Retailers' Product Portfolios and Negotiated Wholesale Prices*

João Granja†

October 3, 2018

Abstract

I investigate how the portfolio of products carried by retailers influences negotiated wholesale prices and, ultimately, retail prices and consumer welfare. To this end, I develop and estimate a model in which manufacturers and retailers bargain over wholesale prices. I use the estimated model to perform counterfactual exercises in which (i) all products of a given manufacturer (including private label products) are excluded from a retailer's portfolio, (ii) all products of a given manufacturer become private label products and (iii) two manufacturers merge. I compute changes in wholesale prices, retail prices and consumer surplus. Focusing on one retail chain, I find small indirect effects on prices: excluding a manufacturer's products leads to only a small increase in other wholesale prices and assuming branded products become private labels decreases other wholesale prices only slightly. However, welfare effects can be large because of reduced product variety and lower prices of the new private label products. Mergers can lead to large increases in wholesale and retail prices, but can also lead to price decreases, depending on the manufacturers involved. A direct consequence is that mergers can be harmful or beneficial to consumers: changes in consumer surplus range from -14.06% to 17.90%.

*Submitted as my third year paper at the University of Pennsylvania.

†University of Pennsylvania, Department of Economics, joaog@sas.upenn.edu

1 Introduction

Consumer goods are a large share of the American economy. In 2016, the food, beverages and tobacco products sector alone accounted for $283.7 billion of the country's GDP[1]. In the past two decades, much has been learned about these industries, from the sources of market power - see Nevo (2001) - to the effects of mergers - see Nevo (2000) and Miller and Weinberg (2017). However, much less is known about the interaction between manufacturers and retailers and its impact on prices and consumer welfare. I partially fill this gap by considering how the portfolio of products carried by retailers affects the wholesale prices they negotiate with manufacturers, which in turn affect retail prices and consumer welfare.

Specifically, I consider the effects of private label products and the upstream market structure. Intuititively, if a retailer sells private label products that are well liked by its customers, this retailer is in a better position when negotiating with the manufacturers of the national brands. This logic extends: if a retailer has a relationship with one manufacturer, this retailer is in a better bargaining position when negotiating with another manufacturer. Moreover, the nature of retailers' relationships with different manufacturers might be different, implying different levels of wholesale mark-ups across manufacturers. My objective is to empirically evaluate how this price negotiation mechanism impacts wholesale prices, retail prices and consumer welfare.

To tackle the questions outlined above, I develop and estimate a model of wholesale price negotiation between retailers and manufacturers. Having estimated the model, I then conduct a series of counterfactual exercises that address the questions above. The counterfactual exercises involve solving the model under the estimated parameter values and different upstream ownership structures[2].

[1]The figure for the entire non-durables sector is $1,005.1B. See `https://www.bea.gov/iTable/iTable.cfm?ReqID=51&step=1#reqid=51&step=51&isuri=1&5114=a&5102=1`.

[2]In this paper, when I refer to "upstream ownership structure", I mean which manufacturers produce which goods, *including private label products*. For example, one structure

The model has two stages. In the first stage, every retail chain - which consist of many stores - negotiates with multiple manufacturers over wholesale prices. Given the wholesale prices agreed upon with each manufacturer, each store then sets its own prices to maximize profits. It is assumed that every store is a local monopolist, so that competition between stores is assumed away. The wholesale prices are assumed to be the outcome of a Nash-in-Nash equilibrium between the retail chain and the multiple manufacturers the chain negotiates with, taking as given the fact that stores price optimally. Given the sequential nature of this game, I refer to an equilibrium of the model as a subgame perfect Nash-in-Nash equilibrium.

The model is estimated using the IRI Academic Dataset[3]. It includes price and quantity information for consumer goods in 30 categories at the week-store-product(UPC) level. The data spans 50 different geographic regions, includes information on product characteristics and all stores are linked to the chains they belong to, which are anonymized. I focus on one category, namely peanut butter. I estimate demand at the chain-region level and, using the demand estimates and the store pricing assumption, I obtain stores' marginal costs. From stores' marginal costs I derive wholesale prices at the chain-product level. These wholesale prices are then treated as data in the estimation of the bargaining part of the model.

I estimate the bargaining part of the model by GMM. Specifically, I show that the bargaining model admits an inversion, which allows me to solve for manufacturers' marginal costs. Together with a model for these costs, this allows me to solve for the econometric error term as a function of (constructed) data and parameters. The bargaining and marginal cost parameters can then be estimated using appropriate orthogonality restrictions. The main identifying assumption is that manufacturers' marginal cost shocks are uncorrelated across different retailers.

The estimated model is then used to conduct counterfactual exercises. In these exercises I solve for subgame perfect Nash-in-Nash equilibria

that will be considered is that in which the private label products seen in the data are actually manufactured and sold by one of the manufacturers of national brands.

[3]For a description of the original release of these data, see Bronnenberg, Kruger, and Mela (2008)

under different assumptions on retailers' product portfolios and the upstream ownership structure. The goal is to understand the effect of these different assumptions on equilibrium wholesale prices, retail prices and consumer surplus. I focus on one chain[4] and consider three types of counterfactual exercises. In the first set of counterfactuals, I consider the scenario in which all the products of a given manufacturer, including private label products, are simply eliminated from the chain's product portfolio. In the second set of counterfactuals, I consider the situation in which all the products of a given manufacturer become private label products. Finally, I consider upstream mergers (including the case in which private label products are instead sold by one of the manufacturers).

I find that indirect equilibrium effects are small. For example, eliminating private label products induces only a small increase in equilibrium wholesale prices of branded products. Assuming all products of a given manufacturer become private label products also induces only a small decrease in the wholesale prices of the products of other manufacturers. Nevertheless, prices of goods produced by one of the manufacturers involved in the counterfactual can change substantially. These prices can increase or decrease, depending on which manufacturers are considered. As a consequence, mergers can be harmful or beneficial to consumers. Assuming products of a given manufacturer become private label products can generate large increases in consumer surplus: even though the indirect effects on the wholesale prices of other manufacturers is small, consumers benefit greatly from the lower prices.

This paper relates to two connected strands of the literature in empirical Industrial Organization. First, it relates to work studying vertical relationshps in retail markets, most notably Villas-Boas (2007) and Draganska, Klapper, and Villas-Boas (2010). Both of these papers try to learn about aspects of vertical relationships between manufacturers and retailers without observing data on wholesale price, much like the present paper. Villas-Boas (2007) considers the implications of 6 non-nested models of verti-

[4]Future versions of this paper will perform the counterfactuals for all chains in the data.

cal relationships and tests which model is most consistent with the available data. Draganska et al. (2010) instead consider a model of bargaining between retailers and manufacturers to study how profits in the vertical chain are distributed between manufacturers and retailers.

Second, this paper is connected to the recent literature on the structural estimation of bargaining models, in particular models employing the Nash-in-Nash solution concept based on Horn and Wolinsky (1988). The first paper to accomplish estimation of a Nash-in-Nash bargaining model was Crawford and Yurukoglu (2012). Other important contributions to this literature are Gowrisankaran, Nevo, and Town (2015), Grennan (2013), Crawford, Lee, Whinston, and Yurukoglu (2015) and Ho and Lee (2017b). A microfoundation for the Nash-in-Nash solution concept was recently provided by Collard-Wexler, Gowrisankaran, and Lee (2014). Another important contribution to the literature on the strucural estimation of bargaining models is Ho and Lee (2017a), who endogenize the network of relationships, an aspect of the data that all the papers cited above take as given, as does the present paper. It should be noted that Draganska et al. (2010) is another paper that estimates a bargaining model, but there are relevant differences, which are discussed below, between their approach and the Nash-in-Nash approach.

This paper complements the literatures above in a number of ways. As Draganska et al. (2010) note, a model in which wholesale prices are determined via Nash bargaining partially generalizes Villas-Boas (2007), as it nests some of the models considered in the latter paper[5]. However, (Draganska et al., 2010) doesn't have a theory for the simultaneous determination of wholesale prices and their model of bilateral negotiations is inconsistent with retailer profit maximization[6]. By fully specifying a model of simultaneous determination of wholesale prices, this paper circumvents

[5]To the extent that bargaining parameters are allowed to vary across manufacturers, retailers or both, a Nash bargaining model generalizes the approach in Villas-Boas (2007) in another direction: the relationship between a retailer and different manufacturers can exhibit varying degrees of inefficiency, something that is not allowed for in Villas-Boas (2007)

[6]As they argue themselves.

that difficulty and is able to embed retailer profit maximization into the vertical relationship model[7]. Moreover, by fully specifying how wholesale prices are determined, I'm able to compute equilibria under counterfactual assumptions.

This paper applies the tools of empirical Nash-in-Nash bargaining models to study retailer-manufacturer interactions[8]. It also extends these tools. I show that the bargaining model introduced in this paper, featuring downstream profit maximization, has an inversion. This enables me to solve for manufacturer marginal costs, which in turn allows me to estimate bargaining and manufacturer marginal cost parameters by GMM, without solving the model. This extends the econometric approach of Gowrisankaran et al. (2015) to the case in which there is downstream profit maximization[9].

The rest of this paper proceeds as follows. Section 2 describes the dataset and provides some reduced form analysis; section 3 introduces the model and provides useful theoretical results; section 4 discusses the estimation of the model; section 5 performs the counterfactual exercises and section 6 concludes.

[7]At the cost of assuming away retailer competition. For more on this and for a more thorough comparison of the model introduced in this paper and Draganska et al. (2010), see Section 3.

[8]Other papers that study retailer-manufacturer relationships and estimate bargaining models are, as already noted, Draganska et al. (2010) and Meza and Sudhir (2010), who study the question of whether private label products incrase retailer bargaining power.

[9]Constructing and estimating a coherent equilibrium model of the determination of wholesale prices while allowing for downstream price *competition* seems to be more challenging. To see why, consider that setting and suppose a manufacturer m increases the wholesale price that chain h pays for good j. Chain h's stores will tend to increase the price of good j. This will shift demand to competing retailers $-h$ and will thus increase manufacturer m's profits from retailers $-h$. A coherent model of vertical relationships with downstream price competition thus needs to consider this demand shifting effect. Interestingly, the predictions of such a model relative to a model in which stores are monopolists is ambiguous: downstream price competition tends to drive prices down, but now the prices a manufacturer charges different retailers for the same good are strategic complements, which tends to drive prices up.

2 Data

The data used in this paper comes from the IRI Academic Database. For a description of the original release of this data, see Bronnenberg et al. (2008). IRI provides information on prices and quantities at the store-week-UPC level, for 30 product categories and 12 years. The data also provides information on the coarse geographic location[10] of each store and to which (anonymized) chain each store belongs to. I focus on one year of data, 2004. Tables 1 and 2 have some motivating evidence for the analysis that follows. Table 1 shows some descriptive statistics for 12 of the 30 categories and table 2 shows results of reduced form analysis for the same 12 categories. In table 1, the second column shows the number of manufacturers in each product category; the third column shows the number of products (UPCs) in each category; the fourth column shows the share of private label products[11].

Table 1: Product Portfolios and Prices

Category	n_manuf	n_prods	share_sb
Blades	17	563	32.5
Cleaning	111	723	15.77
Diapers	14	536	24.44
Facial Tissue	33	346	18.5
Hotdog	155	1100	11.64
Mayonnaise	67	411	25.55
Mustard/Ketchup	242	1101	18.8
Paper Towel	16	540	57.41
Peanut Butter	38	294	24.15
Razors	7	115	7.83
Sugar Substitutes	34	163	26.99
Toothbrush	64	757	11.89

Table 2 is slightly more involved. The second column shows how much

[10]There are 50 values for this location variable and these values are not always at the same geographic level. Most locations are cities but (i) there are cities of varying size, (ii) Two of the locations are regions (New England and West Texas/New Mexico) and (iii) two locations are states (Mississipi and South Carolina).

[11]The number of private label products in the data divided by the total number of products

of the price variation is across chains. Specifically, I estimate a regression of prices on UPC dummies and a regression of prices on dummies for (UPC, Chain) pairs; the first column then reports the difference between the R^2 of the second regression and the R^2 of the first regression, for each category[12]. For some product categories, e.g. diapers and paper towels, variation across chains explains only a small fraction of total price variation; for other product categories, e.g. cleaning products, mustard/ketchup and toothbrush, variation across chains accounts for a non-negligible fraction of total price variation.

Table 2: Product Portfolios and Prices

Category	price_var_chain	sb_effect	sb_portf_effect	competition_effect
Blades	2.39	-1.28	-14.1	-0.83
Cleaning	7.96	-4.71	-13.74	-0.52
Diapers	1.78	-4.63	-8.27	1.05
Facial Tissue	3.12	15.9	19.26	0.63
Hotdog	4.52	15.38	6.86	0.55
Mayonnaise	5.85	4.48	-8.8	0.25
Mustard/Ketchup	7.95	4.73	4.17	-0.23
Paper Towel	2.35	-0.15	9.66	2.92
Peanut Butter	6.41	15.1	20.69	-1.9
Razors	7.59	6.78	2.78	-1.43
Sugar Substitutes	7.37	-13.1	-30.14	-3.08
Toothbrush	11.08	8.81	6.68	-1.05

In the third column of table 2, I evaluate the correlation between prices and the prevalence of store brand products. Specifically, for each category I run regressions of the form

$$\ln(p_{jst}) = \gamma_j + \beta SPL_{st} + \varepsilon_{jst} \tag{1}$$

where p_{jst} is the price of product j in store s in week t, γ_j is a product fixed effect and SPL_{st} in the share of private label products sold in store s and week t, defined by the total quantity of private label products

[12]The R^2 for these regressions is HIGH, BUT GIVE ACTUAL NUMBERS. This implies that there's little price variation across stores of a given chain, consistent with Adams and Williams (2017) and DellaVigna and Gentzkow (2017).

sold in the store divided by the total quantity of products sold. I run these regressions on randomly drawn subsamples of the data with 100,000 observations, restricted to branded products.

The most interesting feature of the results is that five out of the twelve coefficients are negative. Note that equation (1) relates the *price* of branded products and the *share* of private label products. As long as branded products and private label products are substitutes, the correlation between these two variables should be negative, mechanically. Moreover, taking optimal pricing into account also suggests a positive correlation between these variables: suppose, say, that consumers' perceptions about a retailer's private label products improves; the retailer should then increase the price of the private label products and, as prices are strategic complements because of substitutability, the retailer should also increase the price of the branded goods. The fourth column of table 2 shows the results of a similar exercise. I run similar regressions, but instead of using the share of private label products in sales, I use the share of private label products in the store's portfolio, i.e., the number of private label products in store s in week t divided by the total number of products. The signal of the coefficient of interest changes for two product categories, but qualitatively the results are similar.

Finally, in the last column of 2, I evaluate the correlation between prices and the number of manufacturers a retailer negotiates with. Specifically, I run the following regressions:

$$\ln(p_{jst}) = \gamma_j + \beta m_{st} + \varepsilon_{jst} \tag{2}$$

where m_{st} is the number of manufacturers that have some of their products being sold in store s in week t. I also run these regressions on randomly drawn subsamples of size 100,000 and restricting attention to branded products. Seven out of the 12 estimated coefficients are negative, which suggests that retailers that negotiate with more suppliers obtain better wholesale prices.

In summary, tables 1 and 2 show that (i) for some categories there's non-negligible price variation across chains and (ii) in some cases retail prices are lower if retailers' private label products are more prevalent or if they deal with more suppliers. These facts could possibly be explained by different retailers facing different demand and competitive environments. For example, it might be the case that private label products are mostly introduced as cheap alternatives for national brands by chains that tend to locate in poorer neighborhoods, and thus have lower demand for national brands, driving the negative correlations in columns 3 and 4 of table 2. It is also the case that retailers compete both in prices and product offerings, which could help explain the negative correlations in the last column of table 2.

In this paper I'll consider another possible mechanism: store brand products and relationships with other suppliers put retailers in a better position when bargaining with manufacturers. Being in a better bargaining position, retailers will obtain better wholesale prices, which might translate into lower retail prices, depending on the demand function faced by retailers. To empirically evaluate these mechanisms, I'll estimate a model of wholesale price negotiation, which is described in detail in section 3.

The model will be estimated on data for peanut butter[13] for the year 2004. I focus on stores that appear in the data for at least 26 weeks. Conditional on this criterion, I restrict the sample to chains for which I observe at least 5 stores. Finally, conditional on these two criteria, for each store I keep only products that account for at least 5% of total revenues from peanut butter in that store in some week of the year. In the model introduced below, stores are assumed to be local monopolists and every pair (store, week) is treated as a market. The size of the market is assumed to be 1.5 times the maximum total number of units of peanut butter sold in a given store, where the maximum is taken over the weeks in the data.

The selection criteria above leave me with 1,330,205 observations at the store-week-UPC level. There are 22 manufacturers in the data and 199 different products, of which 30.15% are private label products. The

[13]Perhaps not the best choice in light of table 2. A future iteration of this paper will apply the methodology below to a different product category.

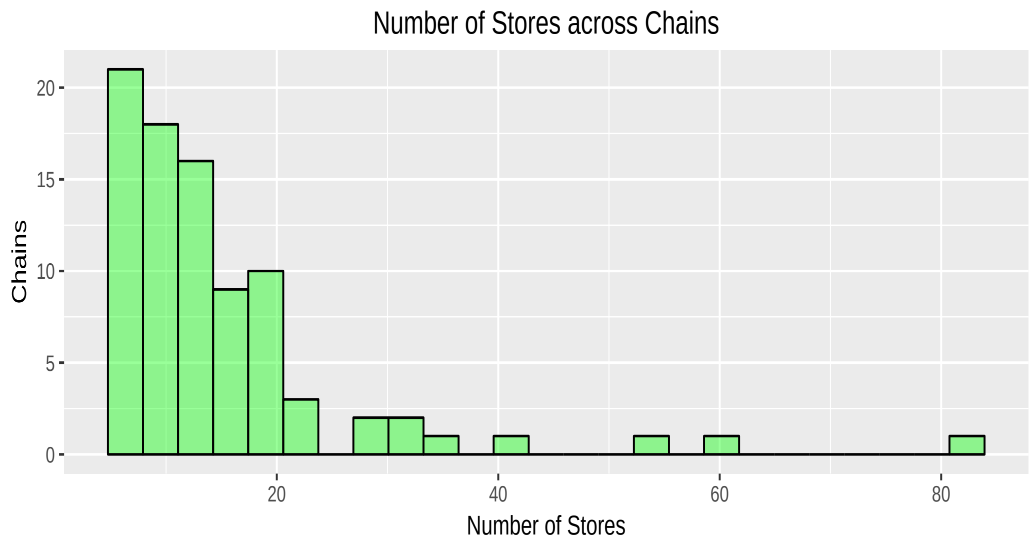

Figure 1: Distribution of the number of stores across chains

observations are distributed across all the 50 geographic locations in the data and across 86 chains. These chains vary considerably in size[14] (see figure 1), number of manufacturers they negotiate with (see figure 2) and the market share of private label products (see figure 3).

3 Model

In section 3.1, I introduce the main components of the model, leaving demand functions unspecified. In section 3.2, I complete the description of the model by specifying the discrete choice problems that consumers face, from which stores' demand functions are derived.

[14]Figure 1 plots the number of stores *in the data*. The data need not be exhaustive, but assuming IRI's sampling of stores across chains is similar, this suggests that these retail chains do differ in size. In the model and empirical implementation, though, the number of stores in the data will be taken to reflect reality: chains' profits are defined to be the sum of stores' profits.

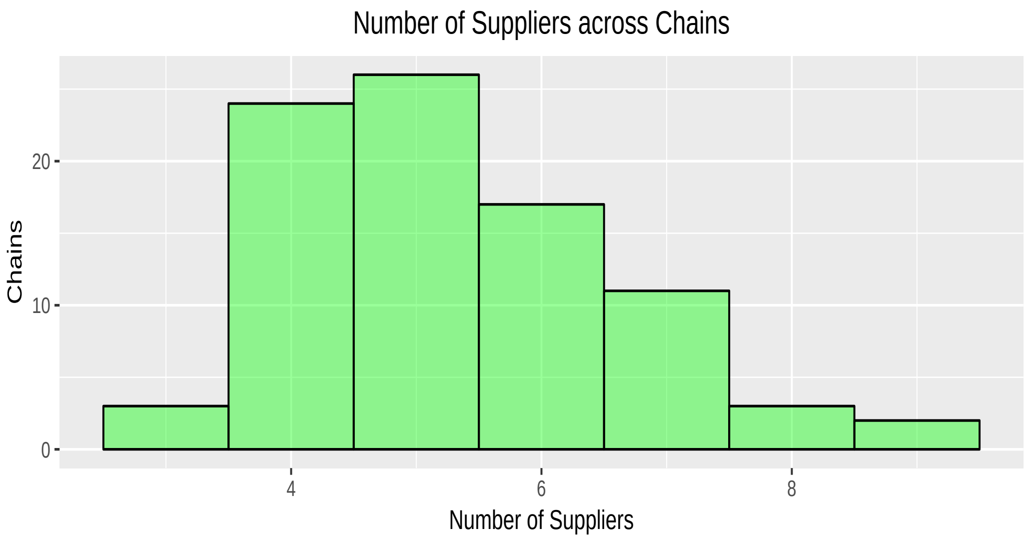

Figure 2: Distribution of the number of suppliers across chains

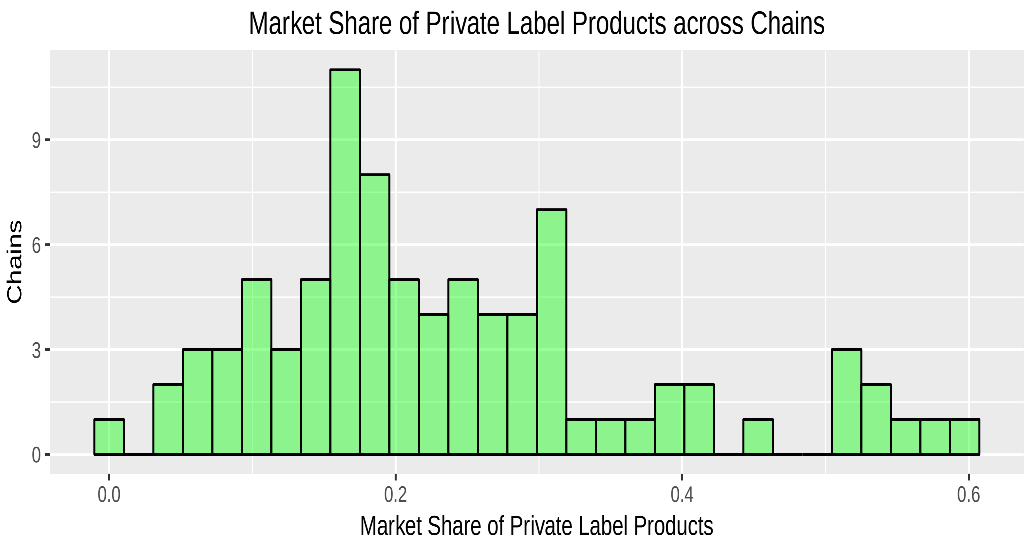

Figure 3: Distribution of the share of private label products across chains

3.1 Model: Retailer Profit Maximization and Negotiated Wholesale Prices

Retail chains are indexed by h. Chain h owns stores $s = 1, \ldots, S_h$. The industry is endowed with a set of products, denoted by[15] $\mathcal{J}$. Products are indexed by $j = 1, \ldots, J$, where $J := |\mathcal{J}|$. A store s carries an exogenously given portfolio of products $\mathcal{J}^s \subseteq \mathcal{J}$ and faces a demand function $D^s : \mathbb{R}^{J^s} \to \mathbb{R}^{J^s}$, mapping that store's prices into quantities demanded by consumers. Note that by specifying the demand of store s as a function of that store's prices only, I'm assuming there's no competition between stores, i.e., stores are assumed to be local monopolists. Given marginal costs c^s, stores set prices to solve

$$\max_{p} \sum_{j \in \mathcal{J}^s} (p_j - c_j^s) D_j^s(p)$$

Let $p^s(c^s)$ denote the solution to this problem. In section 3.3 I show that the solution to this problem is indeed unique given the demand functions I specify.

Let $\mathcal{J}_h := \cup_{s=1}^{S_h} \mathcal{J}^s$ be the set of products sold by chain h. Each such product is either a private label product or a branded product, i.e., a product produced by some manufacturer m. I assume that chains are perfectly vertically integrated with respect to their private label products[16].

Let $\mathcal{J}_{h,m}$ be the set of products sold by chain h and manufactured by m. Chain h bargains with manufacturer m over the wholesale prices of these products, $w_{h,m} = (w_{h,j})_{j \in \mathcal{J}_{h,m}}$. Denote the vector of wholesale prices of the branded products sold by chain h by w_h. I assume that, given w_h, the marginal costs of the stores owned by chain h are given by

$$c_j^s = \begin{cases} k_{h,j} + \tau^s + \eta_j^s, & \text{if } j \in \mathcal{J}_{PL}^s \\ w_{h,j} + \tau^s + \eta_j^s, & \text{if } j \notin \mathcal{J}_{PL}^s \end{cases} \tag{3}$$

[15] Empirically, $\mathcal{J}$ corresponds to the set of all products seen in the data.

[16] An alternative assumption that is also empirically feasible is that all store brand products are produced by a single manufacturer.

where $\mathcal{J}^s_{PL}$ denotes the set of private label products sold by store s and $k_{h,j}$ is the chain's marginal cost of producing good j.

Chain and manufacturers bargain over wholesale prices $w_{h,m}$ *before* the shocks to the stores' marginal costs are observed. After wholesale prices are agreed upon, cost shocks to the stores realize and the stores set prices $p^s(c^s)$. The chain then purchases $D^s_j(p^s(c^s))$ at $w_{h,j}$ and sells $D^s_j(p^s(c^s))$ at price $p^s_j(c^s)$ - for each store s. Hence, the value for the chain of reaching an agreement with manufacturer m at wholesale prices $\hat{w}_m$, holding fixed the wholesale prices agreed upon with other manufacturers, w_{-m}, is given by

$$V_h(\hat{w}_m, w_{-m}; \mathcal{J}_h) = \mathbb{E}_\eta\Bigg[\sum_{s=1}^{S_h}\sum_{j\in\mathcal{J}^s}(p^s_j(\tilde{c}^s(\hat{w}_m, w_{-m})) - \tilde{c}^s_j(\hat{w}_m, w_{-m}))\times \\ \times D^s_j(p^s(\tilde{c}^s(\hat{w}_m, w_{-m})))\Bigg]$$

If chain h and manufacturer m do *not* reach an agreement, the chain caeses to carry all of that manufacturers products. Each store then faces an alternative demand $\bar{D}^{m,s}_j : \mathbb{R}^{|\mathcal{J}^s\setminus\mathcal{J}_{h,m}|} \to \mathbb{R}^{|\mathcal{J}^s\setminus\mathcal{J}_{h,m}|}$, where the superscript m indicates which negotiation failed. In case of disagreement, the chain obtains

$$V_h(w_{-m}; \mathcal{J}_h \setminus \mathcal{J}_{h,m}) = \mathbb{E}_\eta\Bigg[\sum_{s=1}^{S_h}\sum_{j\in\mathcal{J}^s\setminus\mathcal{J}_{h,m}}(p^s_j(\tilde{c}^s(w_{-m})) - \tilde{c}^s_j(w_{-m}))\times \\ \times \bar{D}^{m,s}_j(p^s(\tilde{c}^s(w_{-m})))\Bigg]$$

If chain h and manufacturer m do reach an agreement, the value of the relationship for the manufacturer is

$$V_{m,h}(\hat{w}_m, w_{-m}) = \mathbb{E}_\eta\left[\sum_{s=1}^{S_h}\sum_{j\in\mathcal{J}^s\cap\mathcal{J}_{h,m}}(\hat{w}_j - c^m_j)D^s_j(p^s(\tilde{c}^s(\hat{w}_m, w_{-m})))\right]$$

where c^m_j is the manufacturer's (constant) marginal cost of producing good j. Finally, if an agreement is not reached, the value of the relationship for the manufacturer is zero[17]. I can now define the solution concept for this game.

Definition 1. A vector $w \in \mathbb{R}^{|\mathcal{J}_h \setminus \cup_s \mathcal{J}^s_{PL}|}$ is a **subgame perfect Nash-in-Nash equilibrium (SPNiN) wholesale price vector** if, for every manufacturer m such that $\mathcal{J}_{h,m} \neq \emptyset$, the vector $w_m \in \mathbb{R}^{|\mathcal{J}_{h,m}|}$ solves

$$\max_{\hat{w}_m} V_{m,h}(\hat{w}_m, w_{-m})^{b_{m,h}} \times (V_h(\hat{w}_m, w_{-m}; \mathcal{J}_h) - V_h(w_{-m}; \mathcal{J}_h \setminus \mathcal{J}_{h,m}))^{b_{h,m}} \quad (4)$$

where $b_{m,h}$ is the manufacturer's bargaining power when bargaining with chain h and $b_{h,m}$ is the chain's bargaining power when bargaining with manufacturer m[18].

3.2 Model: Demand

Each store faces a mass M^s of consumers that either buy a single product at the store or don't buy anything. If consumer i buys product j in store[19] s, which is owned by chain h and located in the geographic region l, she enjoys conditional indirect utility

$$u_{ijs} = \gamma_{jhl} + \phi_s + \alpha_{h,l} p_{js} + \psi_{h,l} a_{js} + \xi_{js} + \varepsilon_{ijs}$$

where γ_{jhl} is a product fixed effect, ϕ_s is a store fixed effect, p_{js} is the price of good j in store s, a_{js} is an advertisement dummy, ξ_{js} are product characteristics that are unobserved by the econometrician and ε_{ijs} are preference shocks.

I'll assume, as is standard, that the shocks ε_{ijs} are iid with a Type 1

[17]The assumption underlying the way the value for the manufacturer in both contigencies (agreement or not) is specified is that the bargaining problems that a manufacturer faces with different chains are entirely independent of one another. This is a common assumption in the empirical bargaining literature. It should be noted that it requires (i) absence of downstream price competition and (ii) constant marginal costs for the manufacturer.

[18]Without loss of generality, I impose $b_{h,m} + b_{m,h} = 1$.

[19]Choosing in which store to purchase is *not* a choice in the model.

Extreme Value distribution. Then the share of good j in store s is given by

$$\sigma_j^s(p) = \frac{\exp(\delta_j(p_j))}{1 + \sum_{k \in \mathcal{J}^s} \exp(\delta_k(p_k))}$$

where $\delta_k := \gamma_{khl} + \phi_s + \alpha_{h,l} p_{ks} + \psi_{h,l} a_{ks} + \xi_{ks}$

Store s thus faces the demand function

$$D_j^s(p) = M^s \sigma_j^s(p)$$

The disagreement demand functions $\bar{D}^{m,s}$ are derived from the discrete choice model in the same way.

3.3 Theoretical Results

This section establishes two results that are used in the subsequent analysis. First, I provide a full characterization of the solution - which turns out to be unique - to a monopolist's profit maximization problem under logit demand, as in the case of stores in the model introduced above.

Proposition 1. Suppose a monopolist faces a demand function $D : \mathbb{R}^J \to \mathbb{R}^J$ given by

$$D_j = M\sigma_j(p) = M \frac{\exp(\delta_j(p_j))}{1 + \sum_{k=1}^{J} \exp(\delta_k(p_k))}$$

where $\delta_k(p_k) = \gamma_k + \alpha p_k$ and $\alpha < 0$. Then

(i) There exists a unique solution p^* to the monopolist's profit maximization problem.

(ii) The solution p^* exhibits constant mark-ups, i.e., there exists a $\mu > 0$ such that

$$p_j^* - c_j = \mu, \quad \text{for all } j = 1, \ldots, J$$

(iii) The optimal mark-up μ is given by the unique solution to

$$1 + \alpha\mu\sigma_0(c + \mu) = 0 \tag{5}$$

where $\sigma_0(p) = 1/\left(1 + \sum_k \exp(\delta_k(p_k))\right)$ is the share of the outside good, $c = (c_1, \ldots, c_j)'$ is the vector of marginal costs and $c + \mu$ means that μ is added to every coordinate of c.

(iv) p^* is a continuously differentiable function of $c \in \mathbb{R}^J$ and its derivatives are given by

$$\frac{\partial p_j^*}{\partial c_k} = \begin{cases} 1 + \frac{\partial \mu^*}{\partial c_j}(c) & \text{if } k = j \\ \frac{\partial \mu^*}{\partial c_k}(c) & k \neq j \end{cases}$$

where
$$\frac{\partial \mu^*}{\partial c_k}(c) = \frac{\alpha \mu^*(c) \sigma_k(c + \mu^*(c))}{1 - \alpha \mu^*(c)(1 - \sigma_0(c + \mu^*(c)))}$$

Proof. See appendix A. ♠

The result above is of independent interest, because logit demands are used widely in industrial organization. Existence of a unique solution and the constant mark-up property are private cases of results in Nocke and Schutz (2018), but I provide independent proofs of those facts. The analysis of SPNiN equilibria that follows assumes that retail prices depend smoothly on wholesale prices - see Proposition 2. Part (iv) of Proposition 1 establishes that fact and characterizes the relevant derivatives. This explicit characterization makes computation considerably more efficient: the only step that has to be done numerically is the solution of equation (5), which is a very well behaved equation (see the proof of Proposition 1) and thus easy to solve numerically.

The next result, which characterizes SPNiN wholesale price vectors, is used in the estimation of bargaining and manufacturer marginal cost parameters, which is tackled in section 4.3.

Proposition 2. Let $\mathcal{J}_{h,B} := \mathcal{J}_h \setminus \cup_s \mathcal{J}_{PL}^s$ be the set of branded products sold by chain h. Suppose $w_h \in \mathbb{R}^{|\mathcal{J}_{h,B}|}$ is a subgame perfect Nash-in-Nash equilibrium wholesale price vector. Let $c_h \in \mathbb{R}^{|\mathcal{J}_{h,B}|}$ be the vector of manufacturers' marginal costs of producing the goods sold by chain h and let $m(j)$

be the manufacturer of product j. Then the vector of wholesale markups, $w_h - c_h$, satisfies

$$\left(\sum_{s=1}^{S_h} \Omega^s(w) + \Lambda^s(w)\right)(w_h - c_h) = -\sum_{s=1}^{S_h} \mathbb{E}_\eta[D^{s,h}(p^s(\tilde{c}(w)))]$$

where $D^{s,h}(p) \in \mathbb{R}^{|\mathcal{J}_{h,B}|}$

$$D_j^{s,h}(p) = \begin{cases} D_j^s(p) & \text{if } j \in \mathcal{J}^s \\ 0 & \text{otherwise} \end{cases}$$

,the matrices $\Omega^s(w), \Lambda^s(w) \in \mathbb{R}^{|\mathcal{J}_{h,B}| \times |\mathcal{J}_{h,B}|}$ are given by

$$\Omega^s(w)_{j,k} = \begin{cases} \mathbb{E}_\eta\left[\nabla D_k^s(p^s(\tilde{c}^s(w)))' \frac{\partial p^s}{\partial c_j}(\tilde{c}^s(w))\right] & \text{if } j \in \mathcal{J}^s, k \in \mathcal{J}^s \cap \mathcal{J}_{h,m(j)} \\ 0 & \text{otherwise} \end{cases}$$

and

$$\Lambda^s(w)_{j,k} = \begin{cases} -\frac{b_{h,m}}{b_{m,h} S_h(w)} \left(\sum_{s=1}^{S_h} \mathbb{E}_\eta[D_j^{s,h}(p^s(\tilde{c}(w)))]\right) \mathbb{E}_\eta[D_k^s(p^s(\tilde{c}^s(w)))] & \text{if } k \in \mathcal{J}^s \cap \mathcal{J}_{h,m(j)} \\ 0 & \text{otherwise} \end{cases}$$

and $S_h(w) = V_h(w; \mathcal{J}_h) - V_h(w_{-m}; \mathcal{J}_h \setminus \mathcal{J}_{h,m})$

Proof. See appendix A. ♠

The usefulness of Proposition 2 stems from the fact that it allows me to solve for manufacturers' marginal costs. This, together with a model for those costs, allows me to estimate bargaining and manufacturer marginal cost parameters without solving the model. See section 4.3 for details.

3.4 Comments about the model

The model introduced above does not contain a theory of which manufacturer-retailer relationships occur in equilibrium. The model above assumes, as does almost all of the empirical bargaining literature, that these relationships are exogenously given. The first attempts at relaxing this assumption are very recent - see Ho and Lee (2017a) and I don't incorporate endogenous supply chain relationships as they're not the focus of this paper. The model also takes as given the product variety at each store. It is possible to imagine a model in which the store's profit maximization problem is both over which products to offer - choosing a subset of the products procured by the chain - and prices. For the choice of products to be non-trivial, a constraint - arising, for example, from finite physical space - must be imposed, otherwise the solution with respect to the product variety is to offer all available products. Since I don't have data to estimate a model of optimal product variety, I take the product offerings at each store as exogenous.

I also assume, in line with the empirical bargaining literature, that if disagreement between manufacturer m and chain h occurs, w_{-m} is held fixed. A perhaps natural alternative assumption is that the wholesale prices that occur under disagreement are themselves the outcome of a Nash-in-Nash bargaining game. Computing equilibria for such a model would require the calculation of a large number of Nash-in-Nash equilibria, which might not be computationally feasible. That being said, Proposition 2 goes through without change if the disagreement payoff for the chain is independent of $\hat{w}_m$, which would be true in the alternative model just suggested.

It is also worthwhile to compare the model introduced above to the model in Draganska et al. (2010). There are several differences. First, my model allows for retailer profit maximization, though under the simplifying assumption of no substitution across stores. Draganska et al. (2010) argue that a model of independent negotiations (as my model and theirs) is inconsistent with optimal pricing by retailers, which is true in a model in which substitution across retailers is allowed for, like theirs. Retailer profit maximization can, however, be accomodated in a model in which stores are local monopolists: in that case an incrase in w_{jh} only generates retail price changes for retailer h; there are no equilibrium effects on the prices of other retailers, and thus no change in the manufacturer's value of other negotiations.

Draganska et al. (2010) also assume that manufacturers and retailers bargain separately over each product. Thus, disagreement payoffs in their model involve the elimination of only one product. I assume, instead, that manufacturers and retailers negotiate over the wholesale prices of *all* the goods produced by that manufacturer. I believe this is a more realistic assumption. This assumption also captures the intuition that if a retailer relies heavily on a given manufacturer, that manufacturer has considerable leverage over the retailer and can negotiate larger wholesale prices, a mechanism that is not present in their model (except at the product level). Another difference is that they implicitly assume that negotiations over wholesale prices occur every period (weekly). In my empirical implementation, I instead assume that negotiations over wholesale prices occur once and those wholesale prices are fixed for the duration of the data (1 year). Details are given in the next section.

Finally, Draganska et al. (2010) don't have a theory for the joint determination of wholesale prices[20]. This precludes computation of wholesale prices under alternative scenarios - their counterfactuals analysis is based on a parametrization of bargaining parameters. The model presented here, instead, is able to generate counterfactual wholesale and retail prices, by adopting the Nash-in-Nash solution concept. I should note, however, that even though their model is not of the Nash-in-Nash type, there are similarities in the econometric methodology of both papers.

4 Econometrics

In this section I present the details of how the model is estimated. I start in subsection 4.1 with demand estimation; subsection 4.2 gives the details of how wholesale prices are constructed; subsection 4.3 goes over the estimation of bargaining parameters and manufacturers' cost parameters.

[20]In particular, their model is not of the Nash-in-Nash type.

4.1 Demand Estimation

The conditional indirect utility that consumer i derives from product j when buying it at store s was assumed to be given by

$$u_{ijs} = \gamma_{jhl} + \phi_s + \alpha_{h,l} p_{js} + \psi_{h,l} a_{js} + \xi_{js} + \varepsilon_{ijs}$$

where the ε_{ijs} follow independent Type 1 Extreme Value distributions[21]. To this logit model of demand, the simplest instance of the inversion of S. T. Berry (1994) can be applied, yielding

$$\ln(\sigma_j^s) - \ln(\sigma_0^s) = \gamma_{jhl} + \phi_s + \alpha_{h,l} p_{js} + \psi_{h,l} a_{js} + \xi_{js} \tag{6}$$

As usual, endogeneity of p_{js} is a concern: if the retailer or the manufacturer has some information about ξ_{js} then prices will be correlated with ξ_{js}. I'd thus like to have an instrument for prices. Because of the nature of the questions tackled in this paper, which all ask how chain characteristics explain price variation, I'm interested in estimating demand at the chain level. For this reason, I won't aggregate the data to, say, the geographic location level.

Using the data at the retailer level introduces difficulties in the demand estimation. The reason is that standard instruments - for example, the average price of the product in other markets[22], see Nevo (2001) and Hausman (1996) - are not powerful to explain *within-store* price variation[23]. The reason is that within a store, prices tend to decrease in sale periods and then go back to their "normal level" - a pattern illustrated by figure 4 - but the variation in the price of the product in other markets comes from upstream cost shocks, which are common across stores and retailers.

A common idea for generating instrumental variables for prices is find-

[21] See 3.2 for the definitions of the other terms.

[22] In my context, this would have to be adapted to the average price of the product in other markets and in other chains, because wholesale prices are common across stores that belong to the same chain.

[23] Aggregating across stores within a chain wouldn't help much, because prices within a chain are highly correlated, as discussed above. See also Adams and Williams (2017) and DellaVigna and Gentzkow (2017).

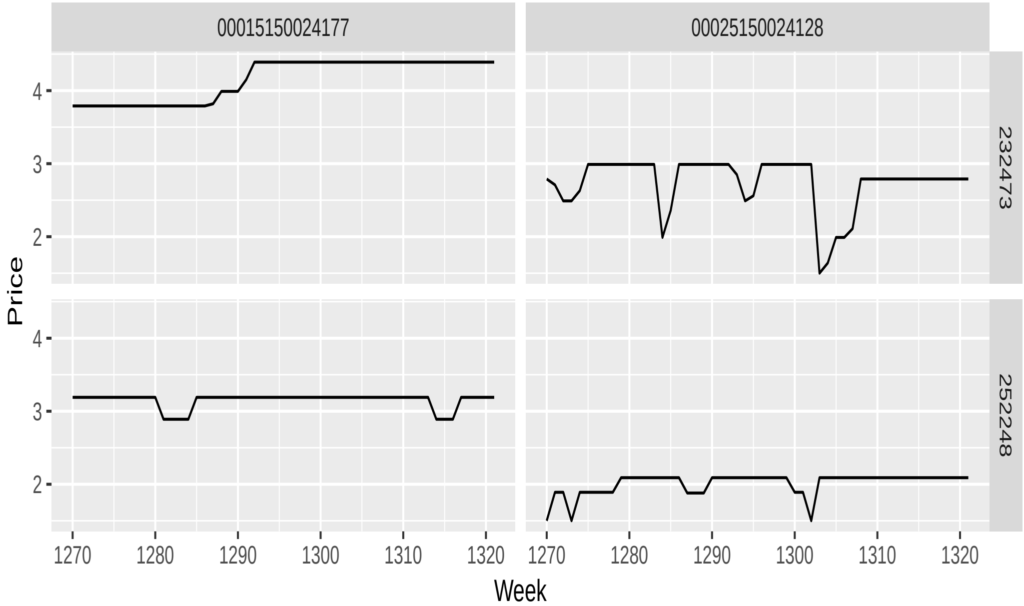

Figure 4: Price paths for two products (columns) in two different stores (rows).

ing exogenous cost shifters. The model introduced above makes cost shocks to the stores explicit. Cost shocks also enter store prices explicitly: stores set their prices equal to $p^s(\tilde{c}^s(w_h, \eta^s))$. Thus, if stores' cost shocks η^s were observed, they'd be an ideal instrument. We can use this intuition to construct a GMM estimator based on the restriction that stores' cost shocks and unobservable demand factors at the store level are uncorrelated, i.e.,

$$\mathbb{E}[\eta_j^s \xi_{js}] = 0$$

Specifically, estimation proceeds as follows. First, I run $\ln(\sigma_j^s) - \ln(\sigma_0^s)$ and p_{js} on product, store and advertisement dummies (see equation 6). Let the resulting residuals be denoted $\hat{\nu}_{sj}$ and $\hat{\rho}_{sj}$, respectively. For a fixed value of α, I invert the store's first order conditions to obtain marginal costs c_j^s - note that this inversion depends only on α and the observed market shares. Then, leveraging the marginal cost model (3), I run the marginal costs implied by α on product and store dummies. Let the resulting residuals be denoted $\hat{\eta}_{sj}$. The GMM estimate of α is the solution to the equation

$$\bar{g}(\alpha) := N^{-1} \sum_{s,j,t} (\hat{\nu}_{sjt} - \alpha \hat{\rho}_{sjt}) \hat{\eta}_{sjt}(\alpha) = 0 \tag{7}$$

where N is the relevant number of observations. Note that this approach is reminiscent of S. Berry, Levinsohn, and Pakes (1995) in that it imposes the equilibrium pricing equations when estimating demand. Standard errors for this estimator can be computed using standard results for the asymptotic distribution of extremum estimators, e.g. Newey and McFadden (1994). Once α has been estimated, the remaining demand coefficients can be obtained by running a OLS regression of $\ln(\sigma_j^s) - \ln(\sigma_0^s) - \hat{\alpha} p_{js}$ on the product, store and advertisement dummies.

I estimate demand separately for each (chain, location) pair in the data, for a total of 206 demands, using OLS and the GMM estimator above. Under the OLS estimator, 4.23% of the own price elasticities are less than one in absolute value. For the GMM estimator, this figure is 0.03%. The median of the distribution of own price elasticities is -2.83 for the OLS estimator and -5.02 for the GMM estimator. Because I'm modelling the demand for peanut butter within a store, the GMM results seem to be more credible than the OLS results. The larger elasticies under the GMM estimator translate into patterns for marginal costs via the inversion of stores' first order conditions. Under the OLS estimator, 13.7% of the marginal costs are estimated to be negative, whereas that figure is equal to 1.24% for the GMM estimator. Of course, these results are the consequence of larger (in absolute value) coefficients obtained with the GMM estimator. Figure 5 plots the distribution of the estimates obtained under each method.

4.2 Construction of Wholesale Prices

Stores choose prices to solve

$$\max_p \sum_{j \in \mathcal{J}^s} (p_j - c_j^s) D_j^s(p)$$

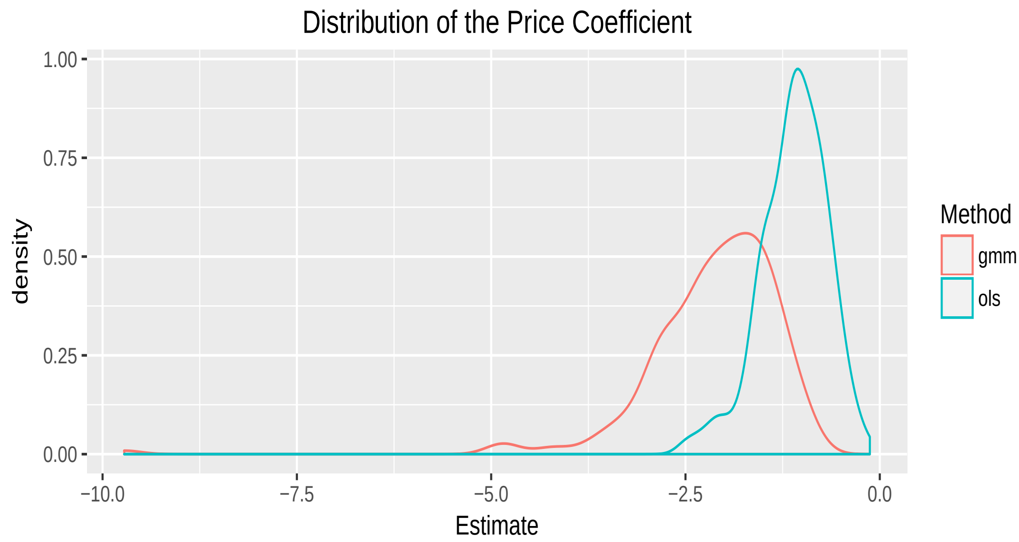

Figure 5: Distribution of price coefficient estimates under OLS and GMM.

The first order conditions can be solved for stores' marginal costs:

$$c^s = p^s + (J_{\sigma^s}(p^s)')^{-1}\sigma^s(p^s) \tag{8}$$

where σ^s is the share function for store s, given by

$$\sigma_j^s(p) = \frac{\exp(\delta_j^s)}{1 + \sum_k \exp(\delta_j^s)}$$

and $J_{\sigma^s}(p)$ is its Jacobian with respect to prices.

From equation (8) I can thus obtain the marginal costs as a function of estimated demand parameters and data[24]. Having recovered stores' marginal costs, I estimate the store marginal cost model (3). Specifically, for each chain I run the stores' marginal costs obtained via equation (8) on product and store dummies. The coefficient estimates on the product

[24] For the logit model things are even simpler. For the share function above we have

$$\frac{\partial \sigma_j^s}{\partial p_k} = \begin{cases} \alpha\sigma_j^s(p)(1 - \sigma_j^s(p)) & \text{if } k = j \\ -\alpha\sigma_j^s(p)\sigma_k^s(p) \end{cases}$$

Marginal costs can thus be recovered as a function of α and data - the remaining demand parameters are not needed.

dummies are the estimated wholesale prices. For the purpose of estimating the bargaining part of the model, these wholesale prices are treated as data. I obtain 3250 observations of wholesale prices at the (Chain,UPC) level.

The following tables provide some reduced form analysis of wholesale prices obtained following the procedure above and their relation to some characteristics of retailers' product portfolios. First, table 3 shows that retailers are indeed able to procure private label products at cheaper prices. The table shows regressions of wholesale prices on a private label dummy. The second column includes chain fixed effects, resulting in no change on the estimated coefficient. The third column controls for weight, and the estimated coefficient now is substantially larger in absolute value, which shows that private label products tend to be heavier than branded products. The result in column (3) shows that private label products are on average 33.23% cheaper than a branded product of the same weight[25].

Table 3: Do Stores Face Lower Marginal Costs for Store Brand Products?

	Dependent variable:		
	$\ln(w_{jh})$	$\ln(w_{jh})$	$\ln(w_{jh})$
	(1)	(2)	(3)
Weight			0.031^{***}
			(0.0004)
Private Label Dummy	-0.241^{***}	-0.241^{***}	-0.404^{***}
	(0.026)	(0.026)	(0.014)
Chain FE	No	Yes	Yes
Observations	3,250	3,250	3,250
R^2	0.026	0.100	0.737
Adjusted R^2	0.025	0.076	0.729
Residual Std. Error	0.601 (df = 3248)	0.586 (df = 3163)	0.317 (df = 3162)
Note:			*p<0.1; **p<0.05; ***p<0.01

[25] $1 - e^{-0.404} = 0.3323$.

The value of private labels for retailers stems from the lower cost of procuring these goods. Assuming they're close substitutes for national brands, retailers are able to obtain a larger mark-up on private labels while pricing these products below the branded products, thus diverting demand towards private labels. Alternatively, under constant mark-ups - as implied by a logit model of demand as the one used in this paper[26] - and assuming consumers perceive private label and branded products to be of similar quality, the retailer obtains greater demand for the private label product, thus generating larger profits from private label products than their national brand counterparts. This is also the mechanism that makes private labels relevant for the determination of equilibrium wholesale prices. The presence of private label products increases retailers' disagreement payoffs in the Nash bargaining problems, which tends to reduce wholesale prices. This mechanism works for all products/manufacturers, not only private labels, but the discussion above suggests it's most relevant for private label products. For the quantitication of these mechanisms in counterfactual analyses, see section 5.

Table 4 shows regressions of the logarithm of wholesale prices of branded products on different measures of the prevalence of private label products in each chain. The first column uses the share of products sold in the year that were private label products and the second column uses the share of private label products in the retailer's portfolio. Both regressions include product fixed effects. The coefficient of interest in the first regression is positive. This might suggest that the prevalence of store brand products brings retailers no benefit in terms of better wholesale prices, but it can be interpreted in the reverse direction: high wholesale prices of branded products will lead to high retail prices, which in turn will imply higher share of private label products. The second regression, on the other hand, shows that retailers that carry a larger share of private label products in their portfolios tend to obtain better wholesale prices. The measure of private label prevalence used in the second regression doesn't capture consumers' optimal purchase decisions. This is a double edged sword: on the

[26]See Proposition 1.

Table 4: Private Label Produts and Wholesale Prices

	Dependent variable:	
	$\ln(w_{jh})$	$\ln(w_{jh})$
	(1)	(2)
Share of PL - Sales	0.155^{***}	
	(0.028)	
Share of PL - Portfolio		-0.225^{***}
		(0.051)
Product FE	Yes	Yes
Observations	2,580	2,580
R^2	0.905	0.905
Adjusted R^2	0.899	0.899
Residual Std. Error (df = 2440)	0.175	0.176

Note: *p<0.1; **p<0.05; ***p<0.01
Estimation sample restricted to branded products.

one hand it shuts down the reverse causality present in the first regression, but it also fails to capture the perceived quality of private label products among consumers. Counterfactuals exercises conducted in section 5 account for both of these factors.

Finally, table 5 shows how the logarithm of wholesale prices correlates with the number of products sold by a retailer and the number of manufacturers the retailer negotiates with. The estimated coefficients are small. The coefficient on the number of suppliers is positive, which suggests that the increase in retailers' disagreement payoffs coming from the larger number of suppliers is not the dominating factor. Instead, the strategic complementarity between wholesale prices of different manufacturers seems to be the relevant factor, which might suggest that manufacturers have more bargaining power in the vertical relationships studied here.

Table 5: Products, Suppliers and Wholesale Prices

	Dependent variable:		
	$\ln(w_{jh})$	$\ln(w_{jh})$	$\ln(w_{jh})$
	(1)	(2)	(3)
Number of Products	-0.002^{***}		-0.003^{***}
	(0.0004)		(0.0005)
Number of Suppliers		0.016^{***}	0.023^{***}
		(0.003)	(0.003)
Product FE	Yes	Yes	Yes
Observations	2,580	2,580	2,580
R^2	0.904	0.905	0.907
Adjusted R^2	0.899	0.900	0.902
Residual Std. Error	0.176 (df = 2440)	0.175 (df = 2440)	0.174 (df = 2439)

Note: *p<0.1; **p<0.05; ***p<0.01
Estimation sample restricted to branded products.

4.3 Estimation of Manufacturer Cost and Bargaining Parameters

The model introduced above allows bargaining parameters to vary both across manufacturers and retailers. When estimating the model, I'll allow them to vary across manufacturers only. Moreover, there are three large manufacturers in the data[27] and many small manufacturers. I'll assume that all the small manufacturers have the same bargaining parameter. In short, there are 4 bargaining parameters to be estimated.

Proposition 2 says that SPNiN wholesale prices for chain h, w_h, satisfy

$$\left(\sum_{s=1}^{S_h} \Omega^s(w) + \Lambda^s(w)\right)(w_h - c_h) = -\sum_{s=1}^{S_h} \mathbb{E}_\eta[D^{s,h}(p^s(\tilde{c}(w)))] \tag{9}$$

Therefore, as long as the matrix on the left hand side of this equation is in-

[27]Conagra Foods Inc, Unilever and J M Smucker Co.

vertible, I can solve for c_h, the manufacturers' marginal costs of producing the goods sold by chain h.

It also turns out that the matrix on the left hand side of equation (9) is block-diagonal, where the blocks correspond to the different manufacturers that negotiate with chain h. This implies that the resulting marginal costs depend only on the bargaining power of that manufacturer. Therefore, I may write

$$c_{jh} = w_{jh} - \mu(w_h, b_{m(j)}), \quad j \in \mathcal{J} \setminus \mathcal{J}_{h,PL}, h = 1, \ldots, H \tag{10}$$

where μ is the mark-up term obtained by solving the system of linear equations given by equation (9), $m(j)$ is the manufacturer that produces good j and $\mathcal{J}_{h,PL}$ is the set of private label products sold by chain h.

Now suppose manufacturers' marginal costs are given by

$$c_{jh} = x'_{jh}\gamma + \nu_{jh} \tag{11}$$

where x_{jh} are observable characteristics of the product and the chain and ν_{jh} is a shock. Specifically, I include in x_{jh} manufacturer fixed effects, chain fixed effects, the product's weight, a dummy for reduced sugar products, a dummy for more expensive production processes[28], a dummy for reduced sodium, a dummy for chunky or crunchy peanut butter and a dummy for flavored products.

Putting equations (10) and (11) together, I can write

$$\nu_{jh}(b_{m(j)}, \gamma, w_h, x_{jh}) = w_{jh} - \mu(w_h, b_{m(j)}) - x'_{jh}\gamma \tag{12}$$

Estimation of manufacturers' marginal cost parameters (γ) and bargaining parameters (b_m) can then be accomplished by GMM, based on the conditional moment restriction

$$\mathbb{E}[\nu_{jh}|z_{jh}] = 0$$

for appropriate variables z_{jh}. Note that in equation (12), the marginal cost

[28]For example, natural, kosher and organic products.

shocks ν_{jh} are written as a function of data and structural parameters. However, the expression in the right hand side of that equation depends on w_h, which in turn depends on ν_{jh}. Therefore, to identify the bargaining parameters b_m, instruments generating exogenous variation in w_h are necessary.

4.3.1 Choice of Instruments and Identification

I use as as instruments z_{jh} the following variables:

(i) the cost covariates x_{jh}.

(ii) Product j's expected demand at chain h, under the optimal retail prices implied by the average wholesale price of product j in *other* chains, interacted with manufacturer dummies.

(iii) Total expected demand for goods produced by the manufacturer of product j, under the optimal retail prices implied by the average wholesale price of product j in other chains, interacted with manufacturer dummies.

(iv) Manufacturer mark-up for product j, as obtained from Proposition 2, computed under the average wholesale prices in other chains (for all the products sold by chain h) and assuming $b_m = 1/2$, interacted with manufacturer dummies.

The cost covariates are all assumed to be exogenous with respect to ν_{jh}. Note that the instruments in bullets (ii) to (iv) all use the average of the wholesale prices in other chains, which depend on $\nu_{jh'}$, $h' \neq h$. Therefore, the main assumption underlying validity of these instruments, and hence the main identifying assumption, is that $\mathbb{E}[\nu_{jh}\nu_{jh'}] = 0$, i.e., that manufacturers' marginal cost shocks are uncorrelated across chains. This seems like a reasonable assumption given that (i) negotiations occur only once[29] and (ii) the cost model includes a rich set of product characteris-

[29] In a model with repeated negotiations, $\mathbb{E}[\nu_{jht}\nu_{jh't}] = 0$ would hardly be a compelling assumption, but in that situation one might use the panel structure of the data to generate alternative moment conditions.

tics[30] and manufacturer dummies. One situation in which this identifying assumption would not hold is when two chains concentrate their stores in a common location that is hard to access, making c_{jh} larger for those chains compared to all other chains.

Power comes from the fact that variables (ii)-(iv) are based on demand factors. For example, if the customers of chain h perceive product j to be of high quality, demand for that product will tend to be high and the manufacturer, being aware of that, can charge larger wholesale prices. As another example, the variable in (iv) predicts wholesale mark-ups assuming $b_m = 1/2$. This variable is correlated with actual wholesale mark-ups, which mechanically influence wholesale prices.

The discussion above shows that identification of bargaining and cost parameters comes from variation across chains. To build intuition, think of one manufacturer (labeled m) producing one good and negotiating with two retailers, $h = 1, 2$. The customers of retailer 1 dislike the manufacturer's product and the customers of manufacturer 2 really enjoy the product. The wholesale price at which the good will be sold to retailer 1, call it w_1, will tend to be close to the manufacturer's marginal cost c_m. The wholesale price at which the good will be sold to retailer 2, w_2, on the other hand, will tend to be larger. Exactly how much larger will depend on how much the manufacturer is able to capitalize on the fact that the customers of retailer 2 enjoy the product. This is governed by the bargaining parameter b_m.

4.3.2 Estimation and Results

A GMM estimate is the solution to the program

$$\min_{\gamma,b} \bar{g}_n(\gamma, b)' W_n \bar{g}_n(\gamma, b) \tag{13}$$

[30]Suppose I had ommitted, say, the reduced sugar dummy. These products might use more expensive sweeteners as ingredients, and thus in that model marginal cost shocks might be correlated across chains

where $\bar{g}_n(\gamma, b) := n^{-1}\sum_{j,h}\nu_{jh}(b_{m(j)}, \gamma, w_h, x_{jh})z_{jh}$ where z_{jh} is a column vector with the instrumental variables described above[31] and W_n is *some* weight matrix. I implement the standard two step procedure to obtain an optimal GMM. In the first step, I take as weight matrix $W_n = (Z'Z)^{-1}$, where Z is the matrix of instruments. Let $\hat{\theta}_1$ be the resulting estimate for $\theta = (\gamma, b)$. With that estimate, I construct an estimate of the optimal weight matrix

$$\hat{W}_n^* = \left(n^{-1}\sum_i g_{jh}(\hat{\theta}_1)g_{jh}(\hat{\theta}_1)'\right)^{-1}$$

where $g_{jh}(\theta) = \nu_{jh}(b_{m(j)}, \gamma, w_h, x_{jh})z_{jh}$. The second step consists of minimizing (13) again, using $\hat{W}_n^*$ as the weight matrix.

It is important to note that, conditional on a value of the bargaining parameters b, estimation is linear on the cost parameters γ - which can be seen from equation (12). One evaluation of the GMM objective (13), for a fixed value of b, consists of the following steps:

1. Apply Proposition 2 to obtain manufacturers' marginal costs as a function of data and bargaining parameters[32].

2. Obtain the implied estimates for γ, given by

   $$\hat{\gamma} = (X'ZWZ'X)^{-1}X'ZWZ'C(b)$$

   where $C(b)$ is the vector of marginal costs for each (product, chain) pair, obtained in the previous step.

3. Compute $\hat{\nu}_{jh}(b_{m(j)}, \gamma, w_h, x_{jh}) = c_{jh}(w_h, b_{m(j)}) - x'_{jh}\hat{\gamma}$.

4. Compute $\bar{g}_n(\hat{\gamma}, b) = n^{-1}\sum_{j,h}\hat{\nu}_{jh}z_{jh}$ and form the GMM objective in (13).

[31] There are 111 instruments in total, 96 of which are cost covariates. The large number comes from the chain dummies.

[32] To compute the terms in equation 9, I (i) set ξ_{js} to zero and (ii) compute the expectations with respect to η^s by simulation, sampling from the residuals obtained from running the stores' marginal cost regressions.

Because γ can be found in closed form for a given value of b, the non-linear search can be restricted to the bargaining parameters. I minimize the GMM objective using a derivative-free global optimizer[33] with somewhat loose termination parameters. Once the global optimizer has found a solution, that solution is used as the starting point for a derivative-free local optimizer[34], now with tighter termination conditions. After the first estimation step (GMM with $W_n = (Z'Z)^{-1}$, as explained above), I do *not* use that estimate as the starting point for solving the second GMM problem. The idea is that we want to use the first GMM estimate to construct an estimate of the optimal weight matrix, but we don't want to "bias" the solution algorithm towards the previously found estimate[35]. Instead I draw a random vector as the starting point and repeat the two-step solution procedure above. The resulting estimate doesn't depend on the randomly drawn starting points.

I now explain how standard errors are computed, acknowledging that they are *not* correct - an updated version of this paper will include appropriately bootstraped standard errors. The standard errors shown below are directly computed from the standard result for the asymptotic distribution of extremum estimators - of which GMM is a special case -, e.g., Newey and McFadden (1994). There are two reasons why these are incorrect. First, they fail to account for the variance coming from the construction of wholesale prices performed before the estimation of the bargaining parameters. Second, the characterization of the asymptotic distribution of extremum estimators is based on the first order conditions for an interior solution for the estimation program, but the estimate I obtain is not interior. From now on, these two caveats will be ignored.

As mentioned above, bargaining parameters are allowed to vary across

[33]Controlled Random Search, see Kaelo and Ali (2006). I use the implemention in NLOPT. See http://ab-initio.mit.edu/nlopt.

[34]COBYLA, see Powell (1994). I use the implementation in NLOPT. See http://ab-initio.mit.edu/nlopt.

[35]Note that the estimator in the first step is also a consistent estimator of θ and thus the resulting estimate should be a good starting point for the second step. However, I want to allow for non-trivial changes from one step to the other and thus follow the procedure described above. This decision doesn't change the results.

Manufacturer	$\hat{b}_m$	Std. Dev.
Conagra	0.0886	0.2393
Unilever	10^{-5}	0.1138
J M Smucker Co.	0.4489	0.0911
Others	1	0.4367

Table 6: Bargaining parameter estimates

manufacturers, but since there are three large manufacturers and many small ones[36], I assume that all the small manufacturers have the same bargaining parameter. Table 6 shows the results[37]. The estimates show that Conagra and Unilever have little bargaining power in the vertical chain. The estimate for Smucker is close to 1/2 and small manufacturers (labeled "OTHERS" in table 6) seem to have substantial bargaining power in the vertical chain, which might be due to small manufacturers focusing on high-end products.

5 Counterfactual Exercises

In this section I present some counterfactual exercises. In the counterfactuals shown here, I focus on one chain[38]. I choose a chain that negotiates with all manufacturers and sells private label products. Specifically, the chain I consider sells 34 different products: 12 produced by J M Smucker Co., 11 produced by Conagra Foods Inc., 6 produced by Unilever, 4 private label products and 1 produced by one of the small manufacturers. For the purpose of the counterfactual exercises, I assume there are no shocks to stores' marginal costs, for computational convenience[39].

I consider three types of counterfactuals: (i) excluding all the products

[36]Conagra is the smallest of the "large" manufacturers: there are 579 observations at the (UPC, Chain) level where the UPC is manufactured by Conagra. The largest of the "small" manufacturers is Hersheys, with 60 (UPC, Chain) pairs in the data.

[37]10^{-5} was the lower bound imposed on the GMM problem.

[38]Future iterations of this paper will perform the same exercise for all chains.

[39]Computing the expectations in equation 9 needs to be done only once for estimation, which is feasible. As explained in this section, solving for SPNiN equilibria involves evaluating those expectations multiple times, which is costly.

of a given manufacturer (including private label products), (ii) assuming all products of a given manufacturer become private label products and (iii) mergers between two manufacturers. I compute changes in wholesale prices, retail prices and consumer surplus.

I start this section by explaining, in subsection 5.1, how I solve for SPNiN wholesale prices. The subsequent sections report the results.

5.1 Solving for SPNiN Wholesale Prices

In a Nash-in-Nash equilibrium wholesale price vector, every Nash product must be maximized with respect to that manufacturer's wholesale prices, given the wholesale prices of every other manufacturer. This yields a system of first order conditions given by[40]

$$b_m \frac{\partial V_m}{\partial \hat{w}_j}(w) \frac{1}{V_m(w)} + (1-b_m) \frac{\partial V_h}{\partial \hat{w}_j}(w) \frac{1}{S_h(w)} = 0, \quad j = 1, \dots, J$$

Explicitly,

$$b_m \frac{\sum_{s=1}^{S_h} \Omega_j^s(w)(w_h - c_h) + \sum_{s=1}^{S_h} D_j^{s,h}(p^s(\tilde{c}(w)))}{\sum_{s=1}^{S_h} \bar{D}_j^s(p^s(\tilde{c}^s(\hat{w}_m, w_{-m}))) \cdot (w_h - c_h)} = \frac{1-b_m}{S_h(w)} \sum_{s=1}^{S_h} D_j^{s,h}(p^s(\tilde{c}^s(w))) \tag{14}$$

For the definition of the terms in the equation above, see the proof of Proposition 2. In a SPNiN equilibrium wholesale price vector, equation (14) must hold for every product $j \in \mathcal{J}_h \setminus \mathcal{J}_{h,PL}$. I solve this system of equations to obtain *candidates* for Nash-in-Nash equilibria. Suppose I can find all the solutions to this system of equations and suppose there are finitely many. The set of solutions is a superset of the set of Nash-in-Nash equilibria. To find the Nash-in-Nash equilibria, I can then test every solution to the system of equations above. Specifically, for each candidate w_h, solve, for each manufacturer m

$$\max_{\hat{w}_m} \mathcal{N}_m(\hat{w}_m, w_{-m})$$

[40]For the derivation of the equations shown here, and the definition of the terms in these equations, see the proof of Proposition 2.

where $\mathcal{N}_m$ is the Nash product for manufacturer m. Let w'_m denote a solution to this problem. If $\mathcal{N}_m(w'_m, w_{-m}) = \mathcal{N}_m(w_h)$ for every m, then w_h is a Nash-in-Nash equilibrium. This is the procedure I follow to solve for SP-NiN equilibria: first find solutions to the system of FOCs and then verify whether the solutions are indeed equilibria.

The system of equations above is a complex one. In particular, it takes into account how stores change their prices once wholesale prices change - this is part of the Ω_j^s terms - again, see the statement and proof of Proposition 2. Moreover, the logic outlined above requires me to try many different starting points[41], in the hope of finding all solutions to the system of FOCs. To be able to efficiently evaluate the left hand side of equation (14), I show in appendix B how the Implicit Function Theorem allows me to further characterize how stores' optimal prices change in response to changes in wholesale prices[42]. I find a unique solution for the FOCs and this solution is indeed a SPNiN: the maximum value (over manufacturers) of $(\max_{\hat{w}_m} \mathcal{N}_m(\hat{w}_m, w_{-m}) - \mathcal{N}_m(w_h^*))$ is equal to 6.46×10^{-7}, where w_h^* is the solution to the FOCs.

5.2 Wholesale Price Changes

In this subsection I present results on equilibrium wholesale prices. All the tables in this section have the same structure as table 7: entry (i, j) shows the results when it is assumed that all the products of manufac-

[41]Good starting points turn out to be critical to solve this system of as many as 34 equations (the number of equations varies with the counterfactual being considered). For many starting values I don't find solutions. A good set of good starting points is given by

$$w_j = c_j(1 + b_{m(j)}\mu)$$

where $\mu \in [0, 1]$. The logic is that under these prices the margins $(w_j - c_j)/c_j$ are proportional to the bargaining parameters. To solve for the benchmark case (where ownership of products is as in the data), I take 100 draws of a $U[0, 1]$ distribution for the value of μ. See the text for some details on the results.

[42]Note that Proposition 1 subsumes Appendix B. However, the results shown below were obtained using the analysis in Appendix B. A future iteration of this paper will use Proposition 1 instead, as it provides a sharper and computationally more efficient characterization of how stores' optimal prices change in response to changes in wholesale prices.

turer i are produced and sold by manufacturer j instead; the column "Excluded" shows the results under the assumption that the products of manufacturer i cease to be sold by the retailer. In counterfactual (i, j) it is assumed that the bargaining power of the resulting firm is equal to the bargaining power of manufacturer j. Counterfactual (i, j) might thus be more appropriately interpreted as manufacturer j acquiring manufacturer i. Counterfactual (j, i) corresponds to the opposite acquisition; the only difference is the bargaining power of the resulting firm.

Table 7 shows the average (across products) wholesale price percent change, where the average is computed across products that are branded products in both the benchmark and in counterfactual (i, j). Excluding a manufacturer's products increases the wholesale prices of the remaining products, but not by a lot. Assuming branded products become private label products decreases the wholesale prices of the other branded products, but again the effect is small. Mergers, however, can have large effects on equilibrium wholesale prices. What drives these effects is the difference in bargaining power between firms i and j. Table 8 shows the average wholesale price percent change, where the average is computed across all products, including private label products. The results involving private label products are substantially different in tables 7 and 8 because when branded products become private label products the retailer is able to procure these products at cost. Conversely, when private label products are sold by one of the upstream firms, the retailer will pay a mark-up on these products.

Table 7: Wholesale Price Inflation: Products Negotiated in Both Scenarios

	Excluded	Private Label	Conagra	Smucker	Unilever	Others
Private Label	0.096		0.050	0.197	0.0001	0.092
Conagra	0.677	−0.044		7.051	−1.124	11.234
Smucker	0.138	−0.053	−4.357		−6.180	7.242
Unilever	0.153	0	0.639	3.670		6.056
Others	0.003	−0.004	7.231	−0.400	1.556	

Table 8: Wholesale Price Inflation: All Products

	Excluded	Private Label	Conagra	Smucker	Unilever	Others
Private Label	0.084		0.584	7.051	0.0001	23.898
Conagra	0.560	−0.992		6.221	−0.992	9.912
Smucker	0.113	−5.453	−3.845		−5.453	6.390
Unilever	0.131	−0.0001	0.564	3.238		5.343
Others	0.002	−0.829	6.381	−0.353	1.373	

5.3 Retail Price Changes

Table 9 shows the average (across products) percent change in retail prices of products that are branded in both the benchmark and in counterfactual (i, j). As shown above, excluding a manufacturer's products leads to a small increase in the wholesale prices of the remaining products. However, with fewer products, the retailer has an incentive to reduce the price of the remaining products. The reason is that with fewer products it's more likely that marginal consumers will move to the outside option after a price increase. This effect dominates in all cases, except for the exclusion of private labels, when the increase in wholesale price dominates - but the resulting effect on retail prices is small.

Assuming that branded products become private label products leads to increases in the retail prices of the remaining branded products. Technically, the reason is that the retailer's optimal mark-up[43] increases when the costs of some of the goods decrease. Economically, the lower costs of the new private label products allow the retailer to charge larger mark-ups while still retaining many consumers. The retailer thus optimally increases its mark-up and a fraction of consumers shifts towards the new private label products. As shown above, upstream mergers can have positive or negative effects on the wholesale prices of branded products, depending on the bargaining parameters of the manufacturers involved. This results in changes in retail prices that similarly depend on which counterfactual is considered.

Table 10 shows the average percent change in retail prices, where the

[43]Remember Proposition 1.

average is now taken across all products, including private label products. Entries in the private label row now show positive and sometimes large retail price increases, because these results account for the fact that former private label products are more costly for the retailer in the counterfactual. The private label column shows that despite the increase in the retail prices of negotiated products (table 9), when considering all products retail prices decrease, because former branded products are now cheaper for the retailer to procure, which maps into cheaper retail prices for those products. Together with the discussion of table 9, this shows that after branded products become private label products, the retailer increases its mark-up but by an amount that makes the retail prices of the new private label products smaller than in the benchmark; prices of the remaining branded products increase, and the retailer thus diverts part of the demand to private label products.

Table 9: Retail Price Inflation: Negotiated Products

	Excluded	Private Label	Conagra	Smucker	Unilever	Others
Private Label	0.037		−0.040	−0.273	0.0001	−0.485
Conagra	−2.173	0.206		3.388	−0.507	5.911
Smucker	−4.210	2.493	−1.320		−1.855	3.163
Unilever	−0.393	0.00000	0.389	2.270		3.879
Others	−0.027	0.046	5.109	−0.274	1.444	

Table 10: Retail Price Inflation: All Products

	Excluded	Private Label	Conagra	Smucker	Unilever	Others
Private Label	0.028		0.263	4.482	0.0001	16.138
Conagra	−2.452	−0.407		2.771	−0.407	4.950
Smucker	−4.395	−1.264	−0.905		−1.264	2.522
Unilever	−0.429	−0.00003	0.335	1.962		3.375
Others	−0.029	−0.537	4.504	−0.240	1.265	

5.4 Changes in Consumer Surplus

Finally, table 11 shows the percent change in consumer surplus. Excluding products from the retailer's portfolio has a large negative impact on consumers' surplus, because of the diminished variety. The results shown in the other entries of table 11 are the direct results of equilibrium price changes. For example, assuming Smucker's products become private label products leads to large benefits to consumers because of the resulting decrease in the prices of Smucker's products. Similar reasoning holds when, e.g., Conagra's products are assumed to be sold by Unilever: because Unilever's bargaining power is smaller than Conagra's, the resulting equilibrium wholesale prices of Conagra's products will be lower than in the benchmark, which in turn will result in lower counterfactual retail prices for these products and larger consumer welfare. Interestingly, because mergers can have both positive and negative effects on prices depending on which is the acquiring firm, some mergers are beneficial to consumers (e.g., when Unilever acquires Conagra or Smucker) and some mergers can decrease consumer surplus.

Table 11: Changes in Consumer Surplus

	Excluded	Private Label	Conagra	Smucker	Unilever	Others
Private Label	−0.227		−0.575	−3.155	0.0002	−4.266
Conagra	−22.401	1.923		−10.868	1.922	−13.148
Smucker	−33.498	17.903	12.703		17.902	−14.063
Unilever	−3.834	0.00004	−0.407	−2.070		−2.419
Others	−0.226	0.387	−0.225	0.083	−0.650	

6 Conclusion

This paper studies vertical relationships between retailers and manufacturers. Specifically, I study what are the effects of private label products and the upstream market structure on equilibrium wholesale prices, retail prices and consumer welfare.

To address these questions, I develop and estimate a model of wholesale price negotiation between retail chains, which consist of multiple stores that are local monopolists, and multiple manufacturers. Estimating demand at the retailer level proves to be challenging. I explain how these challenges can be overcome using stores' pricing equations and the assumption that stores' marginal cost shocks and demand unobservables are uncorrelated. Moreover, I show that the bargaining model has an inversion. This allows me to estimate bargaining and manufacturer marginal cost parameters via GMM without solving the model, extending the methodology of Gowrisankaran et al. (2015) to the case with downstream profit maximization.

I use the estimated model to conduct three types of counterfactual exercises: (i) excluding all the products of a given manufacturer (including private label products), (ii) assuming all products of a given manufacturer become private label products and (iii) mergers between two manufacturers. Indirect equilibrium effects are small. For example, eliminating private label products induces only a small increase in equilibrium wholesale prices of branded products. Assuming all products of a given manufacturer become private label products also induces only a small decrease in the wholesale prices of products of other manufacturers.

Nevertheless, wholesale and retail prices of goods produced by one of the manufacturers involved in the counterfactual can change substantially. These prices can increase or decrease, depending on the manufacturers considered in the counterfactual. As a consequence, mergers can be harmful or beneficial to consumers. Assuming products of a given manufacturer become private label products can generate large increases in consumer surplus: even though the indirect effects on the wholesale prices of other manufacturers is small, consumers can benefit greatly from the lower prices of the new private label products.

In summary, private label products do not seem to greatly improve the retailer's bargaining position vis-à-vis manufacturers. In terms of the upstream market structure, the number of retailers doesn't seem to be relevant by itself. Rather, what matters is the nature of the vertical relationships with different manufacturers, as captured by the bargaining parameters.

Appendix

Appendix A Proofs and Auxiliary Results

Proof of proposition 1

Proof. The first order condition of the profit maximization problem with respect to p_j is

$$\sigma_j(p) + \sum_k (p_k - c_k) \frac{\partial \sigma_k}{\partial p_j}(p) = 0$$

Using the logit functional form this becomes (and dropping the argument p)

$$\sigma_j + \alpha \sigma_j (1 - \sigma_j)(p_j - c_j) - \alpha \sum_{k \neq j} (p_k - c_k) \sigma_j \sigma_k = 0$$

or, since $\sigma_j(p) > 0$ for all j and all p,

$$1 + \alpha (1 - \sigma_j)(p_j - c_j) - \alpha \sum_{k \neq j} (p_k - c_k) \sigma_k = 0$$

Now note that $(1 - \sigma_j) = \sigma_0 + \sum_{k \neq j} \sigma_k$, so that the previous equation can be rewritten as

$$\begin{aligned}
0 &= 1 + \alpha \sigma_0 (p_j - c_j) + \alpha \left(\sum_{k \neq j} \sigma_k (p_j - c_j) - \sum_{k \neq j} \sigma_k (p_k - c_k) \right) \\
&= 1 + \alpha \sigma_0 \mu_j + \alpha \sum_{k \neq j} \sigma_k (\mu_j - \mu_k) \\
&= 1 + \alpha \sigma_0 \mu_j + \alpha \sum_{k=1}^{J} \sigma_k (\mu_j - \mu_k) \qquad (15)
\end{aligned}$$

where $\mu_k := p_k - c_k$. In a solution to the problem, this equation must hold for all $j = 1, \ldots, J$. Take two arbitrary products j_1, j_2 and subtract the corresponding equations above to obtain

$$\begin{aligned}
0 &= \alpha\sigma_0(\mu_{j_1} - \mu_{j_2}) + \alpha \sum_{k=1}^{J} \sigma_k(\mu_{j_1} - \mu_{j_2}) \\
&= \alpha\sigma_0(\mu_{j_1} - \mu_{j_2}) + \alpha(1-\sigma_0)(\mu_{j_1} - \mu_{j_2}) \\
&= \alpha(\mu_{j_1} - \mu_{j_2})
\end{aligned}$$

Since $\alpha < 0$ and j_1, j_2 are arbitrary, this proves the constant mark-up property (ii). The problem thus reduces to finding the optimal mark-up μ, i.e.,

$$\max_{\mu} \mu \sum_k \sigma_k(c+\mu) = \mu(1 - \sigma_0(c+\mu))$$

where c is the vector of marginal costs and $c + \mu$ means that μ is added to every coordinate of c. The first order condition of this problem is

$$(1-\sigma_0) - \mu\sigma_0'(c+\mu) = 0$$

Noting that $\sigma_0'(c+\mu) = -\alpha\sigma_0 \sum_k \sigma_k = -\alpha\sigma_0(1-\sigma_0)$, this reduces to[44]

$$\phi(\mu) := (1 - \sigma_0(c+\mu))(1 + \alpha\mu\sigma_0(c+\mu)) = 0$$

Note that $(1 - \sigma_0(c+\mu)) > 0$ for all μ. Define $\psi(\mu) := 1 + \alpha\sigma_0(c+\mu)\mu$. Note that $\psi(0) = 1$ and that $\lim_{\mu\to\infty} \psi(\mu) = -\infty$, because $\alpha < 0$. By the Intermediate Value Theorem, a solution to $\phi(\mu) = 0$ exists. Moreover,

$$\psi'(\mu) = \alpha[\sigma_0'(c+\mu)\mu + \sigma_0(c+\mu)] = \alpha\sigma_0(c+\mu)[1 - \alpha\mu(1-\sigma_0(c+\mu))] < 0$$

for all $\mu \geq 0$. It follows that there's a unique μ^* such that $\phi(\mu^*) = 0$ and this μ^* maximizes the monopolist's profit. This proves parts (i) and (iii).

[44]Note that imposing the constant mark-up property on equation (15) also yields $1 + \alpha\mu\sigma_0 = 0$. The argument would then deliver the existence of a unique μ^* satisfying the first order conditions of the original problem, but would not imply that the resulting price vector is indeed a solution for that problem. It would then be necessary to establish the validity of a second order condition. The argument given here, which reduces the profit maximization problem to a unidimensional problem, yields existence and uniqueness.

It remains to prove (iv). So far I have shown that the optimal prices satisfy $p_j^* = c_j + \mu^*(c)$, where I make explicit the dependence of μ on the parameter c. It follows that

$$\frac{\partial p_j^*}{\partial c_k} = \begin{cases} 1 + \frac{\partial \mu}{\partial c_j}(c) & \text{if } k = j \\ \frac{\partial \mu}{\partial c_k}(c) & k \neq j \end{cases} \tag{16}$$

As shown above, $\mu^*(c)$ is the unique solution to

$$f(c, \mu) := 1 + \alpha\mu\sigma_0(c + \mu) = 0$$

Since $\frac{\partial f}{\partial \mu}(c, \mu) = \alpha\sigma_0(c+\mu)[1 - \alpha\mu(1 - \sigma_0(c+\mu))] < 0$, the Implicit Function Theorem implies that $\mu^*(c)$ is C^1 and that

$$\frac{\partial \mu}{\partial c_k}(c) = -\frac{\partial f}{\partial c_k}(c, \mu) \Big/ \frac{\partial f}{\partial \mu}(c, \mu)$$

Noting that $\frac{\partial f}{\partial c_k}(c, \mu) = -\alpha^2\mu\sigma_k(c + \mu)\sigma_0(c + \mu)$, I obtain

$$\frac{\partial \mu}{\partial c_k}(c) = \frac{\alpha\mu\sigma_k(c + \mu)}{1 - \alpha\mu(1 - \sigma_0(c + \mu))}$$

as desired. ♠

I now establish an auxiliary result used in the proof of Proposition 2.

Lemma 1. Let the value for a store under wholesale prices w be given by

$$V^s(w; \mathcal{J}^s) = \mathbb{E}_\eta\left[\sum_{j \in \mathcal{J}^s} (p_j^s(\tilde{c}^s(w)) - \tilde{c}_j^s(w_j))D_j^s(p^s(\tilde{c}^s(w)))\right]$$

Then

$$\frac{\partial V^s}{\partial w_j}(w) = -\mathbb{E}_\eta\left[D_j^s(p^s(\tilde{c}^s))\right]$$

Proof. For a given marginal cost vector $\tilde{c}^s$, let the value of the store's profit maximization problem be $\pi^s(p^s(\tilde{c}^s); \tilde{c}^s)$ - where $p^s(\tilde{c}^s)$ is the solution to that problem.

Exchanging the order of differentiation and integration yields

$$\frac{\partial V^s}{\partial w_j}(w) = \mathbb{E}_\eta \left[\frac{\partial}{\partial w_j} \sum_{j \in \mathcal{J}^s} (p_j^s(\tilde{c}^s(w)) - \tilde{c}_j^s(w_j)) D_j^s(p^s(\tilde{c}^s(w))) \right]$$

The derivative inside the expectation operator is $\frac{\partial}{\partial w_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s)$. Because $\tilde{c}_j^s = w_j + \tau^s + \eta_j^s$, $\frac{\partial}{\partial w_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s) = \frac{\mathrm{d}}{\mathrm{d}c_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s)$, where $\frac{\mathrm{d}}{\mathrm{d}c_j}$ denotes the total derivative with respect to c_j. By the Envelope Theorem,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}c_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s) &= \frac{\partial}{\partial c_j}\pi^s(p^s(\tilde{c}^s);\tilde{c}^s) \\ &= -D_j^s(p^s(\tilde{c}^s(w))) \end{aligned}$$

Taking the expectation with respect to η yields the result. ♠

An immediate corollary to Lemma 1 is given below

Corollary 1.

$$\frac{\partial V_h}{\partial w_j}(w) = -\mathbb{E}_\eta \left[\sum_{j=1}^{S^h} D_j^{s,h}(p^s(\tilde{c}^s)) \right]$$

where $D_j^{s,h}$ is equal to D_j^s if $j \in \mathcal{J}^s$ and zero otherwise.

Proof. It follows from Lemma 1 and the definition of V_h. ♠

Corollary 1 is used in the proof of Proposition 2 below.

Proof of Proposition 2 .

Proof. Consider the maximization problem in (4). After taking the logarithm of the objective function, the first order condition with respect to $\hat{w}_j$ is given by

$$b_{m,h}\frac{\partial V_m}{\partial \hat{w}_j}(w)\frac{1}{V_m(w)} + b_{h,m}\frac{\partial V_h}{\partial \hat{w}_j}(w)\frac{1}{S_h(w)} = 0$$

where I write $V_h(w)$ instead of $V_h(w;\mathcal{J}_h)$ to simplify notation and m is the manufacturer that produces product j (denoted $m(j)$ below). Rearranging,

$$\frac{\partial V_m}{\partial \hat{w}_j}(w) = -\frac{b_{h,m}}{b_{m,h}\ S_h(w)}\ \frac{\partial V_h}{\partial \hat{w}_j}(w)\ V_m(w) \tag{17}$$

Now note that

$$\begin{aligned}\frac{\partial V_m}{\partial \hat{w}_j}(w) = &\sum_{\{s:j\in\mathcal{J}^s\}}\ \sum_{k\in\mathcal{J}^s\cap\mathcal{J}_{h,m(j)}} (w_k - c_k^m)\mathbb{E}_\eta\left[\nabla D_k^s(p^s(\tilde{c}^s(w)))'\frac{\partial p^s}{\partial c_j}(\tilde{c}^s(w))\right] \\ &+ \sum_{\{s:j\in\mathcal{J}^s\}} \mathbb{E}_\eta[D_j^s(p^s(\tilde{c}(w)))]\end{aligned}$$

where

$$\nabla D_k^s(p)' = \begin{pmatrix}\frac{\partial D_k^s}{\partial p_1}(p) & \dots & \frac{\partial D_k^s}{\partial p_{J^s}}(p)\end{pmatrix}\ ,\ \frac{\partial p^s}{\partial c_j}(c) = \begin{pmatrix}\frac{\partial p_1^s}{\partial c_j}(c) & \dots & \frac{\partial p_{J^s}^s}{\partial c_j}(c)\end{pmatrix}'$$

Plugging this and the definition of $V_m(w)$ into equation (17) yields

$$\begin{aligned}\sum_{s=1}^{S_h}\Omega_j^s(w)(w_h - c_h) + \sum_{s=1}^{S_h}\mathbb{E}_\eta[D_j^{s,h}(p^s(\tilde{c}(w)))] = &-\frac{b_{h,m}}{b_{m,h}\ S_h(w)}\ \frac{\partial V_h}{\partial \hat{w}_j}(w)\times \\ &\times \sum_{s=1}^{S_h}\mathbb{E}_\eta\left[\bar{D}_j^s(p^s(\tilde{c}^s(\hat{w}_m, w_{-m})))\right]\cdot(w_h - c_h)\end{aligned}$$

where $D_j^{s,h}(p)$ is equal to $D_j^s(p)$ if store s sells product j and it is equal to zero otherwise. Moreover, $\Omega_j^s(w), \bar{D}_j^s(p) \in \mathbb{R}^{|\mathcal{J}_{h,B}|}$ are given by (the extra argument in parenthesis denotes the coordinate)

$$\Omega_j^s(w)(k) = \begin{cases}\mathbb{E}_\eta\left[\nabla D_k^s(p^s(\tilde{c}^s(w)))'\frac{\partial p^s}{\partial c_j}(\tilde{c}^s(w))\right] & \text{if } j\in\mathcal{J}^s, k\in\mathcal{J}^s\cap\mathcal{J}_{h,m(j)} \\ 0 & \text{otherwise}\end{cases}$$

and

$$\bar{D}_j^s(p)(k) = \begin{cases}D_k^s(p) & \text{if } k\in\mathcal{J}^s\cap\mathcal{J}_{h,m(j)} \\ 0 & \text{otherwise}\end{cases}$$

Using corollary 1, the last equation can be rewritten as

$$\sum_{s=1}^{S_h} \Omega_j^s(w)(w_h - c_h) + \sum_{s=1}^{S_h} \mathbb{E}_\eta[D_j^{s,h}(p^s(\tilde{c}(w)))] = \frac{b_{h,m}}{b_{m,h}\ S_h(w)} \sum_{s=1}^{S_h} \mathbb{E}_\eta[D_j^{s,h}(p^s(\tilde{c}^s(w)))] \times$$
$$\times \sum_{s=1}^{S_h} \mathbb{E}_\eta\left[\bar{D}_j^s(p^s(\tilde{c}^s(\hat{w}_m, w_{-m})))\right] \cdot (w_h - c_h)$$

This equation holds for each $j \in \mathcal{J}_{h,B}$. Stacking these equations then yields

$$\sum_{s=1}^{S_h} \Omega^s(w)(w_h - c_h) + \sum_{s=1}^{S_h} \mathbb{E}_\eta[D^{s,h}(p^s(\tilde{c}(w)))] = -\sum_{s=1}^{S_h} \Lambda^s(w)(w_h - c_h)$$

where $\Omega^s(w)$ and $\Lambda^s(w)$ are the matrices defined in the statement of the proposition. Rearranging gives

$$\left(\sum_{s=1}^{S_h} \Omega^s(w) + \Lambda^s(w)\right)(w_h - c_h) = -\sum_{s=1}^{S_h} \mathbb{E}_\eta[D^{s,h}(p^s(\tilde{c}(w)))]$$

as we wanted to show. ♠

Appendix B Characterizing Retail Price Changes with Respect to Changes in Costs

This section characterizes how stores' optimal prices change in response to changes in marginal costs[45]. The results of this section are used in section 5.

Given a product portfolio $\mathcal{J}^s$ and marginal costs $(c_1, \ldots, c_J)$, a store sets prices to solve

$$\max_p \sum_{j=1}^{J} (p_j - c_j) D_j(p)$$

[45]Note that Proposition 1 subsumes the results in this section. However, the results reported in this version of the paper use the characterization in this section instead of Proposition 1.

The first order conditions are

$$\sum_k (p_k - c_k) \frac{\partial D_k}{\partial p_j}(p) + D_j(p) = 0, \quad j = 1, \ldots, J$$

Stacking these equations yields

$$J_D'(p)(p - c) + D(p) = 0 \tag{18}$$

where J_f is the jacobian of the function f. Equation (18) implicitly defines p as a function of c. I'm interested in characterizing $J_p(c)$. Define $H(p, c) := J_D'(p)(p - c) + D(p)$, which is the left hand side of (18). By the Implicit Function Theorem,

$$J_p(c) = -H_p(p, c)^{-1} H_c(p, c) \tag{19}$$

where H_p denotes the matrix of partial derivatives of H with respect to prices and H_c is similarly defined. From (18), $H_c(p, c) = -J_D'(p)$.

Now note that

$$\frac{\partial H_j}{\partial p_l}(p, c) = \sum_k (p_k - c_k) \frac{\partial^2 D_k}{\partial p_l \partial p_j}(p) + \frac{\partial D_l}{\partial p_j}(p) + \frac{\partial D_j}{\partial p_l}(p)$$

and therefore

$$H_p(p, c) = \sum_k (p_k - c_k) \frac{\partial^2 D_k}{\partial p \partial p'}(p) + J_D(p)' + J_D(p)$$

where $\frac{\partial D_k}{\partial p \partial p'}(p)$ is the Hessian matrix of the demand for good k evaluated at p. This result and $H_c(p, c)$ can now be plugged in equation (19) to obtain $J_p(c)$. Computationally, I treat (19) as a collection of J systems of linear equations. The solution to the $j-$th system of equations delivers the $j - th$ column of $J_p(c)$.

References

Adams, B., & Williams, K. (2017). Zone pricing in retail oligopoly.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841–890.

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

Bronnenberg, B. J., Kruger, M. W., & Mela, C. F. (2008). Database paper—the iri marketing data set. *Marketing science*, *27*(4), 745–748.

Collard-Wexler, A., Gowrisankaran, G., & Lee, R. S. (2014). "nash-in-nash" bargaining: A microfoundation for applied work.

Crawford, G. S., Lee, R. S., Whinston, M. D., & Yurukoglu, A. (2015). The welfare effects of vertical integration in multichannel television markets.

Crawford, G. S., & Yurukoglu, A. (2012). The welfare effects of bundling in multichannel television markets. *American Economic Review*, *102*(2), 643–85.

DellaVigna, S., & Gentzkow, M. (2017). *Uniform pricing in us retail chains* (Tech. Rep.). National Bureau of Economic Research.

Draganska, M., Klapper, D., & Villas-Boas, S. B. (2010). A larger slice or a larger pie? an empirical investigation of bargaining power in the distribution channel. *Marketing Science*, *29*(1), 57–74.

Gowrisankaran, G., Nevo, A., & Town, R. (2015). Mergers when prices are negotiated: Evidence from the hospital industry. *American Economic Review*, *105*(1), 172–203.

Grennan, M. (2013). Price discrimination and bargaining: Empirical evidence from medical devices. *American Economic Review*, *103*(1), 145–77.

Hausman, J. A. (1996). Valuation of new goods under perfect and imperfect competition. In *The economics of new goods* (pp. 207–248). University of Chicago Press.

Ho, K., & Lee, R. S. (2017a). Equilibrium provider networks: bargaining and exclusion in health care markets.

Ho, K., & Lee, R. S. (2017b). Insurer competition in health care markets. *Econometrica*, *85*(2), 379–417.

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

Kaelo, P., & Ali, M. (2006). Some variants of the controlled random search algorithm for global optimization. *Journal of optimization theory and applications*, *130*(2), 253–264.

Meza, S., & Sudhir, K. (2010). Do private labels increase retailer bargaining power? *Quantitative Marketing and Economics*, *8*(3), 333–363.

Miller, N. H., & Weinberg, M. C. (2017). Understanding the price effects of the millercoors joint venture. *Econometrica*, *85*(6), 1763–1791.

Nevo, A. (2000). Mergers with differentiated products: The case of the ready-to-eat cereal industry. *The RAND Journal of Economics*, 395–421.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, *69*(2), 307–342.

Newey, W. K., & McFadden, D. (1994). Large sample estimation and hypothesis testing. *Handbook of econometrics*, *4*, 2111–2245.

Nocke, V., & Schutz, N. (2018). Multiproduct-firm oligopoly: An aggregative games approach. *Econometrica*, *86*(2), 523–557.

Powell, M. J. (1994). A direct search optimization method that models the objective and constraint functions by linear interpolation. In *Advances in optimization and numerical analysis* (pp. 51–67). Springer.

Villas-Boas, S. (2007). Vertical relationships between manufacturers and retailers: Inference with limited data. *The Review of Economic Studies*, *74*(2), 625–652.