Bargaining and Merger in Vertical Relationships: Empirics of Packaged Food with Limited Data

Chen Gao

November 2021

Abstract

This paper estimates bargaining power in vertical relationships and simulates vertical mergers, and does so using limited data in the yogurt industry. Vertical mergers promote efficiency by eliminating double marginalization and lowering upstream rival wholesale prices, but harm welfare by increasing downstream rival costs and introducing upward pricing pressure on retail prices. To characterize vertical bargaining and simulate vertical integration in industries with limited data, I first develop a method to estimate vertical bargaining power between retailers and manufacturers, and then simulate vertical mergers of firms with various sizes. I use simulation results to demonstrate the relative magnitude of both pro- and anti-competitive incentives. The overall consumer welfare increases after merger, but consumers purchasing non-vertically integrated brands are worse off. [1]

[1]Researcher's own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1 Introduction

In 2020, the FTC and the DOJ published the new Vertical Merger Guideline. While the previous version in 1984 primarily focused on vertical mergers creating entry barriers, facilitating collusion, and allowing for evasion of rate regulation, the 2020 version added content discussing the efficiency losses due to vertically merged firms foreclosing rivals or increasing their costs and the efficiency gains due to the elimination of double marginalization. However, the policy towards vertical mergers continues to be controversial and there are still disagreements over which pro-/anti-competitive effect matters empirically. The FTC retracted the guideline in September 2021. Despite the rich theoretical literature on various pro- and anti-competitive effects, there's limited empirical evidence that jointly examines their relative magnitude and the overall impact on consumer welfare.

A main goal of my paper is to take a flexible approach to bargaining powers in vertical relationships, which becomes difficult when there is limited data. Vertical mergers become complicated in industries with vertical bargaining, and there are at least four channels through which vertical mergers affect market outcomes in these industries:

1. Vertical mergers eliminate double marginalization (EDM), but the extent to which double marginalization exists in the industry depends on vertical bargaining. If upstream firms provide products at marginal costs, then there's no double marginalization to eliminate.

2. Vertically integrated firms have additional bargaining leverage against both upstream and downstream rivals (Rogerson (2020)). Specifically, vertically merged firms have their own upstream and downstream departments. When they bargain with upstream or downstream rivals, their disagreement payoff is higher compared to cases before vertical merger. This gives them higher bargaining power against both upstream and downstream rivals even if the bargaining weight stays the same. On the one hand, higher bargaining power against upstream rivals allows the firm to get lower wholesale prices, creating downward pressure on its own retail prices; on the other hand, higher bargaining power against downstream rivals results in increased wholesale prices, creating upward pressure on rivals' retail prices.

3. Vertically integrated firms have incentives to increase rival costs (Salop and

Scheffman (1983)). For the vertically merged firm, supplying products to downstream rivals creates competition between these rivals and its own downstream department, hurting its overall profit. To lessen such competition, the merged firm can increase wholesale prices when it supplies products to downstream rivals. Rogerson (2020) highlights the theoretical difference between additional bargaining leverage and raising rival cost and, to my knowledge, my paper is the first to empirically investigate this distinction and to incorporate both effects into one model.

4. The existence of an upstream department creates an upward pricing pressure that works against the elimination of double marginalization (Moresi and Salop (2013)). Since different retailers compete in prices of yogurt with the same brand, if the vertically merged firm sells its products too cheaply in its own stores, they steal market shares and profit of the same brand yogurt in other retailers, indirectly hurting its own upstream department. As a result, vertically merged firms have incentives to increase store-branded product prices.

The theory of EDM and raising rival costs has existed for a long time, but the theory for vertical upward pricing pressure and bargaining leverage is relatively new and requires further empirical studies. By incorporating these effects into one empirical model, my paper takes a step forward in the literature of vertical mergers.

There's a small number of empirical studies that simulate vertical mergers under the environment of vertical bargaining (Crawford et al. (2018), Cuesta et al. (2019)), but they do not allow wholesale price bargaining to affect retail prices, ruling out the raising rival cost effect. They have also done so in data-rich settings where price data at multiple levels on the supply chain are available. When researchers have only retail market level data, estimation of vertical bargaining becomes difficult and researchers have to make restrictive assumptions on vertical relationships like assuming take-it-or-leave-it wholesale prices (Villas-Boas (2007)).

In my paper, I apply a bargaining model between retailers and manufacturers to a setting in which I have only retail-level data and thus show how to flexibly infer vertical relationships with limited data. I structurally estimate the bargaining power allocation between upstream and downstream firms and compute their equilibrium markups. I then simulate horizontal and vertical mergers to investigate the relative magnitude of various pro- and anti-competitive effects as well as the overall impact on consumer welfare.

The main data set I use is the retail scanner data on the yogurt market provided by Nielsen IQ. I focus on the yogurt industry in this paper for the following three reasons: first, yogurt is a non-durable good with static demand. By using yogurt data, I skip the rather difficult dynamic demand modeling problem. Second, yogurt has a high sales volume, a large number of brands, and wide availability in different stores. There is rich variation in brands and products across time, cities, and stores which enables the identification of bargaining power. Lastly, yogurt is a widely studied industry in the consumer demand literature so there are ample opportunities to compare the results of my model to other approaches.

Compared to previous studies in which researchers assume take-it-over-leave-it wholesale prices, I apply a Nash-in-Nash model for a more flexible vertical relationship. In the model, retailers and manufacturers engage in pairwise Nash bargaining over their net surplus of trade to set wholesale prices between them, and each retailer-manufacturer pair is governed by an exogenous bargaining weight parameter. Retailers carry yogurt from multiple manufacturers and manufacturers provide yogurt to multiple retailers. Such a network of supply chains gives both retailers and manufacturers outside options when they bargain with their rivals. As a result, the bargaining power of a retailer or a manufacturer depends on two factors: the bargaining weight parameter in Nash bargaining, and the strength of its outside options. For example, if the yogurt from brand A can be easily substituted by yogurt from brand B, the retailer won't have to worry about losing products from brand A, which gives the retailer more power when it bargains with manufacturer A. The strength of outside options can be computed using demand-side parameters, while the bargaining weight parameter needs to be estimated. In terms of data, unlike Crawford and Yurukoglu (2012) and Crawford et al. (2018), I have access to only retail market level data as in Villas-Boas (2007), so I develop an empirical method of estimating vertical bargaining power. Specifically, by solving the subgame-perfect equilibrium, I back out the sum of retailer and manufacturer marginal costs for any bargaining weight parameters. I then choose the parameters that best fit the input cost data to characterize vertical bargaining relationships in the industry.

Using a rich data set on the yogurt industry, I estimate the bargaining weight parameter of the retailer to be roughly 0.55. The average retailer markup is \$0.50, while the average manufacturer markup is \$0.20. These results suggest that in equilibrium, retailers have much higher markups than manufacturers. In vertical merger simulations, vertically merged firms have significantly higher upstream department markups and lower wholesale prices from upstream rivals. Consumer welfare increases slightly overall, but consumers who purchase non-vertically integrated brand yogurts are worse off.

This paper contributes to the vast literature on vertical mergers. Elimination of double marginalization is formalized in Spengler (1950), while incentives to increase rival cost and foreclosure is documented in Salop and Scheffman (1983), Krattenmaker and Salop (1986), and Hart and Tirole (1990). Moresi and Salop (2013) study the upward pricing pressure in vertical mergers, especially under vertical bargaining. Additionally, Rogerson (2020) points out that vertically merged firms have additional bargaining leverage over rivals. Crawford et al. (2018) is the closest to my paper in terms of vertical relationship and merger simulation, though my research differs in two ways: first, I assume that wholesale prices and retail prices are set sequentially, allowing for the raising rival costs effect; second, my method does not require upstream price and cost data which are difficult to access in most industries and is therefore applicable to a broader range of vertical merger studies.

This paper also adds to the literature on Nash-in-Nash bargaining both theoretically (Horn and Wolinsky (1988), Collard-Wexler et al. (2019)) and empirically (Crawford and Yurukoglu (2012), Ho and Lee (2017), Draganska et al. (2008)). Previous empirical studies applying the Nash-in-Nash bargaining model rely on data of retail price, wholesale price, and manufacturer marginal costs to identify vertical bargaining power, this paper takes the model a step further by reducing the data required to only retail prices. This greatly expands the set of industries in which the Nash-in-Nash bargaining model is applicable for empirical studies.

Lastly, my paper is related to vertical structure inference with limited data (Villas-Boas (2007), Hristakeva (2017)). Instead of assuming wholesale prices equal to manufacturer marginal cost or oligopoly prices, I allow them to be between these two extreme cases and estimate the bargaining power. This provides a more flexible framework of vertical relationships and wholesale price settings.

The rest of the paper is organized as follows: section 2 introduces the data, section 3 describes the model of vertical bargaining and vertical merger, section 4 presents the estimation and simulation method, section 5 and 6 discuss empirical results and counterfactual analysis, section 7 concludes the paper.

2 Data

2.1 Nielsen Scanner Data set

The yogurt retail market-level data come from Nielsen Retail Scanner data. The Nielsen data set contains store-week-UPC[2] level sales data. I observe units sold, price, product size and multiple (number of cups in a package), flavor, and other variables in 2016. During this year, Nielsen collected yogurt sales data in more than 35,000 stores within 69 retail chains. In terms of geographical information, Nielsen divides the United States into 207 designated market areas (DMAs). About 3,500 UPCs from 178 brands are observed in the data set. I aggregate the store level data into retail chain level for two reasons: first, the main players in my model are retail chain firms and manufacturers, not stores within the same retail chain. By aggregating them out, I directly model the competition between retail chain firms. Second, uniform pricing in retail chains is well documented in DellaVigna and Gentzkow (2019), and in the data the price variation across stores within the same retail chain is indeed quite small, so aggregating data to retail chain level maintains data variation pattern without creating unnecessary computational burden. After aggregation, the data is on DMA-week-retail chain-UPC level. I choose data from the seven largest DMAs in terms of yogurt consumption and drop brands with revenue share smaller than 1% to further reduce data size.

Table 1 reports the summary statistics of the best-selling brands. The 11 brands in the table are the brands with market share and revenue share larger than 1%. These 11 brands are all national brands and available in multiple retailers. All other brands' market shares are lower than 1%, so I consider them as unimportant and drop them. There are three major manufacturers with market shares significantly higher than all other brands.

Table 2 lists seven DMAs that consume the most yogurt. The average price per serving is similar across different DMAs, but there is a large variation in average

[2]UPC stands for universal product code, which can be understood as barcode

Brand	Servings	Ave. Price($)	Revenue($)	upcs	Mkt. Share	Rev. Share
Brand1	604M	1.18	715M	353	29.4%	28.3%
Brand2	357M	1.57	561M	184	17.4%	22.2%
Brand3	588M	0.90	531M	379	28.7%	21.0%
Brand4	123M	1.57	194M	73	6.0%	7.6%
Brand5	45M	2.12	97M	46	2.2%	3.8%
Brand6	66M	1.34	89M	97	3.3%	3.5%
Brand7	48M	1.14	56M	61	2.3%	2.2%
Brand8	40M	1.27	51M	40	2.0%	2.0%
Brand9	22M	2.18	48M	27	1.1%	1.9%
Brand10	33M	1.07	35M	50	1.6%	1.4%
Brand11	33M	0.77	26M	63	1.7%	1.0%

Brand names are concealed because Nielsen prohibits researchers to report brand or manufacturer names for research on topics including antitrust, collusion or illegal activities. A serving of yogurt is equal to 8 oz. Market share in this table is defined as servings divided by total servings sold in the data set with 7 DMAs. Revenue share is defined as revenue divided by total revenue in the data set with 7 DMAs.

Table 1: Best Selling Brands

yogurt consumption. This is partly caused by different store coverage rates in different DMAs. For example, Nielsen covers only 42% of all stores in New York yet up to 83% of all stores in Boston. To address this issue, I adjust the average yogurt consumption and divide it by the coverage rate. The last column of table 2 reports the adjusted average yogurt consumption, and the variation across DMAs is smaller after adjustment.

Table 3 lists the ten largest retail chain firms in the data set. Nielsen protects the identity of retailers from data users so I name them firm 1 to firm 10. Apart from one retailer (firm 7), all retailers provide hundreds of yogurt at different prices. Price variations across retailers are small but not negligible.

DMA	Servings	Ave. Price($)	Revenue($)	Population	Ave. Servings	Coverage	Adj. Ave. Servings
New York	142M	1.31	185M	21M	6.75	42%	16.07
LA	125M	1.25	157M	19M	6.38	52%	12.27
Boston	119M	1.21	144M	6.5M	18.09	83%	21.80
DC	79M	1.28	101M	6.8M	11.61	73%	15.93
SF	74M	1.32	98M	7.0M	10.54	48%	21.96
Chicago	76M	1.23	94M	9.6M	7.93	65%	12.2
Denver	68M	1.33	91M	4.6M	14.84	83%	17.88

Average servings consumed is equal to servings divided by population. Coverage is the percentage of stores included in the Nielsen data set. Adjusted average servings is equal to average servings divided by coverage

Table 2: DMAs with Most Yogurt Consumption

Retail Chain	Servings	Revenue	Average Price
Firm 1	316M	389M	1.23
Firm 2	201M	259M	1.28
Firm 3	176M	228M	1.29
Firm 4	79M	108M	1.37
Firm 5	67M	83M	1.23
Firm 6	71M	82M	1.14
Firm 7	81M	75M	0.93
Firm 8	60M	74M	1.23
Firm 9	65M	74M	1.13
Firm 10	54M	74M	1.35

Retailer names are not provided by NielsenIQ.

Table 3: Largest Retail Chain Firms

2.2 Input Cost Data

I collect a set of input cost data and their summary statistics are listed in table 4. Some provide weekly data and others provide monthly data. Cost components of both retailer and manufacturer are collected because both play an important role in the estimation.

	mean	median	std. dev	min	max	obs.
milk	14.91	14.82	1.56	12.76	17.40	12
sugar	18.24	19.12	2.96	12.52	23.42	53
gasoline	1.65	1.68	0.16	1.25	1.90	12
commercial electricity	6.47	6.47	0.16	6.20	6.73	12
industrial electricity	7.63	7.60	0.11	7.48	7.87	12
federal funds rate	0.39	0.38	0.06	0.20	0.66	53

Data source: milk(USDA), gasoline(Petroleum Market Monthly), electricity price(Form EIA-861M), sugar(sugar # 11 futures price), federal funds rate(Federal Reserve)

Table 4: Summary Statistics of Input Cost data

3 Model

In this section, I present a model in which manufacturers, retailers, and consumers play a three-stage game. In the first stage, each pair of retailer and manufacturer bargains over wholesale prices. In the second stage, retailers set retail prices to maximize their profit. In the third stage, consumers make purchase decisions, manufacturers and retailers earn their profit. Since this is a sequential game, I solve the subgame-perfect equilibrium by backward induction.

Vertically merged firms participate in the first two stages like their non-vertically merged rivals, but their bargaining power and retail price setting strategies are different. I derive their optimality conditions separately.

3.1 Consumer Decision

Consumers make discrete choices about which yogurt they buy. Consumer i gets utility u_{ijt} from buying yogurt j in market t:

$$u_{ijt} = x_{jt}\beta_i - \alpha_i p_{jt} + \xi_{jt} + \varepsilon_{ijt} = \delta_{ijt} + \varepsilon_{ijt}, \tag{1}$$

where x_{jt} is observable product characteristics, p_{jt} is the price of yogurt j in market t, β_i and α_i are coefficients of consumer i which are assumed to be heterogeneous across consumers, ξ_{jt} is yogurt j's unobservable product characteristics, $\delta_{ijt} = x_j\beta_i - \alpha_i p_{jt} + \xi_{jt}$ is consumer i's unitility without the error term, and ϵ_{ijt} is an error that follows type-I extreme value distribution.

I define a market as a DMA-week pair. The same yogurt product sold in different retailers are treated as different products as they are often sold at different prices. Treating them as different products allows for competition between retailers. For product characteristics, I use product multiple (number of cups in a package), size, flavor, brand fixed effect, retailer fixed effect, and week fixed effect. Apart from all the yogurts, each consumer has an outside option with utility $\delta_{i0t} = 0$. Consumers choose yogurt j with the highest u_{ijt}. Given such a choice rule, the probability of consumer i choosing j is

$$prob_{ijt} = \frac{e^{\delta_{ijt}}}{1 + \sum_k e^{\delta_{ikt}}},$$

and the market share of yogurt j in market t is

$$s_{jt} = \mathbb{E}_i \left[prob_{ijt}\right] = \mathbb{E}_i \left[\frac{e^{\delta_{ijt}}}{1 + \sum_k e^{\delta_{ikt}}}\right], \tag{2}$$

where the expectation is taken with respect to β_i and α_i. Here I allow constant coefficient β_c and price coefficient α to be different across consumers and assume they follow a normal distribution with standard error σ_c and σ_p.

By assumption $\beta_i^c \sim N(\beta^c, \sigma_c^2)$, $\alpha_i \sim N(\alpha, \sigma_p^2)$, I rewrite $\beta_i^c = \beta^c + \sigma_c \nu_i^c$ and $\alpha_i = \alpha + \sigma_p \nu_i^p$ where ν_i^c and ν_i^p are both standard normal random variables. For simplicity I assume ν_i^c and ν_i^p are uncorrelated, though in real life consumers' preference for yogurt could be correlated with their yogurt price sensitivity. Given this change of expression, I rewrite δ_{ijt} as

$$\begin{aligned}
\delta_{ijt} &= x_{jt}\beta + \sigma_c \nu_i^c - (\alpha + \sigma \nu_i^p) p_{jt} + \xi_{jt} \\
&= x_{jt}\beta - \alpha p_{jt} + \xi_{jt} + \sigma_c \nu_i^c - \sigma_p \nu_i^p p_{jt} \\
&= \delta_{jt} + \sigma_c \nu_i^c - \sigma_p \nu_i^p p_{jt},
\end{aligned}$$

where $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ is the mean utility of product j in market t. Plug this into equation (2):

$$s_{jt} = \int_i \frac{e^{\delta_{jt} + \sigma_c \nu_i^c - \sigma \nu_i^p p_{jt}}}{1 + \sum_k e^{\delta_{kt} + \sigma_c \nu_i^c - \sigma \nu_i^p p_{kt}}} df(\nu_i^c, \nu_i^p). \tag{3}$$

3.2 Retail price setting

Given the consumers' decision, retailer r sets prices to maximize its total profit in market t:

$$\max_{\{p_{jt}\}_{j \in S_t^r}} \Pi_t^r(\mathbf{p}_t, \mathbf{wp}_t, \mathbf{mc}_t^r) = \sum_{j \in S_t^r} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t), \tag{4}$$

where S_t^r is the set of all products provided by retailer r in market t, p_{jt} is the retail price of yogurt j in market t, wp_{jt} is the wholesales price of yogurt j in market t, mc_{kt}^r is the retailer's marginal cost of selling a unit of yogurt j in market t(which includes transportation, storage, wages, etc.), M is the potential market size of market t, and $s_{jt}(\mathbf{p}_t)$ is the market share of yogurt j in market t determined by equation (3). The vectors of all retail prices, wholesale prices and retailer marginal costs are denoted $\mathbf{p}_t$, $\mathbf{wp}_t$, and $\mathbf{mc}_t^r$, respectively.

The first-order condition with respect to p_{jt} is:

$$\frac{\partial \Pi_t^r}{\partial p_{jt}} = s_{jt}(\mathbf{p}_t) + \sum_{k \in S_t^r} (p_{kt} - wp_{kt} - mc_{kt}^r) \frac{\partial s_{kt}(\mathbf{p}_t)}{\partial p_{jt}} = 0. \tag{5}$$

Define Δ_t^p as a N-by-N matrix(N being the total number of products in market t) of partial derivatives of market shares with respect to retail prices such that the (i,j)-th element is $\partial s_{jt}/\partial p_{it}$ if i, j are sold by the same retailer and 0 otherwise, then it can be easily verified that the retailer first order condition in matrix form is

$$\mathbf{p}_t - \mathbf{wp}_t - \mathbf{mc}_t^r = -(\Delta_t^p)^{-1}\mathbf{s}_t(\mathbf{p}_t), \tag{6}$$

where $\mathbf{s}_t(\mathbf{p}_t)$ is the vector of market shares in market t as a function of $\mathbf{p}_t$. Rewrite equation (6) as:

$$\mathbf{p}_t + (\Delta_t^p)^{-1}\mathbf{s}_t(\mathbf{p}_t) = \mathbf{wp}_t + \mathbf{mc}_t^r, \tag{7}$$

Equation (7) gives us the optimal retail prices as a function of wholesale prices and retailer marginal costs: $\mathbf{p}_t = \mathbf{p}_t(\mathbf{wp}_t, \mathbf{mc}_t^r)$, which are taken as given by retailers at this stage.

3.3 Wholesale price setting

Before discussing wholesale price bargaining, it is necessary to define retailer and manufacturer profit. Retailer r's profit is defined by equation (4). Since retailers set retail prices as a function of wholesale prices, I rewrite retailer profit as

$$\Pi_t^r(\mathbf{p}_t(\mathbf{wp}_t), \mathbf{wp}_t, \mathbf{mc}_t^r) = \Pi_t^r(\mathbf{wp}_t).$$

I omit retailer's marginal costs term in this equation since they are exogenous to shocks in yogurt industry. Manufacturer w's profit is:

$$\Pi_t^w(\mathbf{wp}_t) = \sum_{j \in \tilde{S}_t^w} (wp_{jt} - mc_{jt}^w) M s_{jt}(\mathbf{p}_t(\mathbf{wp}_t, \mathbf{mc}_t^r)).$$

where $\tilde{S}_t^w$ is the set of products produced by w (the Tilde above S is added only for manufacturers to distinguish from retailer product sets), and mc_{jt}^w is the manufacturer's marginal cost of yogurt j in market t. Retail prices are determined by equation (7) as a function of wholesale prices and retailer marginal costs, manufacturer marginal costs are determined by yogurt characteristics. Both retailer and manufacturer marginal costs are taken as given by manufacturers, so I write manufacturer's profit as a function that depends only on $\mathbf{wp}_t$. It's worth highlighting that both retailers and manufacturers consider retail prices as a function of wholesale prices, this is different from Crawford et al. (2018) in which retailer and manufacturers set retail prices and wholesale prices simultaneously. Such sequential price setting scheme introduces raising rival costs effect into vertical mergers, which I will discuss in later chapters.

Each pair of manufacturer and retailer bargains over the wholesale price simultaneously, while the market equilibrium is a Nash equilibrium in Nash bargaining(Nash-in-Nash). Wholesale price maximizes the generalized Nash product given the wholesale prices of all other pairs of retailers and manufacturers:

$$\begin{aligned}\mathbf{wp}_t^{rw} =&\arg\max_{\mathbf{wp}_t^{rw}}(\Pi_t^r(\mathbf{wp}_t^{rw},\mathbf{wp}_t^{-rw}) - \Pi_t^r(\infty,\mathbf{wp}_t^{-rw}))^{\zeta}\\ &(\Pi_t^w(\mathbf{wp}_t^{rw},\mathbf{wp}_t^{-rw}) - \Pi_t^w(\infty,\mathbf{wp}_t^{-rw}))^{1-\zeta},\end{aligned}$$

where $\mathbf{wp}_t^{rw}$ is the vector of all wholesale prices between retailer r and manufacturer w in market t, $\mathbf{wp}_t^{-rw}$ is the vector of wholesale prices between all other retailer-manufacturer pairs in market t, ζ is the bargaining weight of retailer r. I rewrite $\Pi_t^r(\mathbf{wp}_t)$, the profit of retailer r, as $\Pi_t^r(\mathbf{wp}_t^{rw},\mathbf{wp}_t^{-rw})$ to separate $\mathbf{wp}_t^{rw}$ from $\mathbf{wp}_t^{-rw}$. This is because retailer r bargains over $\mathbf{wp}_t^{rw}$ while treating $\mathbf{wp}_t^{-rw}$ as given. The disagreement payoff of retailer r when it fails to reach a deal with manufacturer w is $\Pi_t^r(\infty,\mathbf{wp}_t^{-rw})$ in which wholesale prices between retailer r and manufacturer w are set to infinity while all other wholesale prices remain the same. The difference term, $\Pi_t^r(\mathbf{wp}_t^{rw},\mathbf{wp}_t^{-rw}) - \Pi_t^r(\infty,\mathbf{wp}_t^{-rw})$, is retailer r's surplus from coming to an agreement. Such expression also applies to manufacturer w. To simplify computation, I assume all pairs of retailers and manufacturers have the same bargaining weight parameter ζ, though this assumption can be easily generalized such that different pairs of manufacturers and retailers have different bargaining weights.

For retailers, the net surplus of bargaining with manufacturer w is smaller than the sales profit of yogurt w. If a retailer fails to achieve an agreement with a manufacturer w, this retailer would adjust the retail price of all other brands to compensate for the loss of not being able to sell the w-branded yogurt, so the net surplus of achieving an agreement is smaller than the incremental sales profit of selling w. Such net gain from bargaining satisfies

$$
\begin{aligned}
&\Pi_t^r(\mathbf{wp}_t^{rw}, \mathbf{wp}_t^{-rw}) - \Pi_t^r(\infty, \mathbf{wp}_t^{-rw}) \\
&= \underbrace{\sum_{j \in \tilde{S}_t^w, j \in S_t^r} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t) + \sum_{j \notin \tilde{S}_t^w, j \in S_t^r} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t)}_{\text{Profit when a deal is reached, } \Pi_t^r(\mathbf{wp}_t^{rw}, \mathbf{wp}_t^{-rw})} \\
&- \underbrace{\left[\sum_{j \notin \tilde{S}_t^w, j \in S_t^r} (p_{jt}^* - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t^*) \right]}_{\text{Profit when a deal is not reached, } \Pi_t^r(\infty, \mathbf{wp}_t^{-rw})} \\
&= \underbrace{\sum_{j \in \tilde{S}_t^w, j \in S_t^r} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t)}_{\text{Profit of selling yogurt from } w} \\
&- \underbrace{\left[\sum_{j \notin \tilde{S}_t^w, j \in S_t^r} (p_{jt}^* - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t^*) - \sum_{j \notin \tilde{S}_t^w, j \in S_t^r} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t) \right]}_{\text{Profit compensation of non-} w \text{ yogurts}} \\
&= \sum_{j \in \tilde{S}_t^w, j \in S_t^r} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t) - \Delta\pi_{rwt}^r M
\end{aligned}
\tag{8}
$$

where p_{jt}^* is the optimal retail price of yogurt j in the counterfactual case where retailer r and manufacturer w fails to reach a deal, and $\mathbf{p}_t^*$ is the vector of the market equilibrium retail prices when retailer r and manufacturer w fails to reach a deal. There is no analytical solution for p_{jt}^*, so I adapted a contraction mapping-based numerical method to solve for it (see appendix A).

Let

$$
\Delta\pi_{rwt}^r = \sum_{j \notin \tilde{S}_t^w, j \in S_t^r} (p_{jt}^* - wp_{jt} - mc_{jt}^r) s_{jt}(\mathbf{p}_t^*) - \sum_{j \notin \tilde{S}_t^w, j \in S_t^r} (p_{jt} - wp_{jt} - mc_{jt}^r) s_{jt}(\mathbf{p}_t)
$$

be the (per consumer) profit compensation of retailer r not able to sell yogurt produced by w in market t, then $\Delta\pi_{rwt}^r$ affects the result of bargaining by directly entering the net gain of bargaining. This term is constructed under the assumption that wholesale prices of yogurt from other brands remain the same whether or not r and w reach a deal, which is true in the Nash-in-Nash Model. The profit compensation term $\Delta\pi^r_{rwt}$ represents the strength of r's outside options when bargaining with w. If yogurt from w is easily substituted by yogurt from another brand, then retailer r will be able to recover most of the loss caused by losing yogurt from w. This puts retailer r in an advantageous position and allows it to move the wholesale prices towards the direction that benefits them. Conversely, if yogurt from w can't be replaced by other yogurts, then retailer r is in a disadvantageous position while bargaining with w, forcing them to make a concession and accept less favorable wholesale prices. I will discuss how $\Delta\pi^r_{rwt}$ affects bargaining results while deriving the first-order conditions for wholesale prices.

Similarly, for manufacturers,

$$
\begin{aligned}
&\Pi^w_t(\mathbf{wp}^{rw}_t, \mathbf{wp}^{-rw}_t) - \Pi^w_t(\infty, \mathbf{wp}^{-rw}_t) \\
&= \underbrace{\sum_{j\in S^r_t, j\in \tilde{S}^w_t} (wp_{jt} - mc^w_{jt}) M s_{jt}(\mathbf{p}_t)}_{\text{Profit of selling to retailer } r} \\
&- \underbrace{\left[\sum_{j\notin S^r_t, j\in \tilde{S}^w_t} (wp_{jt} - mc^w_{jt}) M s_{jt}(\mathbf{p}^*_t) - \sum_{j\notin S^r_t, j\in \tilde{S}^w_t} (wp_{jt} - mc^w_{jt}) M s_{jt}(\mathbf{p}_t)\right]}_{\text{Profit compensation of non-}r\text{ retailers}} \\
&= \sum_{j\in S^r_t, j\in \tilde{S}^w_t} (wp_{jt} - mc^w_{jt}) M s_{jt}(\mathbf{p}_t) - \Delta\pi^w_{rwt} M,
\end{aligned}
\tag{9}
$$

where mc^w_{jt} is the manufacturer w's marginal cost of yogurt j in market t,

$$
\Delta\pi^w_{rwt} = \sum_{j\notin S^r_t, j\in \tilde{S}^w_t} (wp_{jt} - mc^w_{jt}) s_{jt}(\mathbf{p}^*_t) - \sum_{j\notin S^r_t, j\in \tilde{S}^w_t} (wp_{jt} - mc^w_{jt}) s_{jt}(\mathbf{p}_t)
$$

is the (per consumer) profit compensation of manufacturer w when it fails to reach a deal with retailer r. The wholesale prices between manufacturer w and non-r retailers is still wp_{jt} when w and r fails to reach a deal, but the market share s_{jt} changes due to a new set of optimal retail prices. According to Nash-in-Nash bargaining model, the result of one pair of bargaining doesn't affect the result of other pairs, so using the same wholesale price is valid. Similar to $\Delta\pi^r_{rwt}$, $\Delta\pi^w_{rwt}$ represents the strength of manufacturer w's outside options when bargaining with retailer r. A large $\Delta\pi^w_{rwt}$ means w has strong outside options against r, and the bargaining result biases toward the direction that favors w.

Rewrite the bargaining condition:

$$wp_{jt} = \arg\max_{wp_{jt}} \left[\sum_{j\in \tilde{S}_t^w, j\in S_t^r} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t) - \Delta\pi_{rwt}^r M \right]^{\zeta_{rw}} \cdot \left[\sum_{j\in S_t^r, j\in \tilde{S}_t^w} (wp_{jt} - mc_{jt}^w) M s_{jt}(\mathbf{p}_t) - \Delta\pi_{rwt}^w M \right]^{1-\zeta_{rw}}$$

For simplicity, I take the log of this equation before deriving first order conditions and the result is:

$$\begin{aligned} &\zeta \frac{-s_{jt}(\mathbf{p}_t) + \sum_{k\in S_t^r} \frac{\partial p_{kt}}{\partial wp_{jt}} s_{kt} + \sum_{k\in S_t^r} \frac{\partial s_{kt}}{\partial wp_{jt}} (p_{kt} - wp_{kt} - mc_{kt}^r)}{\sum_{k\in S_t^r, k\in \tilde{S}_t^w} (p_{kt} - wp_{kt} - mc_{kt}^r) s_{kt}(\mathbf{p}_t) - \Delta\pi_{rwt}^r} \\ &+(1-\zeta) \frac{s_{jt}(\mathbf{p}_t) + \sum_{k\in \tilde{S}_t^w} (wp_{kt} - mc_{kt}^w) \frac{\partial s_{kt}(\mathbf{p}_t)}{\partial wp_{jt}}}{\sum_{k\in \tilde{S}_t^w, k\in S_t^r} (wp_{kt} - mc_{kt}^w) s_{kt}(\mathbf{p}_t) - \Delta\pi_{rwt}^w} = 0 \end{aligned} \tag{10}$$

subject to constraints:

$$\sum_{k\in S_t^r, k\in \tilde{S}_t^w} (p_{kt} - wp_{kt} - mc_{kt}^r) s_{kt}(\mathbf{p}_t) - \Delta\pi_{rwt}^r \geq 0 \tag{11}$$

$$\sum_{k\in \tilde{S}_t^w, k\in S_t^r} (wp_{kt} - mc_{kt}^w) s_{kt}(\mathbf{p}_t) - \Delta\pi_{rwt}^w \geq 0 \tag{12}$$

Equation (10) shows how ζ, $\Delta\pi_{rwt}^r$ and $\Delta\pi_{rwt}^w$ affect the result of bargaining. A larger ζ increases the weight of retailer's optimality condition,

$$-s_{jt} + \sum_{k\in S_t^r} \frac{\partial p_{kt}}{\partial wp_{jt}} s_{kt} + \sum_{k\in S_t^r} \frac{\partial s_{kt}}{\partial wp_{jt}} (p_{kt} - wp_{kt} - mc_{kt}^r)$$

in equation (10), leading to more favorable wholesale prices for the retailer. A larger $\Delta\pi_{rwt}^r$ decreases the denominator of the retailer's optimality condition, resulting in a higher weight of retailer's optimality condition and thus more favorable wholesale prices for the retailer. Similarly, lower ζ or higher $\Delta\pi_{rwt}^w$ increases the manufacturer's weight in equation (10) and leads to wholesale prices that benefits the manufacturer.

Equation (10) holds only when constraints (11) and (12) hold because both retailer and manufacturer have to have positive net surplus from reaching a deal. These two constraints set upper and lower bounds for wholesale prices regardless of

ζ specifications. They are unbinding when ζ is far from zero and one. When at least of them is binding under extreme values of ζ (for example zero or one), equation (10) does not hold and cannot be used to compute equilibrium wholesale prices. I will introduce a case where they are binding in the discussion of vertical mergers.

In matrix form, equation (10) can be written as[3]

$$\mathbf{wp}_t - \mathbf{mc}_t^w = T_l^{-1} V, \tag{13}$$

where $\mathbf{mc}_t^w$ is the vector of all manufacturer marginal costs in market t.

3.4 Retail Price Setting and Wholesale Price Bargaining of Vertically Merged Firm

Like their non vertically merged rivals, vertically merged firms optimize retail prices and bargain over wholesale prices as well. Unlike their rivals, however, their profit now consists of an upstream department and downstream department, so their optimality condition differs from firms with just an upstream or downstream department. In this section, I derive these conditions and exploit various incentives that improve/hurt efficiency and competition. For simplicity, I assume there is only one vertically merged firm in the industry, though the mechanism in this section can be extrapolated to markets with multiple vertically merged firms.

3.4.1 Retail Price Setting, Elimination of Double Marginalization, and Vertical Upward Pricing Pressure (VUPP)

In this section, I demonstrate how vertically merged firms have incentives to increase their retail prices to avoid hurting their own manufacturer department. The vertically integrated firm sets retail prices after bargaining, so I treat wholesale prices as given in this section and discuss wholesale price bargaining in later sections.

Suppose a retailer and a manufacturer merge into firm A, then yogurts manufactured and sold by A have their double marginalization eliminated. For simplicity, I call them vertically integrated yogurt, or VI yogurt. The retailer department of A keeps buying yogurt from other brands, and the manufacturer department of A

[3] See appendix B for derivation

keeps supplying yogurt to other retailers. A's total profit in market t is equal to the sum of upstream profit, VI yogurt profit, and downstream profit from yogurt with other brands:

$$
\begin{aligned}
\Pi_t^A = & \underbrace{\sum_{j \in S_t^{Ad}, j \notin S_t^{Au}} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t)}_{\text{retailer profit of yogurt from other brands}} \\
& + \underbrace{\sum_{j \in S_t^{Ad}, j \in S_t^{Au}} (p_{jt} - mc_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t)}_{\text{profit of VI yogurt}} \\
& + \underbrace{\sum_{j \in S_t^{Au}, j \notin S_t^{Ad}} (wp_{jt} - mc_{jt}^w) M s_{jt}(\mathbf{p}_t)}_{\text{manufacturer profit of brand A yogurt in other retailers}}
\end{aligned}
$$

where S_t^{Ad} is the set of products sold by downstream (retailer) department of A in market t, S_t^{Au} is the set of products produced by upstream (manufacturer) department of A in market t. VI yogurt can be viewed as ones supplied to retailer A at manufacturer A's marginal cost, namely $wp_{jt} = mc_{jt}^w$. The total profit of firm A can then be expressed by the sum of retailer profit and manufacturer profit:

$$
\begin{aligned}
\Pi_t^A = & \underbrace{\sum_{j \in S_t^{Ad}} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t)}_{\text{retailer profit}} \\
& + \underbrace{\sum_{k \in S_t^{Au}, k \notin S_t^{Ad}} (wp_{jt} - mc_{jt}^w) M s_{jt}(\mathbf{p}_t)}_{\text{manufacturer profit}}
\end{aligned}
$$

The first order condition of A's retail prices is as follows:

$$
\begin{aligned}
& s_{jt}(\mathbf{p}_t) + \sum_{k \in S_t^{Ad}} (p_{kt} - wp_{kt} - mc_{kt}^r) \frac{\partial s_{kt}(\mathbf{p}_t)}{\partial p_{jt}} \\
& + \sum_{k \in S_t^{Au}, k \notin S_t^{Ad}} (wp_{kt} - mc_{kt}^w) \frac{\partial s_{kt}(\mathbf{p}_t)}{\partial p_{jt}} = 0
\end{aligned}
\tag{14}
$$

The first two terms of equation (14) are the same as in Berry et al. (1995), while the third term, $\sum_{k \in S_t^{Au}, k \notin S_t^{Ad}} (wp_{kt} - mc_{kt}^w) \frac{\partial s_{kt}(\mathbf{p}_t)}{\partial p_{jt}}$, introduces a deviation from optimal retail price setting. Rewrite (14):

$$\sum_{k\in S_t^{Ad}}(p_{kt}-wp_{kt}-mc^r_{kt})\frac{\partial s_{kt}(\mathbf{p}_t)}{\partial p_{jt}}$$
$$=-\,s_{jt}(\mathbf{p}_t)-\sum_{k\in S_t^{Au},k\notin S_t^{Ad}}(wp_{kt}-mc^w_{kt})\frac{\partial s_{kt}(\mathbf{p}_t)}{\partial p_{jt}}$$

In vector form:

$$\Delta_t^p(\mathbf{p}_t-\mathbf{wp}_t-\mathbf{mc}_t^r)=-\mathbf{s}_t-\Delta_t^A(\mathbf{wp}_t-\mathbf{mc}_t^w)$$

Solving for the retailer markup gives:

$$\mathbf{p}_t-\mathbf{wp}_t-\mathbf{mc}_t^r=-(\Delta_t^p)^{-1}(\mathbf{s}_t+\Delta_t^A(\mathbf{wp}_t-\mathbf{mc}_t^w)) \tag{15}$$

where

$$\Delta_t^A(j,i)=\begin{cases}\frac{\partial s_{it}}{\partial p_{jt}}, & \text{if } j\in S_t^{Ar},\ i\in S_t^{Aw},\ i\notin S_t^{Ar}\\ 0, & \text{otherwise}\end{cases}$$

To see how A's optimal retail prices differ from its retailer competitors, consider elements of Δ_t^A. For products i and j such that $\Delta_t^A(j,i)$ is not zero, they can never be the same product (sold by different retailers) and $\frac{\partial s_{it}}{\partial p_{jt}}$ is always a cross-price elasticity thus positive. The vector $\mathbf{wp}_t-\mathbf{mc}_t^w$ is the manufacturer markup and therefore positive. The product term, $\Delta_t^A(\mathbf{wp}_t-\mathbf{mc}_t^w)$, is positive, same as the sign of vector $\mathbf{s}_t$, so the existence of it increases manufacturer A's retailer markup and puts upward pressure on A's retailer prices. Moresi and Salop (2013) call it vGUPPId, the vertical gross upward pricing presssure index of the downstream department. For simplicity, I call it the vertical upward pricing pressure (VUPP). This works against the elimination of double marginalization and reduces the efficiency gain of vertical merger.

Define $VUPP_t=\Delta_t^A(\mathbf{wp}_t-\mathbf{mc}_t^w)$ as the term that generates vertical upward pricing pressure, then $VUPP_t$ is a N-by-1 vector with

$$VUPP_t(j,i)=\begin{cases}\sum_{k\in S_t^{Aw},k\notin S_t^{Ar}}(wp_{kt}-mc^w_{kt})\frac{\partial s_{kt}(\mathbf{p}_t)}{\partial p_{jt}}, & \text{if } j\in S_t^{Ar}\\ 0, & \text{otherwise}\end{cases}$$

this shows firm A has incentives to increase all of the products sold by its retailer department (not just products produced and sold by A), while other retailers are not directly affected. Such vertical upward pricing pressure could make products produced by other brands more expensive in retailer A after merger since these yogurts face upward pricing pressure, and their double marginalization isn't eliminated.

3.4.2 Wholesale Price Bargaining

In this section, I discuss the Nash bargaining between firm A and other non-vertically integrated firms. I highlight two effects in bargaining: raising rival costs (RRC) and additional bargaining leverage over rivals (BLR). Their impacts on market equilibrium are similar, but they work through different channels and exist under different conditions (Rogerson (2020)).

The RRC effect exists under two conditions: (a) retail prices are set after wholesale prices; (b) manufacturer has positive bargaining weight. It is introduced by firm A internalizing the downstream competition and increasing rival costs to steer consumers towards its downstream department. On the other hand, the BLR effect arises as long as retailer has positive bargaining weight, and it works on both upstream and downstream rivals. It exists since A has two departments: upstream and downstream. Before merger, there is only one profit compensation term in equation (10): either $\Delta\pi^r_{rwt}$ or $\Delta\pi^w_{rwt}$. After merger, both enter equation (10), giving A additional leverage over rivals and making firm A more powerful in bargaining. Such leverage increases wholesale prices when A sells to downstream rivals and decreases wholesale prices when A buys from upstream rivals. I illustrate these two effects mathematically in the following sections.

Additional Bargaining Leverage over Rivals (BLR) Suppose firm A bargains with a upstream rival, manufacturer w. The manufacturer's net surplus of bargaining remains the same as in equation (9). For firm A,

$$
\begin{aligned}
&\Pi_t^A(\mathbf{wp}_t^{wA}, \mathbf{wp}_t^{-wA}) - \Pi_t^A(\infty, \mathbf{wp}_t^{-wA}) \\
=& \underbrace{\sum_{j\notin S_t^{Ad}, j\in S_t^{Au}} (wp_{jt} - mc_{jt}^w) M s_{jt}(\mathbf{p}_t)}_{\text{A's upstream profit}} \\
&+ \underbrace{\sum_{j\in S_t^{Ad}, j\in S_t^{w}} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t) + \sum_{j\in S_t^{Ad}, j\notin S_t^{w}} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t)}_{\text{A's downstream profit}} \\
&- \underbrace{\left[\sum_{j\notin S_t^{Ad}, j\in S_t^{Au}} (wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t^*) + \sum_{j\in S_t^{Ad}, j\notin S_t^{w}} (p_{kt}^* - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t^*)\right]}_{\text{A's alternative profit without a deal with retailer r}} \\
=& \underbrace{\sum_{j\in S_t^{Ad}, j\in S_t^{w}} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t)}_{\text{profit of selling yogurt from } w} \\
&- \underbrace{\left[\sum_{j\in S_t^{Ad}, j\notin S_t^{w}} (p_{jt}^* - wp_{kt} - mc_{kt}^r) M s_{jt}(\mathbf{p}_t^*) - \sum_{j\in S_t^{Ad}, j\notin S_t^{w}} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t)\right]}_{\text{A's downstream profit compensation, } \Delta\pi_{wAt}^{Ad}} \\
&- \underbrace{\left[\sum_{j\notin S_t^{Ad}, j\in S_t^{Au}} (wp_{jt} - mc_{jt}^w) M s_{jt}(\mathbf{p}_t^*) - \sum_{j\notin S_t^{Ad}, j\in S_t^{Au}} (wp_{jt} - mc_{jt}^w) M s_{jt}(\mathbf{p}_t)\right]}_{\text{A's upstream profit compensation, } \Delta\pi_{wAt}^{Au}} \\
=& \sum_{j\in S_t^{Ad}, j\in S_t^{w}} (p_{jt} - wp_{jt} - mc_{jt}^r) M s_{jt}(\mathbf{p}_t) - \Delta\pi_{wAt}^{Ad} M - \Delta\pi_{wAt}^{Au} M
\end{aligned}
\tag{16}
$$

Compared to non-vertically-integrated manufacturers in bargaining, firm A has an additional term $\Delta\pi_{wAt}^{Ad}$, which is A's upstream department profit compensation when A and w fails to reach a deal. The bargaining first-order condition becomes:

$$
\begin{aligned}
&\zeta \frac{-s_{jt}(\mathbf{p}_t) + \sum_{k\in S_t^{Ad}} \frac{\partial p_{kt}}{\partial wp_{jt}} s_{kt} + \sum_{k\in S_t^{Ad}} \frac{\partial s_{kt}}{\partial wp_{jt}} (p_{kt} - wp_{kt} - mc_{kt}^r)}{\sum_{k\in S_t^{Ad}, k\in \tilde{S}_t^{w}} (p_{kt} - wp_{kt} - mc_{kt}^r) s_{kt}(\mathbf{p}_t) - \Delta\pi_{wAt}^{Ad} - \Delta\pi_{wAt}^{Au}} \\
&+(1-\zeta) \frac{s_{jt}(\mathbf{p}_t) + \sum_{k\in \tilde{S}_t^{w}} (wp_{kt} - mc_{kt}^w) \frac{\partial s_{kt}(\mathbf{p}_t)}{\partial wp_{jt}}}{\sum_{k\in S_t^{Ad}, k\in \tilde{S}_t^{w}} (wp_{kt} - mc_{kt}^w) s_{kt}(\mathbf{p}_t) - \Delta\pi_{Awt}^{w}} = 0
\end{aligned}
\tag{17}
$$

Compared to equation (10), this equation has an additional term $\Delta\pi_{wAt}^{Au}$ in the denominator of A's optimality condition. This term serves as additional bargaining leverage against downstream rivals, allowing firm A to push wholesale prices towards the direction that favors itself. In this case, wholesale prices between retailer w and A will be lower than cases when A is not vertically merged with a manufacturer. Such lower wholesale prices put downward pressure on A's retail prices, benefiting competition and consumer welfare. The BLR effect also exists when A bargains with downstream rivals, though it increases wholesale prices of yogurt produced by firm A and harms competition.

Raising Rival Costs (RRC) When firm A bargains with downstream rivals, apart from the additional bargaining leverage, there's an additional incentive that makes A push wholesale prices higher, known as the Raising Rival Cost effect. Similar to the case where A bargains with an upstream rival, A's net surplus of trade when bargaining with a downstream rival is:

$$\Pi_t^A(\mathbf{wp}_t^{rA}, \mathbf{wp}_t^{-rA}) - \Pi_t^A(\infty, \mathbf{wp}_t^{-rA}) = \sum_{j \in S_t^{Au}, j \in S_t^r} (wp_{jt} - mc_{jt}^w) M s_{jt}(\mathbf{p}_t) - \Delta\pi_{rAt}^{Au} M - \Delta\pi_{rAt}^{Ad} M$$

with the term $\Delta\pi_{rAt}^{Ad}$ serving as additional bargaining leverage. When we take the first order derivative with respect to the wholesale price of product j, the derivative becomes:

$$s_{jt}(\mathbf{p}_t) + \sum_{k \in S_t^{Ad}, k \notin S_t^{Au}} (wp_{kt} - mc_{kt}^w) \frac{\partial s_{kt}}{\partial p_{jt}} + \sum_{k \in S_t^{Ad}} s_{kt} \frac{\partial p_{kt}}{\partial wp_{jt}} + \sum_{k \in S_t^{Ad}} (p_{kt} - wp_{kt} - mc_{kt}^r) \frac{\partial s_{kt}}{\partial wp_{jt}}$$

The additional term $\sum_{k \in S_t^{Ad}} s_{kt} \frac{\partial p_{kt}}{\partial wp_{jt}} + \sum_{k \in S_t^{Ad}} (p_{kt} - wp_{kt} - mc_{kt}^r) \frac{\partial s_{kt}}{\partial wp_{jt}}$ exists only when A bargains with a downstream rival. Compared to non-vertically integrated manufacturers, firm A's manufacturer department internalizes the profit loss of its retailer department due to competition from downstream rivals. To increase total profit, firm A is incentivized to increase downstream rival costs to divert consumers to its own downstream department. This harms competition and consumer welfare.

Coexistence and Differences between RRC and BLR In most cases, the RRC effect and BLR effect work jointly when firm A bargains with a downstream rival. Though they achieve the same goal, which is higher wholesale prices against downstream rivals, their mechanism and the conditions under which they exist are different. To differentiate them, consider the first-order condition of wholesale price bargaining:

$$
\begin{aligned}
&\zeta\frac{-s_{jt}+\sum_{k\in S_t^r}\frac{\partial p_{kt}}{\partial wp_{jt}}s_{kt}+\sum_{k\in S_t^r}\frac{\partial s_{kt}}{\partial wp_{jt}}(p_{kt}-wp_{kt}-mc_{kt}^w)}{\sum_{k\in S_t^r,k\in\tilde{S}_t^{Au}}(p_{kt}-wp_{kt}-mc_{kt}^r)s_{kt}(\mathbf{p}_t)-\Delta\pi_{rAt}^r}+\\
&(1-\zeta)\frac{s_{jt}(\mathbf{p}_t)+\sum_{k\in\tilde{S}_t^{Au}}(wp_{kt}-mc_{kt}^w)\frac{\partial s_{kt}(\mathbf{p}_t)}{\partial wp_{jt}}}{\sum_{k\in S_t^{Au},k\in S_t^r}(wp_{kt}-mc_{kt}^w)s_{kt}(\mathbf{p}_t)-\Delta\pi_{rAt}^{Au}-\Delta\pi_{rAt}^{Ad}}\\
&(1-\zeta)\frac{\sum_{k\in S_t^{Ad}}\frac{\partial p_{kt}}{\partial wp_{jt}}s_{kt}+\sum_{k\in S_t^{Ad}}(p_{kt}-wp_{kt}-mc_{kt}^r)\frac{\partial s_{kt}}{\partial wp_{jt}}}{\sum_{k\in S_t^{Au},k\in S_t^r}(wp_{kt}-mc_{kt}^w)s_{kt}(\mathbf{p}_t)-\Delta\pi_{rAt}^{Au}-\Delta\pi_{rAt}^{Ad}}\\
&=0
\end{aligned}
\tag{18}
$$

Compared to equation (10), this equation has two additional components: the $\Delta\pi_{rAt}^{Ad}$ on firm A's denominator, and $\sum_{k\in S_t^{Ad}}\frac{\partial p_{kt}}{\partial wp_{jt}}s_{kt}+\sum_{k\in S_t^{Ad}}(p_{kt}-wp_{kt}-mc_{kt}^r)\frac{\partial s_{kt}}{\partial wp_{jt}}$ on firm A's numerator. The downstream department compensation $\Delta\pi_{rAt}^{Ad}$ serves as additional bargaining leverage over rivals (BLR) and works only when the retailer's bargaining weight is positive. The term $\sum_{k\in S_t^{Ad}}\frac{\partial p_{kt}}{\partial wp_{jt}}s_{kt}+\sum_{k\in S_t^{Ad}}(p_{kt}-wp_{kt}-mc_{kt}^r)\frac{\partial s_{kt}}{\partial wp_{jt}}$ indicates A's incentive to raise downstream rival costs to divert consumers to its own downstream department. It exists when retail prices are set after wholesale prices and manufacturers have positive bargaining weight. To further demonstrate their difference, I describe two cases in which only one of these two effects exists.

The RRC becomes the only effect that exists when manufacturer's bargaining weight is one ($\zeta=0$). In this case, equation (18) becomes:

$$
s_{jt}(\mathbf{p}_t)+\sum_{k\in S_t^{Ad},k\notin S_t^{Au}}(wp_{kt}-mc_{kt}^w)\frac{\partial s_{kt}}{\partial p_{jt}}+\sum_{k\in S_t^{Ad}}s_{kt}\frac{\partial p_{kt}}{\partial wp_{jt}}+\sum_{k\in S_t^{Ad}}(p_{kt}-wp_{kt}-mc_{kt}^r)\frac{\partial s_{kt}}{\partial wp_{jt}}=0
$$

Neither $\Delta\pi_{wAt}^{Au}$ nor $\Delta\pi_{rAt}^{Ad}$ exists in this equation, so BLR doesn't exist, but

$$
\sum_{k\in S_t^{Ad}}s_{kt}\frac{\partial p_{kt}}{\partial wp_{jt}}+\sum_{k\in S_t^{Ad}}(p_{kt}-wp_{kt}-mc_{kt}^r)\frac{\partial s_{kt}}{\partial wp_{jt}}
$$

is still present, suggesting that even without bargaining, the vertically-merged firm still has an incentive to increase downstream rival costs.

On the other hand, the BLR becomes the only effect when retailer's bargaining weight is one ($\zeta=1$), though the effect does not work through the profit compensation term $\Delta\pi_{wAt}^{Au}$ or $\Delta\pi_{rAt}^{Ad}$. In discussions above, we assume firm A and its rivals have positive net surplus of bargaining so equation (18) holds. In cases where both sides hold positive bargaining weight, such assumption is unlikely to be violated. When

$\zeta = 1$, however, it is strictly unsatisfied because if firm A continues to supply products to downstream rivals at marginal cost (which is true before vertical mergers), it makes no manufacturer profit while increasing the competition its downstream department faces, making the net gain of trade negative. As a result, even with zero bargaining weight, the upstream department of A can still bargain by saying "no" to downstream rivals, forcing them to pay above marginal cost. This is the effect of additional bargaining leverage.

3.4.3 Overall Effect of Vertical Mergers

From the discussions above, we can see that vertical mergers affect market equilibrium in multiple ways. The vertically merged firm benefits consumers by eliminating double marginalization and negotiating lower wholesale prices from upstream rivals, but harming them by raising downstream rival costs, negotiating higher wholesale prices to downstream rivals, and setting retail prices higher than the level that maximizes retailer profit to lessen upstream competition.

In a vertical merger, the four effects discussed above coexist (EDM, VUPP, BLR, RRC), but they affect different types of yogurt. I separate yogurts in a market into four categories based on whether or not A participates in the supply of them:

1. Yogurts produced and sold by A
2. Yogurts produced by other manufacturers and sold by A
3. Yogurts produced by A and sold by other retailers
4. Yogurts not related to A

For yogurts in the first category (produced and sold by A), they have no double marginalization compared to yogurts in other categories but are subject to vertical upward pricing pressure. For the second category (produced by others and sold by A), these yogurts have lower wholesale prices, but are subject to vertical upward pricing pressure. Retail prices of these two categories may increase or decrease after merger, depending on the relative magnitude of various effects. For yogurts in the third category, however, their prices will certainly increase after merger because these yogurts are under the influence of two effects: RRC and BLR, both increasing their wholesale prices. Yogurts not related to A are not directly affected by the vertical merger, but their prices will change to accommodate the changes of other retail prices.

As for firm profits and consumer welfare, firm A gets cheaper yogurts from other manufacturers and sells yogurts to other retailers at higher wholesale prices, so its profit increases after merger. Consequently, other retailers and manufacturers suffer losses. In the retail market, yogurts in non-A retailers are more expensive, and their prices in retailer A are unpredictable without a full merger simulation, so the change of consumer welfare is ambiguous at this stage.

4 Estimation

4.1 Consumer Demand Estimation

The market share in my model is constructed by a simulation algorithm: take N_s random draws of ν_i^c and ν_i^p and simulate s_{jt} by:

$$s_{jt} = \frac{1}{N_s} \sum_{i=1}^{N_s} \frac{e^{\delta_{jt} + \sigma_c \nu_i^c - \sigma \nu_i^p p_{jt}}}{1 + \sum_k e^{\delta_{kt} + \sigma_c \nu_i^c - \sigma \nu_i^p p_{kt}}}$$

Using the method provided by Berry et al. (1995) I calculate the mean utility of product j in market t, $\delta_{jt} = x_{jt}\beta - \alpha p_{jt} + \xi_{jt}$. I assume the unobservable characteristics, ξ_{jt}, is uncorrelated to the vector of instrument variables $\mathbf{z}_{jt}$, then the moment condition to identify consumer demand is:

$$\mathbb{E}\left[\xi_{jt}\mathbf{z}_{jt}\right] = \mathbb{E}\left[(\delta_{jt} - x_{jt}\beta + \alpha p_{jt})\mathbf{z}_{jt}\right] = 0 \tag{19}$$

4.2 Bargaining power estimation

First of all, equation (13) can not be directly computed. The left-hand side is the vector of manufacturer markups, while the right-hand side contains $\Delta\pi_{rw}^w$, which requires manufacturer markups to compute. I develop a iterative process to numerically compute $\mathbf{wp}_t - \mathbf{mc}_t^w$ [4].

Add up equation (6) and (13), we have

$$\mathbf{mc}_t^r + \mathbf{mc}_t^w = \mathbf{p}_t + (\Delta_t^p)^{-1} s^t + T_l^{-1} V$$

I assume $\log(mc_{jt}^r + mc_{jt}^w) = w_{jt}\gamma + \omega_{jt}$, where w_{jt} includes all the input cost variables and ω_{jt} is the error term. Specifically, w_{jt} includes all variables in table 4. However,

[4] See appendix for details

these variables work equally on all the brands and retailer and cannot explain the cross-brand or cross-retailer cost variation, so I add brand and retailer fixed effects to w_{jt}. Finally, I add the number of retailers and manufacturers to account for the economy of scale. The moment condition to identify supply side parameters is:

$$\mathbb{E}\left[\omega_{jt}\mathbf{z}_{jt}\right] = \mathbb{E}\left[(\log(mc_{jt}^{r} + mc_{jt}^{w}) - w_{jt}\gamma)\mathbf{z}_{jt}\right] = 0 \tag{20}$$

All parameters $(\beta, \alpha, \gamma, \sigma, \zeta)$ are estimated by the generalized method of moments(GMM) using moment conditions (19) and (20).

4.3 Identification

I use two sets of moment conditions to identify parameters:

$$\mathbb{E}\begin{bmatrix} (\delta_{jt} - x_{jt}\beta + \alpha p_{jt})\mathbf{z}_{jt} \\ (\log(mc_{jt}^{r} + mc_{jt}^{w}) - w_{jt}\gamma)\mathbf{z}_{jt} \end{bmatrix} = 0$$

I assume product characteristics are predetermined and thus exogenous. This is indeed the case because I don't observe changes in size and multiple over time. Input costs are also considered exogenous because the yogurt industry is only a small part of the economy and is unlikely to affect prices of electricity, gas, or Federal Funds Rate.

On the demand side, retail prices are endogenous and I use two sets of instrumental variables. The first set of instruments are the same as in Berry et al. (1995), which are the sum of own and rival products characteristics. Specifically, the instruments are:

$$\sum_{k \neq j, k \in S_t^r} x_{kt} \text{ and } \sum_{k \notin S_t^r} x_{kt}$$

where S_t^j is the set of yogurt provided by the firm that provides j. The firm here can either be a retailer or a manufacturer. I use both specifications in the demand estimation, and results are very close to each other. The second set are DMA-retailer fixed effects as yogurt characteristics are the same in different DMAs or retailers.

For the identification of bargaining power, input cost variables are all exogenous, but due to the existence of ζ, there has to be at least one instrument variable to identify it. I use the sum of own and rival characteristics as well.

	OLS	IV
	$\log(s_j/s_0)$	$\log(s_j/s_0)$
Price	-0.99***	-1.77***
	(0.005)	(0.0047)
Multi	0.08***	0.05***
	(0.0009)	(0.0009)
Size	0.024***	0.01***
	(0.0004)	(0.0003)
N	663982	663982

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Logit Demand Estimation

5 Results

5.1 Logit demand estimation

For product characteristics, weekly observations of price, size and multiple are available, while other characteristics(flavor, organic, etc.) are annual observations, so I use price, size, multiple as product characteristics. I use logit model as baseline for demand estimation first (see Berry (1994)), and then present result of the full model.

Table 5 presents the result of logit demand estimation. The price coefficient changes little when IV is applied, suggesting a low level of endogeneity. Adding a brand fixed effect increases the price coefficient, suggesting a higher price sensitivity within brands. Since consumers are less likely to switch to cheaper yogurt when there is no fixed effect, I conclude that consumers tend to stick to their favorite brands. Coefficients of size and multiple are positive so consumers prefer larger sizes and higher multiples.

consumer side			
β^c	constant	-7.35 ***	(0.07)
α	price	-3.16 ***	(0.08)
β^{multiple}	multiple	0.04***	(0.001)
β^{size}	size	0.0007	(0.0005)
σ_p	price standard error	0.98***	(0.03)
σ_c	constant standard error	0.02	(4.38)
producer side			
ζ	bargaining power	0.55***	(0.03)
γ^c	constant	-0.54 **	(0.14)
γ^{sugar}	sugar	-0.0066***	(0.0009)
γ^{milk}	milk	0.0007	(0.0013)
γ^{gas}	gasoline	0.02***	(0.01)
γ^{ffr}	Federal funds rate	0.10***	(0.02)
$\gamma^{\text{elec-c}}$	electricity(commercial)	-0.03 ***	(0.01)
$\gamma^{\text{elec-i}}$	electricity(industrial)	0.08***	(0.02)

Estimated using 663,982 observations. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Values of Coefficients in the Full Model

5.2 Full Model

Table 6 reports the results of coefficients of the full model. The consumer side coefficients are broadly the same as in the logit case. Retailer's bargaining power parameter is 0.55 The mean price per serving of yogurt is \$1.35, mean retailer markup is \$0.50 which is much higher than the manufacturer markup, \$0.20.

6 Bargaining Power, Merger Simulation and Welfare

In this section, I present my counterfactual analysis and merger simulation results under various bargaining power specifications.

6.1 Market Equilibrium under Alternative Bargaining Power

Given the cost of yogurt production, transportation, storage and sales, different values of bargaining weight parameter lead to different market equilibria prices and quantities. The goal of the counterfactual analysis is to investigate price, markups, profits, and welfare under alternative bargaining power parameters. In this session, I use total marginal cost $\mathbf{mc}^r + \mathbf{mc}^w$ and consumer preference parameters α, β, and σ estimated in section 4 to compute new market equilibria under a series of ζ.

Figure 1 presents how the mean of retail price, retailer markup, and manufacturer markup changes when ζ increases. Intuitively, when retailers' bargaining power increases, they can push wholesale prices down towards the marginal cost of manufacturers. Lower wholesale prices provide incentives for retailers to lower their retail prices. In the graph, the average retail price and manufacturer markup decrease when retailers' bargaining weight parameter increases. Retailers' markup decreases slightly.

Figure 2 presents the profit and welfare under different bargaining power parameter specifications. Since sales revenue, retailer and manufacturer profit and consumer welfare are on different scales, I normalized these measures of welfare by dividing them by their value under the true case. So if the value of the vertical axis is 2, it means a variable is twice as much as the real case. When ζ increases, total sales volume increases as a result of lower retail prices. The total revenue increases

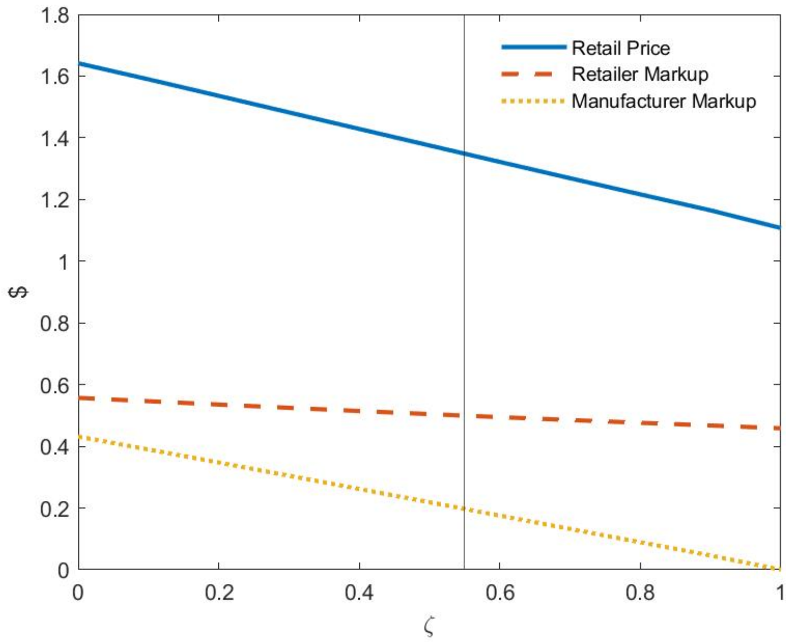

Figure 1: Retail Price, Retailer Markup and Manufacturer Markup

as well despite lower retail prices. Retail profit increases mainly due to larger sales volume, consumer welfare increases due to lower retail prices. Manufacturer profit decreases despite a higher sales volume. Consumer welfare is more sensitive to ζ than retailer profit.

6.2 Horizontal Merger

Horizontal merger is the merger between two firms on the same level of a supply chain, namely two retailers or two manufacturers. After merger, the merged firm optimizes its joint profit instead of individual profit of the merging firms. Since the firm has higher market power, it's likely to increase its prices which lowers consumer welfare. Under different ζ, horizontal merger has different effects. I follow the steps below to demonstrate how misspecification of bargaining power distorts horizontal merger analysis:

- Estimate the model with $\zeta = 0$ and $\zeta = 1$. This step won't change demand parameters and retailer markups, but will produce new sets of manufacturer markups and $\mathbf{mc}_t^r + \mathbf{mc}_t^w$, the total marginal cost.

- Simulate horizontal mergers. In this step, I use the $\mathbf{mc}_t^r + \mathbf{mc}_t^w$ in step one as exogenous marginal costs and simulate market equilibrium prices and quanti-

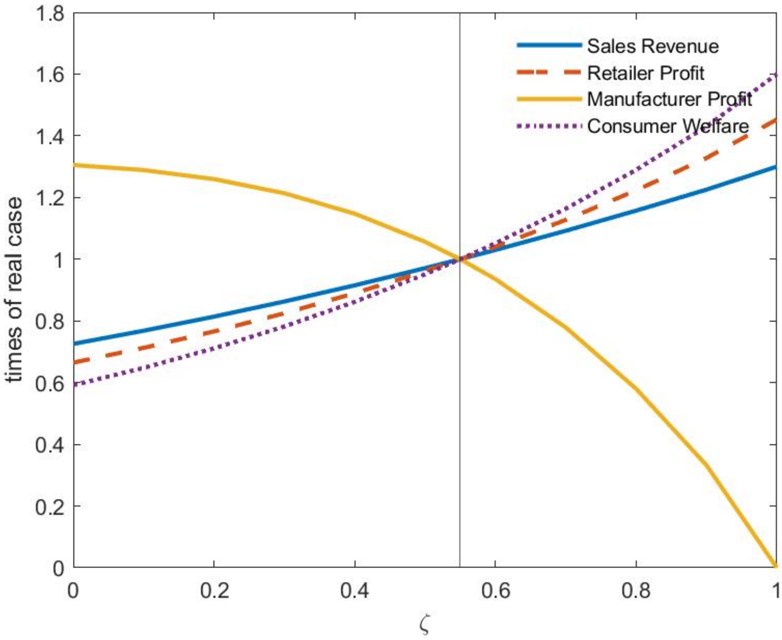

Figure 2: Consumer Welfare, Retailer Profit, Manufacturer Profit and Sales Revenue

ties after merger. Since there are fours sets of $\mathbf{mc}_t^r + \mathbf{mc}_t^w$ corresponding to $\zeta = 0$, $\zeta = 0.55$, $\zeta = 0.95$ and $\zeta = 1$, there will be four sets of prices and quantities.

- For each ζ, compare the profit and welfare before and after merger. Different specifications of ζ introduce different merger effects.

Compared to a traditional market where horizontal merger affects prices through market power, there is an additional affect of horizontal merger in the market with retail price bargaining: strength of outside options. In equation (8), when two retailers merge, manufacturers have fewer outside options when they bargain with the merged firm and their $\Delta\pi_{rwt}^w$ decreases, which weakens their bargaining power. The same story also applies when two manufacturers merge.

Table 7 demonstrates the profit and welfare changes after horizontal mergers. Retailers gain little (2.87%) from monopolizing since their market power before merger is already high, but they can do heavy damage to manufacturers (−11.94%) and consumers (−5.51%). Manufacturers benefit more from becoming a monopoly than retailers (4.12% vs 2.87%), and the damage they do to retailers and consumers are relatively smaller (−2.80% and −3.56%, respectively).

Merger Type	Retailer	Manufacturer	Consumer
Retailer Monopoly	2.87%	-11.94%	-5.51 %
Manufacturer Monopoly	-2.80%	4.12%	-3.56%

For retailer and manufacturer, profit changes after merger are listed. For consumer, welfare changes after merger are listed

Table 7: Profit and Welfare Changes under Horizontal Merger

Retailer Bargaining Weight	Retailer	Manufacturer	Consumer
Retailer Monopoly			
0 (Double Marginalization)	5.20%	-9.21%	-2.81%
0.55 (Correct)	2.87%	-11.94%	-5.51 %
0.95	0.39%	-9.46%	-8.18%
1 (Zero Manufacturer Markup)	0.35%	No Changes	-8.02%
Manufacturer Monopoly			
0 (Double Marginalization)	-4.30%	0.39%	-5.37%
0.55 (Correct)	-2.80%	4.12%	-3.56%
0.95	-0.39%	8.19%	-0.53%
1 (Zero Manufacturer Markup)		No changes	

For retailer and manufacturer, profit changes after merger are listed
For consumer, welfare changes after merger are listed

Table 8: Profit and Welfare Changes under Horizontal Merger

Table 8 demonstrates how misspecifications of ζ lead to biased horizontal merger simulations. In retailer mergers, when retailer's bargaining weight increases, their percentage gains decrease while consumers suffer higher percentage loss of welfare. Manufacturers' percentage loss of profit first increases and then decreases while ζ goes up from zero to one. When manufacturers become a monopoly, their percentage gain increases when retailer bargaining weight increases, and the percentage loss of retailers and consumers decreases.

Table 9 shows the pre- and post-merger prices and markups under various retailer bargaining weights. Consistent with the consumer welfare changes in table 7, the post-merger average prices increase when retailer bargaining weight increases in retailer monopoly simulation. When manufacturers become a monopoly, the post-merger average prices decrease when retailer bargaining weight increases.

Retailer Bargaining Weight	Price	Retailer Markup	Manufacturer Markup
Pre-Merger			
0 (Double Marginalization)	1.35	0.50	0.39
0.55 (Correct)	1.35	0.50	0.20
0.95	1.35	0.50	0.02
1 (Zero Manufacturer Markup)	1.35	0.50	0
Retailer Monopoly			
0 (Double Marginalization)	1.37	0.54	0.36
0.55 (Correct)	1.38	0.54	0.18
0.95	1.39	0.55	0.02
1 (Zero Manufacturer Markup)	1.39	0.54	0
Manufacturer Monopoly			
0 (Double Marginalization)	1.37	0.50	0.41
0.55 (Correct)	1.37	0.50	0.21
0.95	1.35	0.50	0.03
1 (Zero Manufacturer Markup)		No changes	

Table 9: Prices and Markups before and after Horizontal Merger

6.3 Vertical Merger

In this section, I demonstrate the welfare effects of vertical mergers under various market structures. I first construct a hypothetical market with four symmetric retailers and manufacturers (symmetric market hereafter), compute the alternative market equilibrium prices, quantities, and markups. By simulating vertical mergers in this market, I exclude the effect of firm sizes and product assortments on the vertical merger. I then simulate vertical merger in a more realistic market and study how firm sizes change the result of vertical mergers.

The symmetric market is constructed in the following way:

1. Choose a set of popular products from the data, assume all four manufacturers produce the same products in this set, and all four retailers sell products from all four manufacturers. In the market equilibrium, each manufacturer supplies the same products to four retailers, and each retailer sells the same product line-ups from four manufacturers, resulting in 16 identical pairs.

2. Take total marginal cost $\mathbf{mc}_t^r + \mathbf{mc}_t^w$, product characteristics, demand parameters, and the bargaining weight parameter as given, simulate the new equilibrium market prices and quantities as the pre-merger symmetric market

	Retailer A		Other Retailers	
	Brand A	Other Brands	Brand A	Other Brands
Retailer Makrup	VUPP ↑	VUPP ↑		
Manufacturer Markup	EDM ↓	BLR ↓	BLR ↑ +RRC ↑	

Table 10: Four Effects of Vertical Merger

In addition to checking profit and welfare changes, I aim to examine the direct effect of the four channels of vertical merger (EDM, VUPP, BLR, and RRC) on prices and markups. Although they coexist in the vertical merger, they affect different types of yogurt on different levels of a supply chain. Table 10 shows how the four effects change retail prices in the equilibrium. EDM is straightforward since it reduces the manufacturer markup of vertically integrated yogurts to zero. We can tell the effect of VUPP by comparing the retailer markup of firm A's non-vertically integrated yogurts before and after merger. BLR increases wholesale prices of yogurts produced by A and sold by other retailers and decreases wholesale prices when retailer A buys from other manufacturers. The latter effect can be detected by comparing pre-/post-merger manufacturer markups of yogurts produced by non-A manufacturers and sold by retailer A, but the former effect works jointly with the RRC effect so I can't separate them. To obtain the individual effect of BLR and RRC on manufacturer markups of yogurts produced by A and sold by other retailers, I simulate two mergers. In the first merger, I compute the market equilibrium using equation (18). In the second merger, I remove the RRC effect from equation (18). Specifically, I compute market equilibrium using the following equation when manufacturer A bargains with retailer rivals:

$$\begin{aligned}
&\zeta \frac{-s_{jt} + \sum_{k \in S_t^r} \frac{\partial p_{kt}}{\partial wp_{jt}} s_{kt} + \sum_{k \in S_t^r} \frac{\partial s_{kt}}{\partial wp_{jt}} (p_{kt} - wp_{kt} - mc_{kt}^w)}{\sum_{k \in S_t^r, k \in \tilde{S}_t^{Au}} (p_{kt} - wp_{kt} - mc_{kt}^r) s_{kt}(\mathbf{p}_t) - \Delta\pi_{rAt}^r} + \\
&(1-\zeta) \frac{s_{jt}(\mathbf{p}_t) + \sum_{k \in \tilde{S}_t^{Au}} (wp_{kt} - mc_{kt}^w) \frac{\partial s_{kt}(\mathbf{p}_t)}{\partial wp_{jt}}}{\sum_{k \in S_t^{Au}, k \in S_t^r} (wp_{kt} - mc_{kt}^w) s_{kt}(\mathbf{p}_t) - \Delta\pi_{rAt}^{Au} - \Delta\pi_{rAt}^{Ad}} \\
&= 0
\end{aligned} \tag{21}$$

I call the first merger the full merger and the second merger the no-RRC merger. The difference of manufacturer markup between the full merger and no-RRC merger shows us the effect of RRC.

It's worth mentioning that firm A's retailer markups are influenced by both the VUPP effect and the change of wholesale prices. The wholesale prices move in opposite directions for vertically integrated and non-integrated yogurts, so potentially the retailer markups of them move differently after the merger. As a result, I cannot take the change of retailer markups as a measure of the VUPP effect. In the next table, I will show that changes in wholesale prices of yogurts produced by other manufacturers are small in retailer A, so the change of retailer markup is a good approximation of the VUPP effect.

	Retailer A		Other Retailers		Overall
	Brand A	Other Brands	Brand A	Other Brands	All Brands
Average Retail Price($)					
Full Merger	0.94	1.15	1.18	1.15	1.13
No RRC	0.94	1.15	1.16	1.15	1.13
Pre Merger	1.14	1.14	1.14	1.14	1.14
Average Retailer Markup($)					
Full Merger	0.43	0.46	0.45	0.45	0.45
No RRC	0.43	0.46	0.45	0.45	0.45
Pre Merger	0.45	0.45	0.45	0.45	0.45
Average Manufacturer Markup($)					
Full Merger	0	0.17	0.21	0.18	0.17
No RRC	0	0.17	0.20	0.18	0.17
Pre Merger	0.18	0.18	0.18	0.18	0.18

Table 11: Prices and Markups of Vertical Mergers in the Symmetric Market

Assume in this market that a retailer and a manufacturer merges into firm A. Table 11 demonstrates the pre-merger, full merger and no-RRC merger prices and markups of different yogurt. The first column lists the statistics of products produced and sold by firm A, the second column lists the statistics of products produced by other brands and sold by retailer A, etc. Comparing row 5 and row 6, products produced and sold by A have zero manufacturer markup after merger as a result of eliminating double marginalization. Yogurt produced by other brands have lower markup in retailer A (0.18\$ $\rightarrow$ 0.17\$), showing A's higher bargaining power against upstream rivals after merger. A's yogurt has higher manufacturer markup in rival retailers with BLR alone (0.18\$ $\rightarrow$ 0.20\$), while the RRC increases it further (0.20\$ $\rightarrow$ 0.21\$). For firm A, its retailer markups decreases for vertically integrated yogurts (0.45\$ $\rightarrow$ 0.43\$) and increases for non-vertically integrated yogurts (0.45\$ $\rightarrow$ 0.46\$). As discussed above, such difference of changes is due to the difference in the change of wholesale prices As for the retail prices. Since the change in wholesale prices of the second category is only one cent, I take their retailer markup changes as an approximation of the effect of VUPP. As for equilibrium retail prices, A's store-branded yogurt is much cheaper, A's yogurt in other retailers is significantly more expensive. This gives retailer A a huge advantage when selling A-branded yogurt. Other products are slightly more expensive, and the average price of the whole market is slightly lower after merger.

Vertical Merger Type	Manufacturer A	Other Manufacturers	Retailer A
Full	-19.44%	-0.77%	14.22%
No RRC	-21.83%	-0.91%	14.06 %
Vertical Merger Type	Other Retailers	Firm A	Consumer
Full	-2.44%	4.50%	1.62%
No RRC	-1.80%	3.70 %	2.20 %

Table 12: Welfare and Profit Changes of Vertical Mergers in a Symmetric Market

Table 12 shows the profit and welfare changes due to the vertical merger. The manufacturer department of A loses about 19% of profit since it now supplies products to its retailer department at marginal cost. Other manufacturers earn almost the same profit after merger. The retailer department of A earns 14% more profit for cheaper yogurt from its manufacturer department, and other manufacturers suffer a 2% loss due to more expensive products from manufacturer A and the price advantage of retailer A. Firm A's profit increases by 4%. Consumer welfare increases by 1.6%, but the increase comes only from vertically integrated yogurts produced by firm A. For consumers who prefer other brands, they are worse off.

7 Conclusion

This paper offers a method for estimating bargaining power between manufacturers and retailers with retail market data and a set of observable cost shifters, providing a more flexible way of estimating vertical bargaining power parameters than in previous literature. On the consumer side, I apply a random coefficient discrete choice model of yogurt purchases. On the producer/retailer side, I assume a Nash-in-Nash bargaining game and derive the moment condition. All parameters are estimated using GMM. My result demonstrates that retailers have a bargaining weight parameter of 0.55, and retailers have much higher markups than manufacturers. This conclusion is consistent with the existing literature. In counterfactual analysis, I show that when retailers' bargaining power increases, wholesale price and retail price decreases. Retailers gain higher profit, consumers have higher welfare, while manufacturers suffer a loss.

In horizontal merger simulations, when retailers' bargaining weight increases, they earn less from merging but do more damage to manufacturers and consumers. Manufacturers benefit more from merging when their bargaining weight is low, but they do most damage to retailers and consumers when their bargaining weight is in the middle.

In vertical merger simulation, store-branded products become much cheaper due to the elimination of double marginalization, but the vertically merged firm has higher bargaining power against both upstream and downstream rivals, and sets retail prices above the level that maximizes retailer profit. The vertically merged firm extracts most of the efficiency gain from vertical mergers, and consumers gain slightly more welfare, both at the cost of non-vertically merged firms.

Due to data limitations, I make some implicit assumptions about the bargaining model. The first assumption is that the pairs of retailers and manufacturers reaching a deal are exogenous. This assumption also restricted the disagreement payoff, namely $\Delta\pi^{r}_{rwt}$ and $\Delta\pi^{w}_{rwt}$. This assumption makes it impossible for me to simulate foreclosures in vertical merger simulations. With wholesale prices data available, researchers can compute the benefit of strategically excluding a rival as a threat to other rivals, or the alternative payoff of foreclosing a rival. The second assumption is that manufacturers sell yogurt with simple linear pricing. This is probably a strong assumption of this paper since there are lump-sum transfers between retailers and manufacturers in practice. Without wholesale price data, researchers need alternative assumptions on wholesale prices to estimate such transfer.

References

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

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63:841–890.

Collard-Wexler, A., Gowrisankaran, G., and Lee, R. S. (2019). "nash-in-nash" bargaining: A microfoundation for applied work. *Journal of Political Economy*, 127(1):163 – 195.

Crawford, G. S., Lee, R. S., Whinston, M. D., and Yurukoglu, A. (2018). The welfare effects of vertical integration in multichannel television markets. *Econometrica*, 86(3):891–954.

Crawford, G. S. and Yurukoglu, A. (2012). The welfare effects of bundling in multi-channel television markets. *American Economic Review*, 102:643–685.

Cuesta, J. I., Noton, C., and Vatter, B. (2019). Vertical integration between hospitals and insurers. *Working Paper*.

DellaVigna, S. and Gentzkow, M. (2019). Uniform Pricing in U.S. Retail Chains*. *The Quarterly Journal of Economics*, 134(4):2011–2084.

Draganska, M., Klapper, D., and Villas-Boas, S. (2008). A larger slice or a larger pie? an empirical investigation of bargaining power in the distribution channel. *Marketing Science*, 29:57–74.

Hart, O. and Tirole, J. (1990). Vertical integration and market foreclosure. *Brookings Papers on Economic Activity*, 21(1990 Microeconomics):205–286.

Ho, K. and Lee, R. S. (2017). Insurer competition in health care markets. *Econometrica*, 85:379–417.

Horn, H. and Wolinsky, A. (1988). Bilateral monopolies and incentives for merger. *The RAND Journal of Economics*, 19(3):408–419.

Hristakeva, S. (2017). Vertical contracts with endogenous product selection: An empirical analysis of vendor-allowance contracts. *Working Paper*.

Krattenmaker, T. and Salop, S. (1986). Anticompetitive exclusion: Raising rivals' costs to achieve power over price. *The Yale Law Journal*, 96:209–293.

Moresi, S. and Salop, S. C. (2013). vguppi: Scoring unilateral pricing incentives in vertical mergers. *Antitrust Law Journal*, 79(1):185–214.

Rogerson, W. P. (2020). Modelling and predicting the competitive effects of vertical mergers: The bargaining leverage over rivals effect. *Canadian Journal of Economics/Revue canadienne d'économique*, 53(2):407–436.

Salop, S. C. and Scheffman, D. T. (1983). Raising rivals' costs. *American Economic Review*, 73(2):267–71.

Spengler, J. J. (1950). Vertical integration and antitrust policy. *Journal of Political Economy*, 58(4):347–352.

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

Appendices

A Solving counterfactual price numerically

$\mathbf{p}_t^*$ is the counterfactual optimal price given yogurt from $\tilde{s}_t^w$ is unavailable. The equilibrium price satisfies $\mathbf{p}_t + \Delta_t^p(\mathbf{p}_t)\mathbf{s}_t(\mathbf{p}_t) = \mathbf{wp}^t + \mathbf{mc}^r$, and we use this equation to find out this counterfactual price. When constructing counterfactual prices, I fix demand side parameters, product characteristics (including the unobservable ξ), wholesale prices $\mathbf{wp}^t$, and retailer marginal costs mc_{jt}^r. Here are the steps:

- Step 1: collect demand side parameters, mean utility δ_{jt}, and $\mathbf{wp}^t + \mathbf{mc}^r$. Make sure to delete products from w in retailer r's lineup
- Step 2: make an initial guess of *price adjustment*, say $\Delta\mathbf{p}_t = 0$
- Step 3: compute

$$lhs = (\mathbf{p}_t + \Delta\mathbf{p}_t) + \Delta_t^p(\mathbf{p}_t + \Delta\mathbf{p}_t)\mathbf{s}_t(\mathbf{p}_t + \Delta\mathbf{p}_t)$$

- Step 4: let $\Delta\mathbf{p}_t' = (\mathbf{wp}^t + \mathbf{mc}^r - lhs) + \Delta\mathbf{p}_t'$, use it as the new guess of price change, go back to step 3
- Step 5: repeat step 3 and 4 until the difference between two guesses is smaller than designated tolerance level (I use 10^{-9} in practice)

This fixed point algorithm resembles the one in Berry et al. (1995). I do not have a proof that my algorithm is a contraction mapping, but in practice it converges to my tolerance level within ten iterations.

B Matrix Form of Equation (13)

The matrix form of equation (10) is equation (13):

$$\mathbf{wp}_t - \mathbf{mc}_t^w = T_l^{-1}V$$

where

$$T_l = \zeta(V_1 \cdot \mathbf{s}_t').*T_{rw} + (1-\zeta)*diag(\Pi_t^r - \Delta\pi_{rwt}^r)\cdot(\Delta_t^{wp}.*T_w)$$

$$V = \zeta(V_1.*\Delta\overrightarrow{\pi}_{rwt}^w) - (1-\zeta)\cdot(\mathbf{s}_t.*\overrightarrow{(\Pi_t^r - \Delta\pi_{rwt}^r)})$$

$$V_1 = -\mathbf{s}_t + (\Delta_t^{wp}.*Tr)\cdot(\mathbf{p}_t - \mathbf{wp}_t - \mathbf{mc}_t^r) + (\Delta_t^{p/wp}.*T_r)\mathbf{s}_t$$

Here T_{rw}, T_r, T_w, Δ_t^{wp}, $\Delta_t^{p/wp}$ are N-by-N matrices with

$$T_{rw}(i,j) = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are provided by the same retailer and manufacturer} \\ 0, & \text{otherwise} \end{cases}$$

$$T_r(i,j) = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are provided by the same retailer} \\ 0, & \text{otherwise} \end{cases}$$

$$T_w(i,j) = \begin{cases} 1, & \text{if } i \text{ and } j \text{ are provided by the same manufacturer} \\ 0, & \text{otherwise} \end{cases}$$

$\Delta_t^{wp}(i,j) = \frac{\partial s_{jt}}{\partial wp_{it}}$, $\Delta_t^{p/wp}(i,j) = \frac{\partial p_{jt}}{\partial wp_{it}}$, $.*$ is the element-by-element product of two vectors or matrices with the same size, $\overrightarrow{\pi}_{rwt}^{w}$ is a vector with its j-th element being the manufacturer profit compensation $\Delta\pi_{rwt}^{w}$ such that j is provided by retailer r and manufacturer w, $\overrightarrow{(\Pi_t^r - \Delta\pi_{rwt}^r)}$ is a vector such that its j-th is the net surplus of trade of the retailer-manufacturer pair that provides j. For both Δ_t^{wp} and $\Delta_t^{p/wp}$, Villas-Boas (2007) provides method to compute them, I discuss this in the next section.

C Matrix Form of Δ_t^{wp} and $\Delta_t^{p/wp}$

Villas-Boas (2007) derived the analytical solution for $\frac{\partial p_j}{\partial wp_k}$ in equation (9) of the original paper. In this section, I derive the matrix form of this partial derivative. For simplicity, I omit the t in subscripts

The most complicated part of equation (9) is $\sum_{i=1}^{N} T_r(i,j)\frac{\partial^2 s_i}{\partial s_j \partial s_k}(p_i - wp_i - mc_i^r)$, so I start from here.

First order derivative:

$$\frac{\partial s_i}{\partial s_k} = \begin{cases} \alpha s_k(s_k - 1), & i = k \\ \alpha s_i s_k, & i \neq k \end{cases}$$

Second order derivative:

$$\frac{\partial^2 s_i}{\partial s_j \partial s_k} = \begin{cases} \alpha^2(2s_j^2 - 3s_j + 1)s_i, & i = j = k \\ \alpha^2(2s_j^2 - s_j)s_i, & i \neq j = k \\ \alpha^2(2s_j s_k - s_k)s_i, & i = j \neq k \\ \alpha^2(2s_k s_j - s_j)s_i, & i = k \neq j \\ \alpha^2 2s_i s_j s_k, & \text{otherwise} \end{cases}$$

I put a single s_i in each of these derivatives because equation requires me to sum across all i's.

Denote M_1 the matrix such that $T_1(j,k) = \sum_{i=1}^N T_r(i,j) \frac{\partial^2 s_i}{\partial s_j \partial s_k}(p_i - wp_i^r - mc_i^r)$. When $j = k$,

$$\begin{aligned} &\sum_{i=1}^N T_r(i,j) \frac{\partial^2 s_i}{\partial s_j \partial s_k}(p_i - wp_i^r - mc_i^r) \\ =&\sum_{i \neq j} T_r(i,j)\alpha^2(2s_j^2 - s_j)s_i mk_i^r + \alpha^2(2s_j^2 - 3s_j + 1)s_j mk_j^r \\ =&\alpha^2(2s_j^2 - s_j)\sum_i T_r(i,j)s_i mk_i^r + \alpha^2(-2s_j + 1)s_j mk_j^r \end{aligned}$$

When $j \neq k$,

$$\begin{aligned} &\sum_{i=1}^N T_r(i,j) \frac{\partial^2 s_i}{\partial s_j \partial s_k}(p_i - wp_i^r - mc_i^r) \\ =&\sum_{i \neq j,k} \alpha^2 s_j s_k s_i mk_i^r + \alpha^2(2s_j s_k - s_j)s_j mk_j^r + \alpha^2(s_j s_k - s_j)T_r(i,j)s_i mk_i^r \\ =&2\alpha^2 s_j s_k \sum_i T_r(i,j)s_i mk_i^r - \alpha^2 s_k s_j mk_j^r - \alpha^2 s_j s_k T(j,k) mk_k^r \end{aligned}$$

Here mk_i^r means the retailer markup of yogurt i.

Given these formula, T_1 can be separated into four matrices that are easy to compute:

$$T_{1a} = \alpha^2 \begin{bmatrix} 2s_1^2 - s_1 & \cdots & 2s_1 s_N \\ \vdots & \ddots & \vdots \\ 2s_N s_1 & \cdots & 2s_N^2 - s_N \end{bmatrix} .* \begin{bmatrix} \sum_i T_r(i,1)s_i mk_i^r & \cdots & \sum_i T_r(i,1)s_i mk_i^r \\ \vdots & \cdots & \vdots \\ \sum_i T_r(i,N)s_i mk_i^r & \cdots & \sum_i T_r(i,N)s_i mk_i^r \end{bmatrix}$$

$$T_{1b} = -\alpha^2 \begin{bmatrix} s_1^2 & \cdots & s_1 s_N \\ \vdots & \ddots & \vdots \\ s_N s_1 & \cdots & s_N^2 \end{bmatrix} .* \begin{bmatrix} mk_1^r & \cdots & mk_1^r \\ \vdots & \cdots & \vdots \\ mk_N^r & \cdots & mk_N^r \end{bmatrix}$$

$$T_{1c} = -\alpha^2 \begin{bmatrix} s_1^2 & \cdots & s_1 s_N \\ \vdots & \ddots & \vdots \\ s_N s_1 & \cdots & s_N^2 \end{bmatrix} .* \begin{bmatrix} mk_1^r T_r(1,1) & \cdots & mk_N^r T_r(1,N) \\ \vdots & \vdots & \vdots \\ mk_1^r T_r(N,1) & \cdots & mk_N^r T_r(N,N) \end{bmatrix}$$

$$T_{1d} = \alpha^2 \begin{bmatrix} s_1 mk_1^r & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_N mk_N^r \end{bmatrix}$$

Here $.*$ means element-by-element multiplication. It's easy to verify that $T_1 = T_{1a} + T_{1b} + T_{1c} + T_{1d}$

D Fixed Point Algorithm of Computing wp

Equation (10) isn't analytically solvable due to the existence of $\Delta\pi_{rw}^w$. The left-hand side of equation (13) is the vector of manufacturer markup of products $k \in S_r, k \in S_w$, while the right-hand side contains $\Delta\pi_{rw}^w$, which requires manufacturer markup of product $k \notin S_r, k \in S_w$ to compute. I develop an iteration algorithm to solve for equation (13) numerically:

1. Set the initial value of $wp - mc^w$ to be a vector of zeros
2. Use $wp - mc^w$ to compute $\Delta\pi_{rw}^w$ and the corresponding $wp - mc^w$
3. Use the $wp - mc^w$ in step 2 as the new input for $\Delta\pi_{rw}^w$, go back to step 2
4. Repeat step 2 and 3 until $wp - mc^w$ converges

In practice, $wp - mc^w$ converges quickly. After 3 iterations, the difference between the second and the third value of $wp - mc^w$ decreases below 10^{-7}.
There's a trick that greatly reduces the computation time. Note that

$$\begin{aligned}\Delta\pi_{rw}^w = &\sum_{k \notin S_r, k \in S_w} (wp_k - mc_k^w) s_k(p^{t*}) - \sum_{k \notin S_r, k \in S_w} (wp_k - mc_k^w) s_k(p^t) \\ =&(wp_{-rw} - mc_{-rw}^w)' * (s_{-rw}(p^{t*}) - s_{-rw}(p^t))\end{aligned}$$

where wp_{-rw} is the vector of wholesale prices of market t without products $k \in S_r, k \in S_w$, wp_{-rw} is the vector of manufacturer marginal cost of this hypothetical market, $s_{-rw}(p^{t*})$ is the market share under hypothetical optimal prices, and $s_{-rw}(p^t)$

is the market share in the data. Vector $wp_{-rw} - mc^{w}_{-rw}$ is iterated, and $s_{-rw}(p^{t*})$, the vector that requires the most time to compute, remains the same in iteration, so it should be pre-computed and stored.