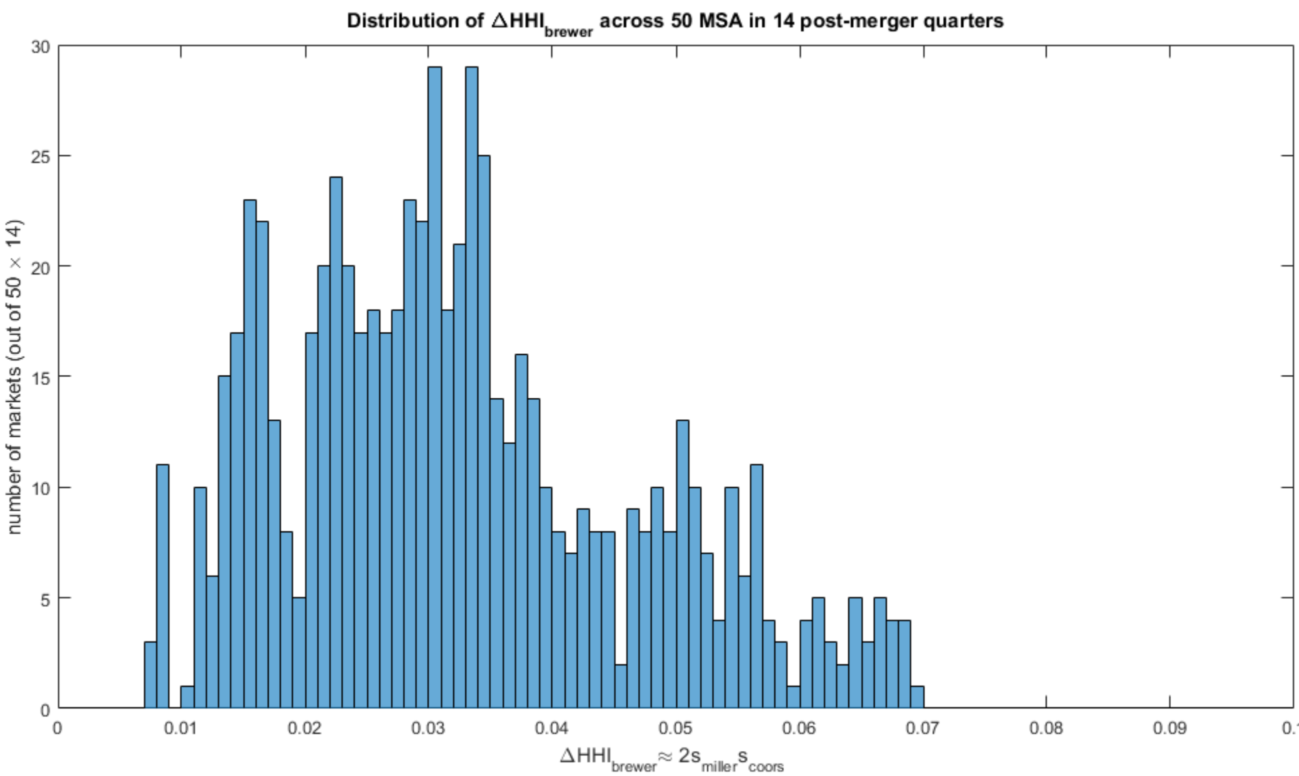
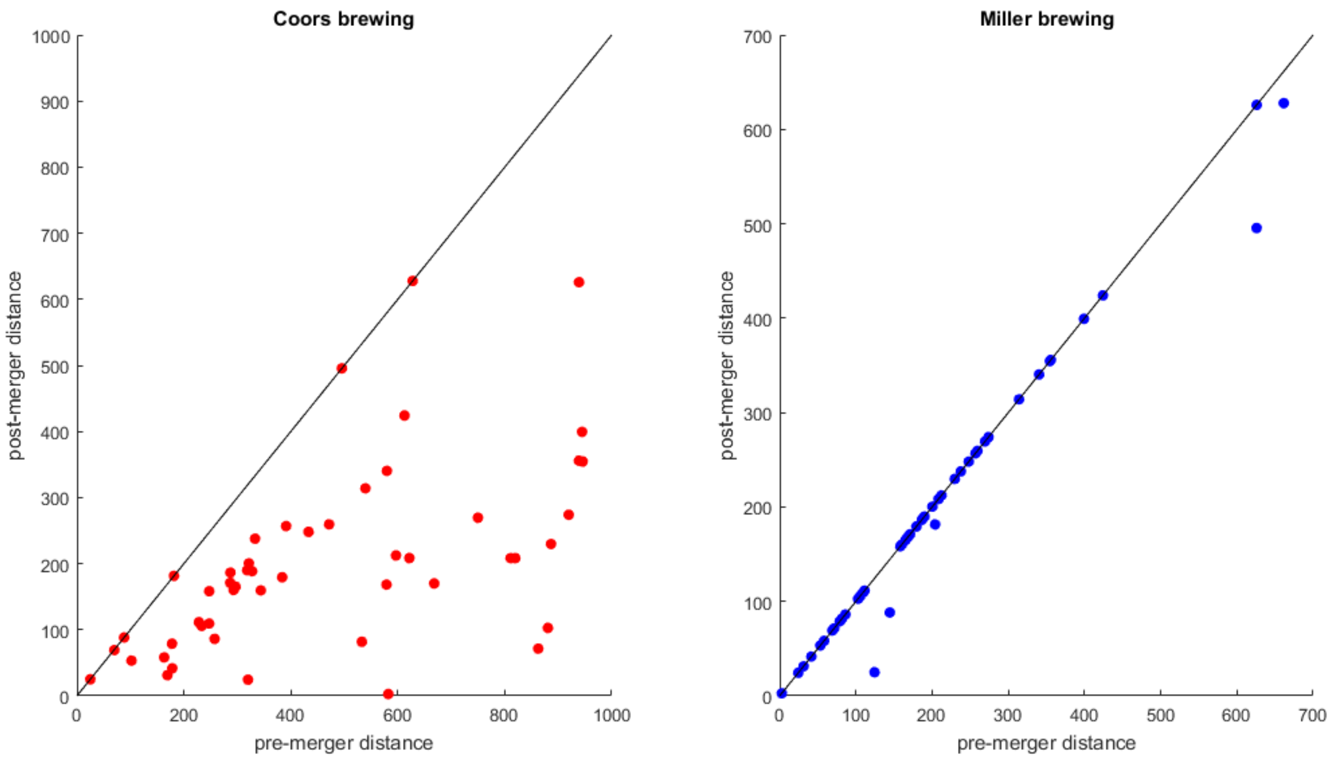


Figure 2: Distribution of HHI increases of post-merger



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

Figure 3: Distance(miles) between breweries and markets for pre/post-merger

The following logarithmic regressions are used to help investigate the impacts of merger and market concentrations on retail prices.

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

and the second specification is,

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

where the subscript j stands for brand j; c stands for chain c; m is geographic market MSA; 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. More importantly, the interaction term of two HHIs are added to estimate the counterveiling effect on retail price. In other words, the interaction term measures how downstream market concentration interferes with exercising market power of the upstream market. The coefficients are flexible by adding interaction terms of HHIs with the merger dummy. The merger dummy is not in the equation because its coefficient is not significant due to correlation with HHI and cost variables which already control for the effects of merger. 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). It differes from Ashenfelter, Hosken and Weinberg (2015) by adding interaction of $HHI^{retailer}$ and ΔHHI^{brewer} in order to study how downstream concentration affects the pass-through of increased market power to retail price. Since the market structure and merger are treated exogenously, the possible endogeneity is not addressed in the OLS regression.

The regression results are given in table 1. The first column is the regression result of equation 4.1 without the interaction terms of the merger dummy. The second column is the full specification of 4.1. The result shows that both HHI^{brewer} and $HHI^{retailer}$ positively affect retail prices. The coefficient on brewer HHI means increasing HHI by 0.01 points will increase the retail price per 12 oz by 0.463%. The estimate of HHI interaction term is negative for both column 1 and 2. It implies that increasing the upstream concentration will raise retail prices less in concentrated downstream markets than in markets with competitive downstream. In other words, competitive downstream markets have less power to mitigate the upstream shocks. The estimates $\hat{\beta}$ for interaction terms are significant and have the same signs. The estimated coefficient of large packs is negative since large packs often have discounts due to price discrimination or cost saving on packages. The shipping distance positively affects retail price such that 1% increase of shipping distance will increase retail price by 0.031%.

Column 3 of table 1 shows a regression result of 4.2. Like the first specification, increasing upstream concentration will positively affect retail price. However, the effect is mitigated by the downstream concentration according to the negative coefficient -1.232 of the interaction term. The preliminary regression only shows how HHIs and cost varaibles affect final retail price. The exact changes of explicit/implicit costs and markups and their pass-through on retail price are not quantified. The next section introduces the structural model used to estimate changes.

Table 1: 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

5 Structural Model

5.1 Beer Demand

The demand side models a consumer's decision for purchasing one 12 oz standard serving of beer using random coefficient discrete choice model following Berry (1994), Berry et al. (1995), and Nevo (2001). In each MSA-quarter, a consumer's choice set includes beer brands sold in all in-sample chains of the market and outside option such that each choice is a brand-size-chain combination. Stores of the same chain are assumed to set a uniform price for the same brand-size. This assumption reduces the number of price setting agents. The incentive of price differences across stores in the same chain is not focus of this paper which is studied by Dearing (2016)[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 by 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 I extreme-value distribution. This is the standard simple logit demand model.

[8]Dearing's paper finds that upstream adjustment will significantly dampen the profit earned of retailers if they set store-level prices rather than uniform price.

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

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

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

The mean utility of choosing outside option is:

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

which is normalized to be $u_{i0t} = \varepsilon_{i0t}$ for both the simple and mixed logit model.

Consumers with demographics $\{D_i, v_i, \boldsymbol{\varepsilon}_i\}$ choose one product which gives them the highest utility such that $jc^* = \operatorname{argmax}_{jc} u_{ijcmt}$. 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 TIEV 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 the simple logit model, and the market share is:

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

the δ_{jcmt} of simple logit model can be simply uncovered 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 can be calculated in the sample, the dependent variable of equation 5.9 can be constructed and regressed on the variables of the right-hand side.

Back to the full random coefficient model, BLP shows that the contraction mapping can help uncover mean utility δ from equation 5.7 by matching modelled predicted shares to observed market shares. To be more specific, given each pair of parameter values, one can randomly draw ns persons' (D_i, v_i) from "known" distribution of P_m such that the monte carlo simulated share is:

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

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, it is simple to regress the mean utility to estimate Θ_1. The common issue occurring to both the simple and mixed logit model is the endogeneity of p_{jcmt} which correlates with unobserved demand shock ξ_{jcmt}. This correlation is clearly shown in the first order conditions of sellers' optimal prices. In the estimation, I instrument prices with cost side variables such as local wage, rent, hop, and malt prices.

5.2 Beer Supply

The three-tier system of beer distribution includes beer manufacturers, distributors, and retailers. To decompose the price setting, it is ideal if all three tiers are modelled in the supply side. However, there are more than 3300 licensed beer wholesalers throughout 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. Without collecting data on beer distributors or even the number of distributors in a given MSA, I follow the literature (Hellerstein (2008), Goldberg and Hellerstein (2013)) by integrating manufacturers and distributors into one layer which jointly sets wholesale prices. Retailers set retail price in the second stage after observing wholesale prices . In the following sections, the distributor is ignored whenever I infer to brewer-distributor. 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 their marginal costs. The double markups are determined by price elasticity of demand and market structure. The former set requires demand estimates and the latter one is directly observed in data. Marginal costs are uncovered by subtracting markups from retail prices.

In order to calculate the double markups, the price setting is assumed to follow Bertrand-Nash linear pricing competition between upstream and downstream firms. Since vertical contracts between brewer and retailer are not observed, linear pricing contract is assumed as in Hellerstein (2008), Goldberg and Hellerstein (2013) and Dearing (2016). One evidence is that wholesale prices are posted publicly in some states and discrimination over retailers are not encouraged. It does not rule out the possibility of various types of contracts across states. Faheem and Gayle (2017) study the possible nonlinear pricing contract between brewers and retailers and whether it fits better to the data. Since my paper has a different focus, I simply assume linear pricing contract.

In first stage, brewers set wholesale prices for differentiated brands accounting for the downstream retailers' best response retail prices. The wholesale price is uniform across retailers without price discrimination. In the second stage, after observing the wholesale prices of all brands set by brewers, the retailers set optimal retail prices for the brands they carry. I do not model retailers' product portfolio choice which could be an interesting extension. Instead, the brands sold by each chain retailer are deemed as exogenously given and retailers only set endogenous retail prices. This assumption is reasonable since only selected "popular" brands are chosen in the sample. This two-stage game can be solved by backward induction.

5.2.1 Retailers

The retailer c representing a food chain in a market m at period t chooses retail price p_{jcmt} for all products it carries to maximize its profit[9]:

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

where J_{ct} is the set of products sold by chain c; p^{ω}_{jt} is wholesale price of brand j; mc^r_{jct} is marginal cost of chain c for selling product (j, c). To avoid the duplication of subscript, 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 does 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:

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

for all $(j, c) \in J_{ct}$. Expressing all products (j, c) of market t in matrix notation and inverse markups, the first order conditions are rewritten as:

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

with T_r as the retailer's ownership matrix. The dimension of matrix T_r equals the number of products in the market, denoted as 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.

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

5.2.2 Manufacturers

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

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

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 with respect to (w.r.t.) p^{ω}_{jt} becomes:

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

Similarly, let T_{ω} be a matrix of ownership for the manufacturers. The dimension of matrix T_{ω} is different from the dimension of T_r because the number of brands is less than the 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, denoted as 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 w.r.t. wholesale prices according to the chain rule:

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

Define a N^U_t-by-N_t matrix Δ_{pt} with element $(i,j) = \partial p_{jt}/\partial p^{\omega}_{it}$. This matrix reflects the pass-through rate of wholesale price to retail price. Once Δ_{pt} is calculated, derivatives of product market shares with respect to wholesale prices can be computed as $\Delta_{pt}\Delta_{rt}$ which is a N^U_t by N_t matrix. The matrix Δ_{pt} is derived from total differential of retailer's first order condition. To avoid too many subscripts, t and c are dropped in the following derivation. The total differential of first order condition 5.13 for product

[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.

j is:

$$\Sigma_{k=1}^{N}\underbrace{[\frac{\partial s_j}{\partial p_k}+\Sigma_{i=1}^{N}(T_r(i,j)\frac{\partial^2 s_i}{\partial p_j \partial p_k}(p_i-p_i^{\omega}-mc_i^r))+T_r(k,j)\frac{\partial s_k}{\partial p_j}]}_{\text{g(j,k)}}dp_k-$$

$$\underbrace{\Sigma_{f\in F}T_r(f,j)\frac{\partial s_f}{\partial p_j}}_{\text{h(j,F)}}dp_F^{\omega}=0 \quad (5.18)$$

where F is the set of products j of the same brand but sold in different chains. F can equivalently represent the brand out of N^U. Let G be the matrix with element $g(j,k)$ and let H_F be the N_t-dimensional vector with element $h(j,F)$. Equation 5.18 can be rewritten in matrix form $G\mathrm{d}\boldsymbol{p} - H_F\ \mathrm{d}p_F^{\omega} = 0$. Since there are N^U brands F, the matrix expression is $\Delta_p' = G^{-1} * [H_1 H_2 \dots H_{N^U}]$ with dimension N by N^U. Finally, 5.17 in matrix notation can be written as $\Delta_{\omega t} = \Delta_{pt}\Delta_{rt}U$, 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}. In analogy to retailer markup, the manufacturer markup is :

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

Note that the vector of brand market share in 5.19 has N_t^U elements different from the vector of product market share 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 w.r.t. 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 wholesale prices $\boldsymbol{p}_t^{\omega}$ are unobserved, it is impossible to calculate marginal costs respectively using 5.14 and 5.19. Instead, by combing these two equations, I can uncover the joint costs of retailer and manufacturer such that:

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

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 proxy for markups but not equivalent. To see that, even controlling for HHI, the markup can changes after cost savings. 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. After backing out marginal costs for both pre- and post-merger periods, cost savings through shipping distance or production can be estimated.

6 Estimation

6.1 Demand estimation

Demand estimation follows the standard procedure of estimating the random coefficient discrete choice model. For any given values of Θ_2, contraction mapping solves for fixed point of mean utility $\boldsymbol{\delta}(\Theta_2)$ such that model simulated market share equals observed share in data. Then, unobserved demand shock $\boldsymbol{\xi}(\Theta_1, \Theta_2)$ are uncovered by equation 5.2. 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:

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

where W is weight matrix. Following Nevo (2001), the number of parameters can be reduced by substituting the estimator $\hat{\Theta}_1$ for any guess of Θ_2 :

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

into the GMM estimation such that the estimation algorithm only searches over Θ_2 rather than $\{\Theta_1, \Theta_2\}$ to minimize the objective function. For the simple logit model, $\boldsymbol{\delta}$ can be calculated using market shares and the estimator for Θ_1 above is equivalent to 2SLS 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. Costs of manufacturers are also used 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 further use mean demographics interacting with exogenous product characteristics in X as instrument for estimating parameters Π in the random coefficients.

6.2 Supply estimation

Given demand estimates, the first and second order partial derivatives of market shares to retail prices are calculated. Based on 5.14 and 5.19, markups for retailers and manufacturers are calculated respectively. Then the joint marginal cost can be uncovered by subtracting double markups from retail price based on 5.20. Cost function is estimated in a linear regression model:

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

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, coefficients on these interaction terms are expected to be negative. One thing to note is that I do not add brand dummies interacting with merger dummy for other brewers. The marginal costs of other brewers could also change over time due to technology innovation or scale economy. Such common supply shocks are controlled by year and season dummies. The estimation results of cost function are used to predict costs of post-merger periods without merger in the counterfactual. This counterfactual helps to disentangle impacts of cost saving and increased market power on market equilibrium.

6.3 Results

6.3.1 Simple logit demand

Estimation results start with the simple logit model which regresses the mean utilities on product characteristics using OLS and 2SLS estimation. The regression results are listed in table 2. The price coefficients are negative which means the utility of choosing a product decreases with the price. Since price is endogenous and positively correlated with demand shock, OLS underestimates the price elasticity. The estimate of price coefficient with IV is larger than OLS as expected which means that the cost variables address the endogeneity issue of price. The second product characteristics is the logarithm of radius per store. This variable is calculated by dividing MSA area by the number of stores in the chain as proxy for shopping distance. Chains with more stores have a smaller radius per store which implies less traveling cost for consumers to visit. The estimated coefficient is negative which means that consumer utility decreases in traveling distance to the nearest store of the chain. The third coefficient is for the pack size dummy. I define package size less or equal to 12 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 when controlling for other characteristics consumers are less likely to purchase a large size. It makes sense considering the moving cost, storage cost, and beer's perishability. Finally, the estimates of selected brand fixed effects are listed. Imported bear has high fixed effects followed by domestic flagship brands such as Bud Light, Coors Light, and Miller Lite. Due to the limited substitution patterns and I.I.A problem of the simple logit demand model, random coefficient model is estimated in the next subsection.

6.3.2 Random Coefficient Model

The random coefficient model adds the interaction terms of consumer demographics and product characteristics according to 5.3. For each market, 300 consumers are randomly drawn from the joint distribution of income and age. The product characteristics interacting 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 among products with similar characteristics are stronger. Identification of coefficients

Table 2: Demand estimates of the 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

on these interaction terms comes from the different consumption patterns across markets with different demographics. For example, if consumers in market A have higher income than consumers in market B, we observe that 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 levels are less sensitive to price such that the coefficient of income interacting with price is positive.

The estimates of random coefficient model are shown in table 3. The same set of instrumental variables are used in the GMM estimation as in 2SLS regression. The first column of table 3 estimates Θ_1 in comparison to table 2. Estimates of Θ_2 are provided in the last three columns. Since brand dummies 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, 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. The interaction term of income and price has estimated coefficient 1.034 which means that consumers with higher incomes level are less price sensitive. The interactions of income with light, domestic dummies, and ABV are all negative indicating 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 provide more flexible substitution among products. The estimates of Θ_1 are similar to 2SLS with IV. The coefficient on radius of chain stores is -1.688 close to -1.684 in table 2. The estimated coefficient on size dummy is also negative. The coefficient on light dummy, domestic dummy, and ABV are retrieved from minimum-distance estimation. The bottom of table 3 reports the statistics of estimated own price elasticity. These statistics are a reference for the justification of estimated elasticity. The cutoff 0 is to check whether price elasticity is negative and the cutoff -1 comes from single-product optimal price elasticity $= -\dfrac{p}{p-c} < -1$.

The detailed own price elasticity and cross price elasticity by brand are provided in table 4. These numbers are obtained by aggregating market shares of a given brand over chains in a market, then calculating partial derivatives of brand shares to brand prices, and finally averaging the derivatives over markets. 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 4.924%. If the price of Coors Light increase by 1%, demand for Bud Light will increase by 0.757%. The elasticities are quite reasonable such that substitution among light beer such as Bud Light, Coors Light, and Miller Lite is stronger than substitution between light and lager. Moreover, substitution among domestic brands are stronger than imported brands. Finally, the second panel of table 4 lists almost all the 50 brands in my sample including large brewers and top craft brewers such as New Belgium, Yuengling, and Boston brewing.

Table 3: Demand estimates of the 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 4: 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 Bevada Pale Ale	-6.703		
Yuengling			
Yuengling	-5.060		
Boston Brewing			
Samuel Adams	-6.366		

6.3.3 Supply estimates

With demand estimates, the markups of retailer and manufacturer can be calculated and implicit marginal costs are uncovered 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 table 5. 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 on 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 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 of the merger from the common supply shock.

As for the effect of increased market power, the table compares the markups of retailers and brewers for Miller and Coors before and after the merger. The finding is that retailer's profit of selling one serving of Miller or Coors beer decreases by 2 cents on average. The brewer's markup of Coors increases from 15 cents to 19 cents and markup of Miller increases from 16 cents to 19 cents.

The conclusions are twofold. First, upstream brewers' profits increase with their market power. Second, downstream retailers may sacrifice their profits as a buffer against the positive shock of upstream on retail price. It is important to note that the decrease of retailer markup is a compound response to both changes of brewer's market power and marginal cost. For example, even without change of brewer's market power, retailers still adjust markups for change of marginal cost. Moreover, retailer's markups of selling Anheuser-Busch, Heineken, Modelo, Boston, and Sierra Nevada decrease. 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 5 provides the average quarterly profits of brewers. Miller and Coors profits increase by 4.3 million and 2.2 million after the merger. Anheuser-Busch profits also increase by 0.8 million which is mainly due to the shift of demand to Anheuser-Busch due to high MillerCoors new prices. Lastly, in the bottom of table 5, it shows 13% of uncovered marginal costs are negative. This is because the sum of calculated double markups exceeds the observed retail price. This could happen to multi-product firm even though each product has own price elasticity less than -1.

Table 5: 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						

To investigate how downstream market concentration impacts the pass-through of upstream shocks, the derivative of retail price to wholesale price, $\partial p/\partial p^{\omega}$, and Δ_{pt}, are calculated. This derivative measures the pass-through rate of the wholesale price change, either from cost saving or increased markup, on retail prices. If powerful downstream retailers dampen the upstream shock, it should have two effects. First, downstream concentration affects brewer markup as in 5.19. Second, it affects the pass-through rate. To study the relationship, two regression results are shown in table 6. The first column shows that concentrated retailers have smaller pass-through rates than competitive retailers such that when wholesale price increases/decreases, concentrated retailers will increase/decrease retail prices much less. In other words, retailers with power adjust their markups in the opposite direction (counteract) to the change of wholesale price more intensively than those with less power. This result is intuitive as competitive retailers have limited markups to adjust against the change of wholesale price. Moreover, the negative coefficient on upstream HHI implies that concentrated upstream markets negatively affect pass-through rate. This implies that high-concentrated upstream brewers charge high markups which increase the price elasticity of demand. Therefore, retailers are forced to squeeze their markups to mitigate the upstream shock resulting in a low pass-through rate. The second column shows how brewers' markups are affected by downstream concentration. The positive sign of coefficient on HHI-brewer indicates the standard relationship between market power and markup. Estimated coefficients on HHI-retailer is negative implying that brewers tend to charge slightly lower markups when the downstream market is concentrated as the double marginalization model predicts.

Table 6: Effects of retailer HHI on pass-through rate

	$\partial p/\partial p^{\omega}$	markup$^{\omega}$
HHI-retailer	-0.029	-0.017
	(0.004)**	(0.004)**
HHI-retailer × merger	0.016	0.016
	(0.001)**	(0.001)**
HHI-brewer	-0.031	0.171
	(0.005)**	(0.006)**
HHI-brewer × merger	-0.023	-0.040
	(0.002)**	(0.003)**
merge dummy	X	
Year dummies	X	
Season dummies	X	
Brand dummies	X	
Market dummies	X	
R-square	0.368	0.666
Observation	155,973	

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

Finally, given the uncovered 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 the cost per serving correlates 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, the 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 figure 3 that Coors benefits more than Miller in terms of shipping distance. The estimates of changes of brand dummies after merger are listed in table 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 post-merger production synergy. Most importantly, they cannot be estimated without backed out cost level and sample covering both pre- and post-merger. Without these estimates, 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 in order to disentangle merger effects and understand the role of vertical relationship in upstream merger.

7 Counterfactuals

With the estimates, several counterfactual scenarios are simulated by solving for a new equilibrium in each counterfactual. The first counterfactual solves market equilibrium without the joint venture. To do that, the marginal cost without merger cost saving is calculated by subtracting reduced shipping cost and synergy from the uncovered marginal cost. Then, new equilibrium markups and prices are calculated. This counterfactual is treated as a benchmark. In order to decompose the merger effects, the second counterfactual only allows cost saving of the merger by keeping the price-setting independent for Miller and Coors. In other words, Miller and Coors only optimize their own profit such that they do not exercise their market power by setting prices jointly. The equilibrium estimated with the sample is the third scenario with both cost saving and market power effects. In addition, two other counterfactuals are simulated when there is no vertical framework. The brewers set retail price such that there is only

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

one tier in the price setting. The purpose of two counterfactuals are twofold. First, it compares equilibrium with or without vertical relationship and to what extent they differ. Some literature on beer only models one tier which may draw incorrect welfare conclusion if supply side is improperly specified. Second, supply with one tier is equivalent to a perfectly competitive market in the downstream market. Therefore, the welfare changes of merger in one tier can be compared to two tiers to understand the buffer effect of downstream.

The simulation results are listed in table 8. Panel A shows the average number of each variable over products by brewer and downstream concentration[11]. Panel B shows the change of aggregate welfares. The columns without brackets represent the five scenarios. For example, the first column comes from equilibrium without merger. The second column is for the counterfactual with only cost saving. The fourth column comes from the estimates of data. The last two columns are counterfactuals with only one tier. The sampled MSA markets are categorized into “high” and “low” concentrated markets with cutoff $HHI_{chain} = 0.33$ in order to compare the differences in performance. Comparison between columns shows the decomposed effects of merger. Comparison between high and low reflects the heterogeneous responses to upstream merger for markets with different concentrations.

First, compare column (1) and (2) to see the cost saving effects on equilibrium. In the first block of panel A, the merger saves Coors the marginal cost of 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 downstream concentration and 1.1 cents in less concentrated ones. The differences in costs are mainly due to the location of markets. The second block shows changes of retailer markups. Due to the price adjustment, double markups change for both retailers and brewers even with only cost saving. It means that HHI is not as accurate as proxy for markup since markup can still change even without change of HHI. Since markups are adjusted after cost saving, cost saving does not fully pass through to retail price. The change of average markup is negligible. As supplement, column 3 shows the percentage of positive changes, $\Delta^{+}\% = (\text{col 2 - col 1})^{+}\%$, among products. For example, around 17.6%–30.7% of Coors products and 74.9%–80% of Miller product observations have higher retailer markup after the cost reduction. It implies that when cost decreases, some retailers will increases markup in the opposite direction and the

[11]For example, Coors(low) = 0.419 is obtained by averaging the marginal costs over all brands of Coors sold by all chains of the high concentrated MSAs in 14 post-merger quarters. In other words, the statistics is the average cost over a subset of 112 thousand “products”.

pass-through rate is less than one. However, retailers in competitive retail markets are less likely to increase markups (e.g., 17.6% vs. 30.7% of Coors) than concentrated ones due to the exercise of market power by retailers. The third block shows markup change of brewers. Similar to retailers, brewer's markup increases for most products which is 91.5%–93.5% for Coors and 50.8%–61.9% for Miller. By comparing "high" and "low" markets, it shows that brewer's markup is less likely to increase in market with more powerful downstream buyers (e.g., 50.8% vs. 61.9% of Miller). Another interesting finding is that Anheuser-Busch decreases its wholesale price to compete with low-cost products of Miller and Coors, while it is mitigated by the increase of retailer's markup ($+\Delta\% = 92\%$).

Comparing (2) and (4) demonstrates the changes due to increased market power conditional on cost saving. As market power increases, average retailer markups in low concentrated market increases by 0.1 cents for MillerCoors, while in high concentrated market it decreases by 0.1 cents for Coors and by 0.5 cents for Miller[12]. This results indicates that powerful retailers tend to reduce their markups to counteract the increase of upstream market power especially for Miller's products. Column 5 also displays this fact that markups of only 25.4% Miller products are increased by retailers in contrast to 64.6% in "low" markets. Surprisingly, retailers also reduce their markups of selling Anheuser-Busch products. For instance, the average markup of selling AB products in "high" markets decrease by 0.6 cents on average. As for brewers' markups, all three domestic brewers significantly increase markups after the consolidation which is 3.6–5.4 cents for Coors, 2.7–3.1 cents for Miller, and 1 cent for Anheuser-Busch.

The column 3 and 5 that represent the share of positive changes among products between different scenarios provide several crucial insights. First, percentages other than 0 and 100% indicate the heterogeneous responses of either retailers or brewers over products. In other words, the markup adjustment is not uniformly in one direction such that some products have increased markups while others have decreased markups. It may indicate that a theoretical model with both muti-product firms in upstream and downstream market of this kind may also end up with an ambiguous merger prediction. Second, the possible reason of driving the diverse responses among products is the complex substitution pattern among large number of products. Further work to investigate the composition of products with increased/decreased markups may provide

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

suggestions corresponding to divesture of this merger.

Panel B of table 8 shows the changes of brewers' and retailers' profits of selling MillerCoors brands in the first block and all brands in the second block. The second column shows that with cost saving the total profits of MillerCoors increases by 22 million dollars and retailers' profits increase by 73 million dollars compared to scenario 1. The fourth column shows profits' changes when MillerCoors maximizes joint profits to exercise its market power. MillerCoors' profit increases further, whereas retailers' profits decrease significantly. One reason is the decreased consumption of Miller and Coors products (by $3.27 * 10^8$ servings) due to higher prices and the other reason is that retailers sacrifice their markups to counter the upstream market power. At the bottom, it shows the changes of surplus and social welfare. Column (2) shows that total profits of all brewers decrease (by $1.81 * 10^7$) under cost saving of MillerCoors, although MillerCoors' profits increase (by $2.2 * 10^7$) which results from demand competition. The cost savings eventually benefit the consumers and welfare increases. After considering change of market power as in column (4), brewers' profits increase and consumption shifts from MillerCoors to other brands. But retailers' profits and consumer welfare decrease due to the buffer effect and anti-competitiveness. Considering the aggregate merger effect, consumers and retailers are worse off in this merger. However, their welfare loss 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. There are several important findings. 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. Third, similar to the second finding, the effect of increased market power on brewer markup is smaller with vertical relationship than without which indicates that concentrated downstream market restricts the brewers' exercising market power. Last, the social welfare after the merger in one tier framework is less than that in two tiers scenario, which implies that welfare may be underestimated in a horizontal merger analysis without accounting for reaction of downstream market.

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

Panel A: mean statistics							
	Two tiers					One tier	
Brewer name	No merger	w/CS	$+\Delta\%$	w/CS+Power	$+\Delta\%$	w/CS	w/CS+Power
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MC							
Coors(low)	0.419	0.391	(0%)			0.419	0.391
Coors(high)	0.055	0.026	(0%)			0.055	0.026
Miller(low)	0.289	0.276	(21.2%)			0.289	0.276
Miller(high)	-0.101	-0.112	(24.6%)			-0.101	-0.112
Markup_r							
Coors(low)	0.236	0.235	(17.6%)	0.235	(79.4%)		
Coors(high)	0.557	0.557	(30.7%)	0.556	(76.6%)		
Miller(low)	0.225	0.225	(74.9%)	0.226	(64.6%)		
Miller(high)	0.543	0.545	(80.0%)	0.540	(25.4%)		
AB(low)	0.235	0.236	(92.5%)	0.235	(2.3%)		
AB(high)	0.551	0.553	(92.6%)	0.547	(0.6%)		
Heineken(low)	0.261	0.262	(99.8%)	0.260	(0.3%)		
Heineken(high)	0.577	0.580	(99.9%)	0.574	(0%)		
Markup_w							
Coors(low)	0.157	0.160	(93.5%)	0.196	(100%)	0.176	0.214
Coors(high)	0.149	0.152	(91.5%)	0.206	(100%)	0.172	0.232
Miller(low)	0.160	0.159	(61.9%)	0.190	(100%)	0.175	0.209
Miller(high)	0.175	0.174	(50.8%)	0.201	(99.9%)	0.195	0.226
AB(low)	0.251	0.246	(0%)	0.255	(100%)	0.268	0.278
AB(high)	0.262	0.257	(1.9%)	0.268	(99.6%)	0.287	0.299
Heineken(low)	0.172	0.173	(95.9%)	0.172	(4.6%)	0.191	0.189
Heineken(high)	0.173	0.174	(98.5%)	0.173	(0.8%)	0.198	0.196
Panel B: welfare							
(MillerCoors only)		(2)-(1)		(4)-(2)			(7)-(4)
Brewer profits$		$2.2 * 10^7$		$4.5 * 10^7$			$5.6 * 10^7$
Retailer profits$		$7.3 * 10^7$		$-1.78 * 10^8$			
Cost saving$		$-3.53 * 10^7$		$2.34 * 10^7$			
Servings		$1.49 * 10^8$		$-3.27 * 10^8$			$-3.8 * 10^8$
(All brewers)		(2)-(1)		(4)-(2)			(7)-(4)
Brewer profits$		$-1.81 * 10^7$		$1.28 * 10^8$			$1.68 * 10^8$
Retailer profits$		$2.56 * 10^7$		$-5.40 * 10^7$			
Cost saving$		$-4.46 * 10^7$		$2.86 * 10^7$			
Servings		$3.99 * 10^7$		$-7.86 * 10^7$			$-5.52 * 10^7$
Consumer welfare$		$5.7 * 10^7$		$-1.17 * 10^8$			$-1.78 * 10^8$
Social welfare$		$6.45 * 10^7$		$-4.3 * 10^7$			$-1.00 * 10^7$

Note: "high" indicates $HHI_{chain} > 0.3$ where 0.3 is the median HHI over markets in post-merger

8 Conclusion and extension

This paper studies and quantifies the impacts of cost synergy and increased market power of upstream consolidation in the U.S. beer industry. Nielsen retail data of beer sales in food stores from 2007–2011 is used to estimate demand for beer in 50 selected MSA markets. I build and estimate a static demand model and a double markup supply model assuming a Bertrand0Nash linear pricing game between upstream brewers and downstream retailers. Implicit costs for both pre- and post-merger periods are backed out by subtracting estimated markups from retail price. By regressing 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 costs. The finding is 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, several counterfactual scenarios are simulated. I find that brewers will increase markup in both scenarios of cost savings and increased upstream market power. Retailers in markets with high concentrated downstream are more likely to adjust retail markups to dampen upstream shocks. The evidence for the latter is the low pass-through rate of concentrated retailers who are more likely to counteract when wholesale price decreases/increases. The diverse directions of markup adjustments among products reflects complex substitution patterns among products which could provide insights on divesture of the merger. Finally, the simulation finds that MillerCoors increases markups after the merger which dominates the cost savings. In terms of welfare, the mega-merger increases MillerCoors profits but hurts retailers' profits significantly. The total consumption of MillerCoors beer for all 50 markets from mid-2008 to 2011 decrease by $1.78*10^8$ servings. As for the change in 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. Most importantly, ignoring the vertical relationship in merger analysis would lead to inaacurate welfare conclusion.

As for extension, this paper only focuses on brewers' and retailers' price setting. One possible extension is to study how an upstream merger affects retailers' choice of brand portfolios. It is also interesting to study how this merger and price rise affect entry or profits of "small" brewers. For example, I find non-participating brewers' profits also increase after the merger as in table 5 due to the positive cross price elasticity and shift of demand. Another interesting aspect is to study the effects on advertising and introductoin of new products[13]. Chandra and Weinberg (2017) find positive effects of market concentration on advertising expenditure. If a merger increases marginal revenue of advertising, it may contribute to introducing new products by decreasing learning cost through intensive advertisement. Moreover, if positive spillover effect of advertising exists, a merger can further increase firms' incentive to introduce new products. Finally, as the counterfactual part shows, retailers adjust markups in different directions among the products. Some products have higher markups, while some have lower markups. This variation may suggest how to efficiently divest products in a merger which could be studied in future work.

[13]Usually, firms advertise new products for informative reason or to encourage learning.

References

[1] Ashenfelter, O, Hosken, D. and M. Weinberg (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 A. Pakes (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 (2018): "How does advertising depend on competition? Evidence from U.S. brewing", *Management Science*, 64(11), 4967-5460.

[7] Chen, Z. (2003):"Dominant Retailers and the Countervailing-Power Hypothesis", *RAND Journal of Economics*, 34, 612-625.

[8] 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.

[9] Dana, J.D. (2012): "Buyer groups as strategic commitments," *Games and Economic Behavior*, 74(2), 470-485.

[10] Dearing, A. (2016),"Pricing policy and welfare in retail chains: a case for vertical incentives," working paper.

[11] Dobson, P. and M. Waterson (1997), "Countervailing Power and Consumer Prices," *Economic Journal*, 107, 418-430.

[12] Ellison, S..F. and C.M. Snyder (2010), "Countervailing Power in Wholesal Pharmaceuticals," *Journal of Industrial Economics*, 58(1), 32-53.

[13] Faheem, A. and P. Gayle (2017): "Mergers, competitive conduct, and vertical relationships between firms: evidence from US beer industry," working paper.

[14] Fan, Y. (2013): "Ownership consolidation and product characteristics: a study of the US daily newspaper market," *American Economic Review*, 103, 1598-1628.

[15] Galbraith, J. K. (1952), "American Capitalism: The Concept of Countervailing Power". Boston: Houghton Mifflin..

[16] Goldberg, L. and R. Hellerstein (2013): "A structural approach to identifying the sources of local currency price stability," *Review of Economic Studies*, 80, 175-210.

[17] 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.

[18] Horn, H. and A. Wolinsky (1988):"Bilateral Monopolies and Incentives for Merger," *RAND Journal of Economics*, 19, 408-419.

[19] Houde, J. (2012): "Spatial differentiation and vertical mergers in retail markets for gasoline," *American Economic Review*, 102(5), 2147-2182.

[20] Katz, M.L. (1987):"The Welfare Effects of Third Degree Price Discrimination in Intermediate Goods Markets," *American Economic Review*, 77, 154-167.

[21] Lee, R. S. (2013): " Vertical integration and exclusivity in platform and two-sided markets," *American Economic Review*, 103(7), 2960-3000.

[22] Loertscher, S. and L.M. Marx(2019):"Merger Review for Markets with Buyer Power," *Journal of Political Economy*, 127, 2967-3017.

[23] 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.

[24] Melnik, G.A., Zwanziger, J., Barnezai, A. and R. Pattison (1992) ,"The Effects of Market Structure and Barganing Position on Hospital Prices", *Journal of Health Economics*, 11, 217-233.

[25] Miller, N. and M. Weinberg (2017): "Understanding the price effects of the Miller/Coors joint venture," *Econometrica*, 85, 1763-1791.

[26] Murry, C. (2017): "Advertising in vertical relationships: an equilibrium model of the automobile industry," working paper.

[27] 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.

[28] Nevo, A. (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69, 307-342.

[29] Nevo, A. and M. Whinston (2010): "Taking the dogma out of econometrics: structural modeling and credible inference," *Journal of Economic Perspectives*, 24(2), 69-82.

[30] 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.

[31] Sorensen, A.T. (2003), "Insurer-Hospital Bargaining: Negotiated Discounts in Post-Deregulation Connecticut", *Journal of Industrial Economics*, 51, 471-492.

[32] Sweeting, A., Tao, X. and X. Yao (2021): "Dynamic oligopoly pricing with asymmetric information: implications for horizontal mergers," working paper.

[33] Snyder, C.M. (1996):"A Dynamic Theory of Countervailing Power," *RAND Journal of Economics*, 27, 747-769.

[34] Tyagi, R. (2001):"Why Do Suppliers Charge Larger Buyers Lower Prices?" *Journal of Industrial Economics*, 49, 45-61.

[35] Villas-Boas, S. B. (2007): "Vertical relationships between manufacturers and retailers: inference with limited data," *Review of Economic Studies*, 74, 625-652.

[36] 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.

[37] von Ungern-Sternberg, T. (1996):" Countervailing Power Revisited," *International Journal of Industrial Organization*, 14, 45-61.

[38] Weinberg, M. (2011): "More evidence on the performance of merger simulations," *American Economic Review: Papers & Proceedings*, 101(3), 51-55.

[39] Weinberg, M. and D. Hosken (2013): "Evidence on the accuracy of merger simulations," *Review of Economics and Statistics*, 95(5), 1584-1600.

[40] Yang, C. (2020): "Vertical structure and innovation: a study of the SoC and smartphone industries," *RAND Journal of Economics*, 51(3), 739-785.

Appendix

Table 9: 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

continued from previous page					
Market	No. chains	No. products	Inside share	Market size(10^7 oz)	DMA food coverage
Medford	3	179	0.68	1.88	
Milwaukee	2	105	0.81	20.9	0.72
Myrtle Beach	3	199	0.90	7.74	
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 10: Couterfactuals Robustness: use 0.28 as cutoff for "high"/"low" markets

Panel A: mean statistics							
	Two tiers					One tier	
Brewer name	No merger	w/CS	+Δ%	w/CS+Power	+Δ%	w/CS	w/CS+Power
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
MC							
Coors(low)	0.440	0.412	(0%)			0.440	0.412
Coors(high)	0.158	0.130	(0%)			0.158	0.130
Miller(low)	0.318	0.305	(20.2%)			0.318	0.305
Miller(high)	0.003	-0.006	(24.3%)			0.003	-0.006
Markup_r							
Coors(low)	0.225	0.224	(18.6%)	0.225	(75.7%)		
Coors(high)	0.457	0.457	(25.1%)	0.456	(81.6%)		
Miller(low)	0.213	0.213	(73.8%)	0.214	(73.2%)		
Miller(high)	0.446	0.447	(79.2%)	0.444	(30.8%)		
AB(low)	0.224	0.225	(92.3%)	0.224	(2.8%)		
AB(high)	0.455	0.456	(92.6%)	0.452	(0.5%)		
Heineken(low)	0.250	0.251	(99.7%)	0.249	(0.3%)		
Heineken(high)	0.481	0.483	(99.9%)	0.478	(0%)		
Markup_w							
Coors(low)	0.159	0.162	(94.5%)	0.194	(100%)	0.177	0.212
Coors(high)	0.150	0.152	(91.5%)	0.203	(100%)	0.171	0.228
Miller(low)	0.156	0.156	(61.5%)	0.189	(100%)	0.170	0.206
Miller(high)	0.173	0.172	(54.7%)	0.199	(99.9%)	0.192	0.222
AB(low)	0.249	0.243	(0%)	0.253	(100%)	0.263	0.273
AB(high)	0.260	0.254	(1.3%)	0.266	(99.7%)	0.284	0.296
Heineken(low)	0.171	0.171	(94.9%)	0.171	(5.7%)	0.189	0.188
Heineken(high)	0.174	0.175	(98.8%)	0.173	(0.7%)	0.197	0.196

Panel B: welfare			
(MillerCoors only)	(2)-(1)	(4)-(2)	(7)-(4)
Brewer profits$	$2.2 * 10^7$	$4.5 * 10^7$	$5.6 * 10^7$
Retailer profits$	$7.3 * 10^7$	$-1.78 * 10^8$	
Cost saving$	$-3.53 * 10^7$	$2.34 * 10^7$	
Servings	$1.49 * 10^8$	$-3.27 * 10^8$	$-3.8 * 10^8$
(All brewers)	(2)-(1)	(4)-(2)	(7)-(4)
Brewer profits$	$-1.81 * 10^7$	$1.28 * 10^8$	$1.68 * 10^8$
Retailer profits$	$2.56 * 10^7$	$-5.40 * 10^7$	
Cost saving$	$-4.46 * 10^7$	$2.86 * 10^7$	
Servings	$3.99 * 10^7$	$-7.86 * 10^7$	$-5.52 * 10^7$
Consumer welfare$	$5.7 * 10^7$	$-1.17 * 10^8$	$-1.78 * 10^8$
Social 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