Essays in Industrial Organization

by

Adam Dearing

**A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2017**

Doctoral Committee:

Professor Daniel Ackerberg, Chair
Assistant Professor Ying Fan
Associate Professor Jeremy Fox, Rice University
Associate Professor Jagadeesh Sivadasan

Adam Dearing

adamdea@umich.edu

ORCID iD: 0000-0002-2752-503X

DEDICATION

For my mother, Jacqueline Denise Dearing. Rest in peace.

ACKNOWLEDGMENT

I thank my advisors, family, and friends who have provided guidance and support toward the completion of this dissertation. My committee members Dan Ackerberg, Jeremy Fox, Ying Fan, and Jagadeesh Sivadasan have provided countless contributions and extraordinary patience. Mel Stephens provided advice and guidance in my early years as a graduate student. Chris Sullivan, Sarah Johnston, Evan Taylor, Alan Griffith, Dan Schaffa, Dan Jaqua, Morris Hamilton, Alecia and Traviss Cassidy, Ben Lipsius, and Mark Fontana provided helpful discussions and brightened the hallways and offices of our beloved Lorch Hall. My father, Tony, and his wife Pei-Pei, my aunts Talia, Liz and Lisa, my uncle Bobby, and all of my extended family have all provided love and support beyond measure.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

This work considers the effects of market structure and regulation on firm conduct and welfare in retail markets. The first chapter considers market structure as a potential explanation for uniform (non-discriminatory) pricing across stores in a retail chain. I use structural empirical methods to show that the presence of an imperfectly competitive upstream wholesale market that is vertically separated from the retailer can reduce the retail chain's incentive to price discriminate across stores. I also discuss the welfare and policy implications of the retailer's uniform pricing. The second chapter considers the effects of a minimum markup policy imposed on retailers in an oligopolistic, differentiated products market where consumers have a "taste for variety." In particular, I consider the effect of the minimum markup on the number of products offered by firms. I present a model in which the introduction of a minimum markup induces firms to expand their number of products offered and show that this product line expansion has important implications for the economic variables of interest to policymakers.

CHAPTER I

Pricing Policy and Welfare in Retail Chains: A Case for Vertical Incentives

1.1 Introduction

Retail chains are frequently observed setting uniform prices for a product across stores in a region, despite demographic and other unobserved differences amongst the stores' customer bases. When consumers' preferences are sufficiently different across store locations, these policies can lead to unrealized profit gains resulting from the failure to optimize prices at the store or "zone" level. Some papers empirically investigate these potential profit gains and find that while they are relatively small in some contexts, they can be quite large in others (Chintagunta et al., 2003; Hoch et al., 1995; Montgomery, 1997).

When profit gains from differentiated pricing are small, it is reasonable to posit additional costs associated with the more complex pricing scheme, known as managerial menu costs, as an explanation for the observed uniform pricing policy (McMillan, 2007). However, when the potential profit gains are large, such an explanation holds little persuasive power and other explanations should be considered. Some potential explanations focus on the demand side and on non-standard models of demand. For example, demand may be a function not only of observed prices but also of the chain's policy choice of uniform or discriminatory pricing.[1] Other explanations focus on the supply side. In this chapter, I contribute to the literature on supply-side explanations for uniform pricing by exploring the potential for upstream market power to affect the retailer's choice of pricing policy.

I also contribute to the literature on the welfare effects of third-degree price discrimination (in this case, store-level pricing). To my knowledge, no previous studies on the welfare effects of third-degree price discrimination account for the possibility of upstream wholesale price adjustment when the retailer changes its pricing policy.[2] An implication of my empirical work is that failure to include such upstream adjustment can result in inaccurate welfare predictions for markets where upstream firms have significant market power. By including the upstream firms' decisions in the analysis of a retailer's pricing policy change, I am able to generate more accurate welfare predictions for markets with an imperfectly competitive upstream segment.

In the main analysis of this chapter, I empirically explore the impact of upstream wholesale price adjustment induced by the (downstream) retailer's change in pricing policy. I evaluate the effect of wholesale price adjustment on the retailer's profit gains, the (upstream) manufacturers' profits, consumer surplus, and total welfare using data from Dominick's Fine Foods, a supermarket chain in the Chicago area, from the early-to-mid 1990s. In the data, I observe the retailer pricing its products semi-uniformly across stores. That is, while not all stores charge the same price for a product, many of them do - far more than might be

[1] For example, Kahneman et al. (1986) find evidence that consumers' perceptions of "fairness" in a firm's pricing policy can influence their purchasing behavior.

[2] Crawford and Yurukoglu (2012) also allow for upstream adjustment in their study on the welfare effects of bundling, a different type of price discrimination.

expected given the differences in demographic characteristics between the stores' customer bases. I focus on the category of mass-produced beer, in which there is scope for wholesale price adjustment due to significant upstream market power. I estimate the (residual) demand system faced by the retailer. I then use the estimated demand system, along with data on both retail and wholesale prices, to back out both downstream and upstream marginal costs. Once these are obtained, I am able to evaluate counterfactual experiments in which the retailer charges fully discriminatory prices, setting profit maximizing prices at an individual store level instead of committing to pricing at the semi-chain level. I evaluate the predicted changes in profits, consumer welfare, and total welfare both with and without upstream adjustment to the retailer's new policy. My counterfactual simulations suggest that upstream adjustment has a large effect on these welfare outcomes. I find that wholesale prices increase when the retailer moves to store-level pricing. The increase in wholesale prices results in a decrease in retailer profits, vertical profits, and consumer welfare. In my counterfactual simulations, only the upstream firms benefit from the wholesale price adjustment once the retailer is already price discriminating.

In addition to evaluating welfare implications, I also use the empirical results to explore the potential for upstream adjustment as an explanation for the retailer's observed uniform pricing. This explanation could be valid for either of two mutually exclusive reasons. First, the retailer may anticipate that upstream adjustment will undo much of the retailer's profit gains from price discrimination. Second, it could be the case that the retailer is "naive" when considering potential profit gains from price discrimination, meaning that it (erroneously) considers wholesale prices to be invariant to its choice of pricing policy. In particular, the naive retailer may predict negligible profit gains from price discrimination, when in reality an upstream response to the policy change could lead to much higher profit gains for the retailer. In either of these cases, the retailer will choose a uniform pricing policy if there are substantial additional managerial costs of implementing price discrimination across stores. In my application, the counterfactual analysis suggests that the first case applies (and rules out the second): the retailer predicts that upstream adjustment will greatly reduce the profitability of a change to store-level pricing.

The remainder of the chapter proceeds as follows. Section 1.2 describes the pricing model used in the chapter, describes the incentive for upstream adjustment to changes in retail pricing policy, and motivates the use of empirical analysis. Section 1.3 describes the beer industry and data used in the chapter. Section 1.4 describes the modeling, estimation, and identification of demand and supply. Section 1.5 presents the results from demand and supply estimation. Section 1.6 presents the counterfactual simulations and discusses their implications. Section 1.7 concludes.

1.2 Double Marginalization and Upstream Adjustment

In this section, I aim to motivate the research question and methods. First, I discuss the standard two-stage linear pricing model for an upstream monopolist and a downstream monopolist, along with the resulting double marginalization problem. I then incorporate the retailer's choice of pricing policy and discuss some of the welfare implications when wholesale prices are constant. Third, I examine the role of upstream adjustment and offer some predictions for two commonly-used demand models. Finally, I discuss some of the possible implications of including upstream competition and multi-product firms in the model.

1.2.1 Two-Stage Linear Pricing

Consider an upstream monopolist (wholesaler) who sells to a downstream monopolist (retailer) in a single market. The market's demand function is $q_m(p_m)$. In stage one of the game, the wholesaler, whose constant marginal cost is c, sets a linear wholesale price, w. In stage two, the retailer observes the wholesale price, which is its marginal cost.[3] The retailer then sets the own profit-maximizing retailer price, $p_m^*(w)$, and sells to consumers at this price. The game is solved by backward induction, and the wholesaler's problem can be characterized as

$$\max_{w} \quad q_m(p_m^*(w))(w-c)$$

where retailer's profit-maximizing price is given by

$$p_m^*(w) = \arg\max_{p_m} \quad q_m(p_m)(p_m - w)$$

This leads to the classic "double marginalization" problem, examined extensively by Rey and Tirole (1986). The main result of the problem is that in the equilibrium, both firms earn positive profits but vertical chain profits are not maximized. This is because the equilibrium retail price is too high; a vertically integrated firm would charge a lower price to maximize vertical profits. There are, however, multiple ways in which vertically separated firms can "solve" the double marginalization problem and price to maximize vertical profits. If such a solution is implemented, then one of the tiers is effectively pricing at marginal cost. Examples of these contracting choices include two-part tariffs, resale price maintenance, and quantity discounts. However, some features of the market, such as the regulatory environment, may

[3] The retailer's constant marginal cost could also include an additional component, so that the marginal cost is $w + c^r$. For parsimony, this term is not included in the model because it does not affect the main results presented in this section.

prevent firms from implementing these joint profit maximizing contracts. The market for beer is one such market - this will be discussed in more detail in Section 1.3.

Now I consider a similar problem where the retailer operates in multiple markets. The total quantity sold across both markets is $Q(p_1, p_2) = q_1(p_1) + q_2(p_2)$. The wholesaler must sell to the retailer at one price, no matter which market the item is sold in. I will consider two cases: 1) the retailer commits to charging a uniform price in both markets; and 2) the retailer price discriminates and may charge different prices in each market. For the moment, I will focus on the retailer's problem. Under uniform pricing, the retailer has one choice variable, p, and its profit-maximizing price can be expressed as

$$p^* = \arg\max_{p} \quad Q(p)(p - w)$$

When the retailer price discriminates, it chooses two prices, and their profit maximizing levels are characterized as

$$(p_1^*, p_2^*) = \arg\max_{p_1, p_2} \quad q_1(p_1)(p_1 - w) + q_2(p_2)(p_2 - w)$$

and I assume, without loss of generality, that $p_1^* > p_2^*$. This assumption implies that demand in Market 1 is less elastic at the optimum.

Several papers compare these two problems when retailer marginal cost is the same under both pricing regimes. Aguirre et al. (2010) provide a review of many main results derived in previous papers and derive some new results of their own. The majority of papers on this topic attempt to predict the sign of the welfare change from implementing price discrimination, relative to welfare under uniform pricing. One key result described by Schmalensee

(1981) and Varian (1985) is that an increase in the total quantity sold is a necessary condition for an increase in total welfare. Intuitively, this is because the transition to price discrimination exacerbates the inefficiency due to monopoly pricing by introducing a misallocation issue. The misallocation arises from selling more units to consumers who value the item less while also selling fewer units to consumers who value the item more. So, an increase in quantity is needed to offset this negative misallocation effect on welfare. Cowan (2012) and Aguirre et al. (2010) derive some sufficient conditions for increases in consumer and total welfare, respectively, when the retailer implements price discrimination. These sufficient conditions require technical assumptions on the demand curve that are beyond the scope of this chapter.

While these predictions are quite useful, they do not take into account a potential change in retailer marginal costs between the different pricing regimes. Such a change will arise if the wholesaler adjusts the wholesale price. Under either pricing policy, the wholesaler's profit maximization problem can be written as

$$w^* = \arg\max_{w} \quad Q(w)(w-c)$$

but the form of $Q(w)$ depends on the pricing policy of the retailer. Explicitly, when the retailer sets a uniform price, $Q(w) = Q(p^*(w))$; when the retailer price discriminates, $Q(w) = Q(p_1^*(w), p_2^*(w))$. So, due to changes in the function $Q(w)$, the wholesaler may have incentive to change w^* when the retailer moves toward price discrimination. Either way, the wholesaler's profit maximizing wholesale price must satisfy

$$\frac{w^*-c}{w^*} = -\frac{1}{\varepsilon_{w^*}}$$

where $\varepsilon_w = \frac{dQ}{dw} \cdot \frac{w}{Q} < 0$ is the wholesale price elasticity of aggregate demand. When this elasticity is larger in magnitude, the equilibrium wholesale price will be lower.

1.2.2 Examples of Upstream Adjustment

Here I discuss some predictions for upstream wholesale price adjustment when the retailer faces some commonly-used demand functions. For the moment, I continue to focus on the setting described in Section 1.2.1. If both markets have a constant elasticity of demand and the elasticities differ across markets, then some analytic predictions are possible.

Proposition 1. *(Constant elasticity) If demand in each market,* $i \in \{1, 2\}$*, is given by* $q_i(p_i) = A_i p_i^{\varepsilon_i}$*, then relative to the equilibrium when the retailer commits to uniform pricing:*

1) The equilibrium wholesale price is lower when the retailer price discriminates.

2) Wholesaler and retailer profits are both higher when the retailer price discriminates.

A technical proof of the proposition is presented in Appendix A, but here it is useful to discuss the intuition behind the decrease in wholesaler price when the retailer's policy changes from uniform pricing to price discrimination. To begin, it is useful to note that the wholesale price elasticity of aggregate demand can be expressed as

$$\varepsilon_w = \varepsilon_1 \kappa_1 s_1 + \varepsilon_2 \kappa_2 s_2$$

where $\varepsilon_i = \frac{dq_i}{dp_i} \cdot \frac{p_i}{q_i}$ is a market's price elasticity of demand, $\kappa_i = \frac{dp_i}{dw} \cdot \frac{w}{p_i}$ is the pass-through elasticity, and $s_i = \frac{q_i}{Q}$ is the market's share of aggregate demand. This is similar to the standard result that, under uniform pricing, the own-price elasticity of aggregate demand is equal to the share-weighted own price elasticities of demand in each market. However, the formula of interest in this application must also account for the wholesale price passing through to the retail price. Note that $\kappa_1 = \kappa_2$ under uniform pricing but that $\kappa_1 \neq \kappa_2$ is possible under price discrimination.

Suppose that, in each market, the pass-through elasticity is invariant to the retailer's choice of pricing policy. The intuition behind the change in wholesale price is then clear. When the retailer moves from setting a uniform price to price discrimination, it increases the price in the relatively inelastic market and decreases the price in the more elastic market. So, demand in the inelastic market decreases while demand in the elastic market increases. This leads to an increase in the more elastic market's share of aggregate demand. Because the market own-price elasticities are constant, this implies that ε_w becomes more elastic and the wholesaler should then lower its wholesale price. Simulations suggest that the effect of the wholesale price change on profits, consumer welfare, and total welfare can be large.

The intuition from the constant elasticity case is compelling, but it does not apply for all forms of demand. In particular, when demand in each market is linear, the analytic predictions change.

Proposition 2. *(Linear demand) If the demand function in each market is linear, then relative to the equilibrium when the retailer commits to uniform pricing:*

1) The wholesale price is the same when the retailer price discriminates.

2) Wholesaler profits are the same when the retailer price discriminates.

Proof. Pigou (1920) shows that $Q(w)$ has the same form under both retailer price discrimination and uniform pricing. $\square$

The analytic results from Propositions 1 and 2 show the potential for either a decrease or no change in the wholesale price when the retailer switches from uniform pricing to price discrimination. The following numerical example shows a situation in which the wholesale price rises.

Example 1. (Concave and linear demand) Suppose the demand function in market 1 is $q_1 = \sqrt{100^2 - p_1^2}$, demand in market 2 is $q_2 = 100 - p_2$, and the wholesaler's marginal cost is $c = 10$. Then it can be shown via numerical simulations that the equilibrium wholesale price is higher under price discrimination than when the retailer commits to uniform pricing.

Taken together, Propositions 1, Proposition2, and Example 1 have some important implications. First, they serve to motivate the need for empirical analysis of the question at hand by illustrating the potential for upstream adjustment to affect (or not affect) the equilibrium predictions. Second, the propositions show the need for a sufficiently flexible demand system in empirical analysis, so that the predicted effects of upstream adjustment are not simply an artifact of demand model selection.

1.2.3 Competition and Multi-Product firms

Throughout this section, I have assumed that there exists a monopoly at both levels of the vertical supply chain and that only one good is traded. This assumption will not extend to the empirical application, where there is an upstream oligopoly and multiple goods are traded in each market. With an upstream oligopoly in place, analytical analysis appears to become intractable without imposing unrealistically restrictive assumptions on the demand curve, even when only one good is traded. The same issue applies when multiple goods are included in the model. As discussed above, overly-restrictive assumptions on demand should not be used in empirical analysis. So, I do not pursue analytic predictions for the game with an upstream oligopoly and/or multiple goods.

Despite the absence of analytic results, some intuition can still be applied to more complex models. When there are multiple goods and upstream competition, it is no longer only the own-price elasticities that matter but also cross-price elasticities. This leaves room for a good's aggregate quantity traded to be less sensitive to changes in its wholesale price when the retailer moves toward price discrimination. If cross-price elasticities are lower in markets where own-price elasticities are higher, then the transition to retailer price discrimination will "soften" competition. That is, the markets with higher own-price elasticities and lower cross-price elasticities will now have greater shares of aggregate output, implying that aggregate quantities will be less responsive to changes in the wholesale prices of other goods. With a downstream oligopoly, such softening of competition would cause wholesale prices to rise. And so, the net effect on wholesale price will depend on the relative size of market-level own- and cross-price elasticities; it is reasonable to suspect that the wholesale prices could either rise or fall.

1.3 Industry and Data

1.3.1 Double Marginalization in the Illinois Beer Industry

The beer industry features some regulatory and economic distinctions that make it an ideal industry for this study. Recall from the previous section that the double marginalization problem can arise when retailers and manufacturers are vertically separated and both have some market power. All of these conditions are satisfied in the beer industry. Vertical separation is imposed by law, with the beer industry in most U.S. states separated into three tiers: manufacturers, distributors, and retailers. The manufacturers brew and sell beer to distributors, who then sell to retailers. In Illinois, all three tiers are privately owned and operated. For mass-produced beers, prior work and industry lore suggest that during the early-mid 1990s, a manufacturer-distributor pair acted as a single vertically integrated entity, with the manufacturer imposing (possibly illegal) resale price maintenance (RPM) on the distributor (Hellerstein, 2008; Asker, 2015). Therefore, I model the industry as two-tiered in my analysis and assume that the manufacturers make the wholesale pricing decisions and fully internalize their distributor's marginal costs.

In addition to vertical separation, the industry exhibits market power at both tiers. Manufacturer market power is well-documented in the beer industry and is universally assumed in the literature. Retailer market power is captured in the retailer's residual demand curve. On the manufacturer side, the industry exhibits a large degree of concentration, with the top four brewers maintaining a total national market share of at least 87% from 1990-2002 (Tremblay and Tremblay, 2005). There is also significant product differentiation through branding, even though many mass-produced beers have similar physical characteristics.[4]

While these features of the industry structure allow for the possibility of double marginalization, there is still a concern that firms may find a way to eliminate this problem by implementing joint-profit maximizing contracts. Most commonly, such contracts feature either one or more of three mechanisms: two-part tariffs with fixed payments, resale price maintenance, and quantity discounts. If any of these contracting strategies are implemented, then the industry will effectively be vertically integrated. Fortunately for this empirical application, the regulatory environment eliminates the first two options and the data do not

[4]Tremblay and Tremblay (2005) provide a detailed analysis of the beer industry.

support the third. The 1935 Federal Administration of Alcohol Act prohibits slotting allowances or other fixed payments in the retailing of alcoholic beverages. So, firms cannot enact two-part tariff contracting without violating federal law. Additionally, during my sample period, resale price maintenance was *per-se* illegal. Although RPM was apparently overlooked between manufacturers and distributors, there is a lack of similar evidence to support a claim that such a practice was also imposed on retailers. Furthermore, wholesale prices observed in the data do not appear to vary with quantity sold.

Along with the above evidence supporting a vertical structure featuring double marginalization, Hellerstein (2004) implements a series of non-nested tests developed by Villas-Boas (2007) and finds that the best-fit model for the industry is one of simple linear pricing by both upstream and downstream firms.[5] Taking all this evidence into account, I conclude that simple two-stage linear pricing is the most reasonable model for the supply side of the data.

1.3.2 Description of the Data

The data come from Dominick's Fine Foods, a supermarket operating in the metropolitan Chicago area with an approximately 20% market share of all supermarket business. The data are publicly available for academic use from the Kilts Center. The data used in this study include weekly scanner data on consumer purchases of beer from 71 Dominick's stores in 220 weeks from late-1991 to early-1995. The dataset also includes demographic data for the stores' ZIP codes, obtained from the 1990 Census, and some data on consumers' shopping habits within each ZIP code. Most ZIP codes contain at most one store. Because some stores

[5] I am unable to replicate these tests due to a lack of data on marginal cost shifters.

are not present in all weeks of the data, I use a subsample of stores and weeks. I use data from 49 of the stores in 186 weeks. The sales data are collected at the store-week-UPC level. For each UPC, the data include the brand, container size and type (12oz cans, 20oz bottles, etc.), package size (6-pack, 12-pack, etc.), and a manufacturer code. Each observation in the data also includes the unit sales, retail price, wholesale price, and promotional activity. For the beer category, almost all promotional activity is in the form of a "bonus buy," which is a simple price reduction.

I restrict my analysis to beers produced by three brewers, which together account for 73% of category revenue in the sample period: Miller (44%), Anheuser-Busch (15%), and Heilman(14%).[6] I further restrict my analysis to a few "flagship" brands from each retailer, which are listed in Table 1.1. The categories presented in the table are industry classifications. The chosen brands from each manufacturer represent over 80% of the revenue collected by Dominick's for all of that manufacturer's brands. I make these selections in order to focus on the products for which upstream adjustment is most likely to affect Dominick's incentives while also keeping the sample size reasonable for structural analysis.

I define a unit as a 12oz container of beer, as is standard in the literature, and restrict the analysis to 6- and 12-packs of such containers. Although Dominick's also sells 24- and 30-packs of some brands, these seem to be considered distinct products from the smaller package sizes. Meza and Sudhir (2006) note that the large package sizes are marketed toward a "heavy user segment." The large package sizes are also sold as loss-leaders in the data, while 6- and 12-packs are both sold at considerable (positive) profits. The selected products included account for 58% of category profits for the store-weeks included in the

[6]Nationally, Anheuser-Busch commanded the largest market share, followed by Miller and then Coors during the sample period (Tremblay and Tremblay, 2005).

final sample. Within a store-week, I aggregate unit sales to the brand-size level; prices within a store-week do not vary across UPC's within a brand-size. So a product is defined as a set of 12oz containers where the brand, pack size, retail price, week, and store are all the same. I define the market size as the average weekly beer consumption of the adult population in the store's ZIP code.[7] The discrete choice demand model requires an outside option, which here is defined as beer purchased from other stores plus beer purchased from Dominick's not included in the sample.

Dominick's appears to set retail prices "semi-uniformly" across the chain; meaning that while there is a high degree of uniformity in the observed price for a product across stores, prices are not completely uniform. This pattern is illustrated in Figure 1.1, which plots the retail and wholesale prices for 12-packs of Budweiser in each store-week. That is, the prices of all 49 stores are plotted on the graph. In some weeks, all 49 stores set the same price. In other weeks, there are multiple prices observed in the data, but the number of distinct prices is not so large as to indicate full price discrimination. Section 1.4.2 provides a further discussion of this semi-uniform pricing and describes how it is incorporated into the supply-side estimation.

A potential problem arises from the way Dominick's computes wholesale prices. The wholesale price recorded in the data reflects the average acquisition cost (AAC) of the retailer's inventory of each UPC. If different wholesale prices were paid for different units in the inventory, then the wholesale price in the data may not reflect the true marginal cost to the retailer. However, there is very little variation in the recorded wholesale price across stores within a UPC-week, suggesting that this is not a concern for this particular product cate-

[7] Average beer consumption per adult is 4.6 12oz containers per week (Hellerstein, 2008). Market size is then equal to $AdultPop_{ZIP} \times 4.6$ in a store-week.

gory. Within a week, I use the average recorded AAC across all stores for a brand-size as the wholesale price in my analysis. In other words, the wholesale price used in my analysis is the chain-level average of a within-store average of the wholesale prices for units of a product. For over 90% of the observations, the recorded AAC (within-store average) differs from this chain-level averaged wholesale price by less than 2%.

Another concern with the data arises from the frequency of the observations. If consumers are able to store beer, then the retailer will have dynamic pricing incentives, and a static weekly demand model may not capture the relevant elasticities of interest. Instead, intertemporal substitution will cause the static demand model to over-predict own-price elasticities. However, there is substantial evidence to suggest that intertemporal substitution is not a concern in the market for beer. Hendel and Nevo (2006) describe the the implications of storability on the timing and effectiveness of price promotions by the retailer. Goldberg and Hellerstein (2012) test these predictions in the Dominick's beer data and find no evidence of the patterns predicted by Hendel and Nevo (2006). Additionally, marketing research in the beer industry suggests that beer is usually consumed within a few hours of purchase (Hellerstein, 2008).

Table 1.2 presents some summary statistics of the data. The final sample includes $N = 269,146$ products as defined above (i.e. store-week-brands). All prices are deflated to 1991 prices using the Consumer Price Index, available from the Bureau of Labor Statistics. Promotional activity is observed in just above one-fourth of the observations, although these are simple price reductions rather than feature or display promotions. In about 10% of the observations, the retail price is below the wholesale price. Product market shares have a mean of 0.09%, a median of 0.04%, and a standard deviation of 0.17%. The mean share of the outside option is 97.33%, with a standard deviation of 1.96%. Note that these low market shares for products and high outside option shares are partly an artifact of the products' definition and selection and do not necessarily reflect an almost perfectly competitive market. For example, the outside option shares decrease substantially if I include the 24-packs in the product sample, although they are excluded for aforementioned reasons. Including other products sold by Dominick's can further reduce the outside option's share. Hellerstein (2008) uses a wider selection of products and finds an average outside option share of 81.5%. The mean retail price (per 12oz) is 57.25¢ with a median of 58.19¢ and standard deviation 5.91¢. Wholesale price has a mean of 46.40¢ with median 46.41¢ and standard deviation 2.33¢. The mean markup is then 10.85¢ suggesting a mean retail margin of about 19%, not taking into account unobserved marginal costs for the retailer. This suggested mean retail margin is confirmed in the data.

Table 1.2 also includes some summary statistics of the demographic variables used in the analysis, drawn from 49 stores across the metropolitan Chicago area. Two variables drawn from the 1990 Census at the ZIP code level are median income and the percentage of the population over age 60. The third variable is an ability-to-shop index provided by Dominick's. This variable measures consumer mobility in the market, with a higher value indicating that consumers have more ability to visit multiple locations in a shopping trip. This can be due to greater mobility via automobile ownership, greater availability of time to dedicate to shopping trips, etc. Variation in these measures across stores is likely to provide incentive to for Dominick's to price discriminate through their influence on consumers' preferences. In particular, I expect that consumers in areas with higher income will be less price sensitive and that consumers with higher ability to shop will be more price sensitive. The summary statistics presented in Table 1.2 show substantial variation in demographics across stores. The minimum (of median) income is $21,300 and the maximum is $75,600. The standard deviation across stores is \$11,500. The other two demographic variables also vary significantly across stores. The large variation in demographics is likely to induce heterogeneity in consumer preferences across ZIP codes, providing opportunity for the retailer to increase profits by pricing at the store level.

There may be a concern about whether the 49 stores selected for the sample can be considered as representative of the chain at large. While this is difficult to formally test due to the fact that the excluded stores are missing in many weeks (otherwise, they would be included), some informal checks can still be applied. One method is to examine the joint distribution of the demographic variables. If the joint distribution across the selected stores is similar the joint distribution across the excluded stores, then this provides some informal evidence that the selected stores can be taken as representative of the chain as a while. Figure 1.2 shows kernel density plots of median income for the stores in the sample and for the excluded stores. Each store represents one observation. As the figure shows, these plots are very similar. Figure 1.3 shows a median income plotted against the ability-to-shop index for stores both in and excluded from the sample. Again, these plots are very similar. Together, these figures suggest that the joint densities of demographics for stores both in and out of the sample are quite similar.[8]

1.4 Model, Estimation, and Identification

In this section, I present the structural demand and supply models used in estimation. I discuss the estimating equations, instruments, and identification where appropriate. The demand model is a discrete choice model, a class of model which is commonly used in structural

[8]Similar patterns hold for the age variable. However, income and the ability-to-shop index have the most predictive power in the estimated model, so I limit the joint density discussion to those two variables.

analysis of the market for beer (Asker, 2015; Hellerstein, 2008; Miller and Weinberg, 2016; Goldberg and Hellerstein, 2012). The supply side is modeled as a static three-stage game. In Stage 1, the retailer commits to a pricing policy - either semi-uniform or discriminatory pricing. In Stage 2, manufacturers are Nash-Bertrand competitors and set profit-maximizing linear wholesale prices. In Stage 3, the retailer observes the wholesale prices set in the previous stage and sets profit-maximizing retail prices. The subgame beginning in Stage 2 is solved via backward induction and is repeated each week. The retailer's decision in Stage 1 is assumed to hold for entire time period covered by the sample. This modeling approach is consistent with McMillan (2007), who notes that the decision of general pricing policy is a long-run decision and is not reevaluated at a weekly level. In the observed data, the retailer has selected semi-uniform pricing. The supply models used to estimate retailer and manufacturer marginal costs presented below assume that this semi-uniform pricing policy is in place.

1.4.1 Demand

The demand model used here is the nested logit model (McFadden, 1978; Cardell, 1997) with products in the sample in one nest and the outside option in another.[9] The nested logit model helps alleviate some of the issues with the standard logit model. Given my specific modeling choice, the main attractive feature is that the nested logit model is better-able to capture within-nest substitution. This is of particular importance due to the high observed outside option shares in the data. In the standard logit model, the relative sizes of the cross-price elasticities depend solely on the market shares. So, a standard logit model would likely

[9] My approach allows modifications for further levels of nests, e.g. for package size, brewer, industry category, etc.

over-estimate substitution to the outside option in my sample. This prediction is confirmed in the estimation results, presented later.

An alternative to the nested logit model in this context would be the flexible random coefficients model of BLP, although the dearth of meaningful product characteristics in my sample may limit the advantages of using such a model over the nested logit. Another alternative model is the random coefficients nested logit model used by Miller and Weinberg (2016) in their study of the beer market. However, their application of that model would only allow for somewhat more flexible incorporation of the demographic characteristics, as they do not allow for an unobserved term in the random coefficients. Furthermore, Grigolon and Verboven (2014) show that the nested logit, random coefficients, and random coefficients nested logit models all yield similar counterfactual predictions in a data structure similar to mine. One caveat to my modeling choice is that there is not a natural way to include within-ZIP-code heterogeneity in the demographic variables, whereas the models with random coefficients would allow me to do so.[10]

The underlying utility function for individual, i, of product,j, in group, g, sold in store (market), m, at time, t, is given by

$$u_{ijmt} = V_{jmt} + \zeta_{igmt} + (1 - \sigma)\epsilon_{ijmt}$$

where the mean utility level is composed of an observed component, δ_{jmt}, and an unobserved market-level shock, ξ_{jmt}, such that $V_{jmt} = \delta_{jmt} + \xi_{jmt}$. The individual product-level shocks, ϵ_{ijmt}, are distributed i.i.d. Type 1 extreme value, and the distribution of the group-level

[10]Such heterogeneity could be included in the standard logit model by, for example, including the standard deviations of the demographic variables as extra linear regressors. However, the coefficients on such parameters would not have a clear economic interpretation. The random coefficient models allow for such variation to be included in a way that is more natural but at the cost of greater computational burden.

shocks, ζ_{igmt}, is the unique distribution such that $\zeta_{igmt} + (1-\sigma)\epsilon_{ijmt}$ is also distributed Type 1 extreme value (Cardell, 1997).

The observed component of the mean utility level is modeled as

$$\delta_{jmt} = \alpha_m p_{jmt} + X_{jmt}\beta$$

where p_{jmt} is the product's retail price and X_{jmt} is a vector of covariates, all of which are dummy variables. Dummy variables indicating a holiday, a promotion, and whether the product comes in a 12-pack are included. The holiday dummy is set to 1 in the week prior-to and including Memorial Day, July 4th, Labor Day, Thanksgiving, and Christmas. I also include brand, brewer-year, and store dummies. Demographic variables are not included in X_{jmt} because store dummies are included, while demographics are collected at the store level and do not vary over time in my sample. The coefficient on price, α_m, is expected to be negative in all markets and can be decomposed into

$$\alpha_m = \alpha_1 + D_m\alpha_2$$

where D_m contains the market's demographic variables, which do not vary over time in the data. In this specification, the demographic variables affect the mean utility in two ways: through their level effect (via store dummies) and through their effect on consumers' price sensitivity. The mean utility of the outside option ($j = 0$) is normalized to be zero.

Berry (1994) shows that these modeling assumptions lead to a linear estimating equation of the form

$$ln(s_{jmt}/s_{0mt}) = \alpha_1 p_{jmt} + (p_{jmt}D_m)\alpha_2 + \sigma s_{j|gmt} + X_{jmt}\beta + \xi_{jmt}$$

where s_{jmt} is product j's share of the market, s_{0mt} is the outside option's share, and $s_{j|gmt}$ is product j's within-group share in the market. The parameters to be estimated are α_1, α_2, σ, and β. The regression residual is ξ_{jmt} and all other variables are observed or can be computed from the data. The nesting parameter, $\sigma \in (0, 1)$, measures the correlation in individual-level unobserved taste shocks for the products within a nest. A higher value of σ implies greater within-nest substitution, while a value of $\sigma = 0$ corresponds to the standard logit model.

Estimation proceeds using the generalized method of moments (GMM). The GMM assumption used in estimation is

$$E[\xi_{jmt}|Z_{mt}] = 0$$

where Z_{mt} is a vector of instruments. The need for instruments arises from the endogeneity of two terms: p_{jmt} and $s_{j|gmt}$. Both of these terms should be positively correlated with the unobserved residual, ξ_{jmt}, leading to an upward bias in both the (negative) estimate of α_m and the estimate of σ. I must therefore instrument for both of these terms. The instrument used for price is the wholesale price. The justification for this instrument is similar to that of Chintagunta (2002) and is somewhat non-standard due to the presence of upstream market power in the supply-side model. In order to use wholesale price as an instrument, it must be correlated with price and uncorrelated with the market-level shocks. The former property is confirmed in the data. The latter is an assumption that requires justification. As noted earlier, the wholesale price is set at the chain level, rather than the store level. The retailer may be better-able to observe changes in local demand than the wholesalers, suggesting that wholesale prices will be less correlated with ξ_{jmt} than are retail prices. If the wholesalers perfectly observe changes in local demand before setting the wholesale price, then the wholesale price will not be a valid instrument.

This argument motivates the inclusion of the holiday and brewer-year dummies in the vector of covariates, X_{jmt}. The holiday dummies capture a predictable component of demand. Furthermore, the brewer-year dummies capture the effect of brewers' advertising decisions on demand. Figure 1.1 shows a plot of all observed wholesale and retail prices for a 12-pack of Budweiser over the sample period. This includes the prices observed in all stores for each week. It is clear from this plot that the wholesale prices show little variation within a year but vary substantially from year to year. Some other products exhibit more within-year variation, but the pattern is consistent: there is much less within-year variation in wholesale price than is present in the retail price.[11] The significant year-to-year variation in wholesale prices can be explained by variation in costs or by the brewers' advertising decisions. The effect of cost changes on wholesale prices does not present an endogeneity problem, but the the extent to which advertising decisions affect wholesale prices can be problematic. If national advertising decisions are made by the brewers on a yearly basis, then year-level changes in the wholesale prices will partially reflect the expected changes in demand. Inclusion of the brewer-year dummies in the demand system helps alleviate this concern.

The instruments for a product's within-group share are similar to those of Meza and Sudhir (2006). I use the average wholesale price of other products within the nest, along with the square of the difference between this average and the product's own wholesale price. These elements capture the product's relative position in the group along the price dimension. Controlling for product characteristics (via the brand and pack size dummies), the products

[11] Some within-year variation can be accounted for by the inclusion of the holiday dummy in the demand model.

with lower relative prices should have higher within-group shares. So, these instruments will be correlated with the within-group share and will also be uncorrelated with the residual for the reasons explained previously.[12] With these instruments in place, the full GMM instrument vector is

$$Z_{jmt} = \left(w_{jmt}, w_{jmt} \times D_m, \bar{w}_{(i \neq j)mt}, \left(w_{jmt} - \bar{w}_{(i \neq j)mt}\right)^2, X_{jmt}\right)$$

Identification of the coefficient vector β is standard. Identification of the element of price sensitivity common across stores, α_1, comes from the variation in retail prices driven by cross-sectional and over-time variation of wholesale price within brewer-year, within-brand, within-store, and within-pack-size variation of wholesale prices. Identification is possible because dummy variables at each of these grouping levels are included separately (e.g. there are separate brand and store dummies instead of brand-store dummies). Product-level wholesale prices show very little within-year variation, but the inclusion of brewer-year dummies instead of brand-year dummies allows for the required variation of wholesale price. Identification of α_2 comes from the variation of demographic variables across stores (interacted with the wholesale price). The nesting parameter, σ, is identified by variation in the instruments for within-group share, which is due to the variation in wholesale prices across all products in a market. There is also variation in the instruments for within-group share due to product introduction and attrition in the sample. One of the brands included only appears after 1991 and another brand only appears before 1995.

[12]Formally, $E[\xi_{jmt}|w_{jmt}] = 0$ is required for the price instrument to be valid, while the stronger condition that $E[\xi_{jmt}|w_{mt}] = 0$ is required for the within-group share instruments to be valid.

1.4.2 Downstream Retailer Pricing

Dominick's is assumed to set prices at the retail level on a weekly basis after observing all of the ξ_{jmt}'s. Dominick's appears to set prices semi-uniformly across its stores. The fact that prices are not set completely uniformly introduces a problem with determining the number of prices that Dominick's sets in a week. I resolve this issue by assuming that Dominick's does not set uniform prices by "simple coincidence." Consider, for example, a week in which I observe that Dominick's sells 12-packs of Budweiser at a uniform price in 47 stores and that the remaining 2 stores each have their own prices. in this case, I assume that Dominick's optimal pricing problem for 12-packs of Budweiser required 3 prices: one price for the 47 stores and two separate prices for the remaining stores. Another interpretation is that Dominick's considers the 12-packs of Budweiser sold in the 47 stores to be one product, while the 12-packs sold in the remaining stores are two other distinct products.

Following this logic, I define a product in the retailer's semi-uniform pricing problem as a set of observations where the brand, pack size, week, and retail price are the same. This results in $N = 19,728$ total products for the retailer's pricing problems across the sample. Note that if Dominick's priced uniformly across all stores in all weeks, the number of products resulting from this procedure would be $6,221$. And if Dominick's practiced full price discrimination in all weeks, the number of products would be $269,146$. So, while the pricing is not completely uniform, it still exhibits a high degree of uniformity.

Prices are set week-by-week, so I will omit the time subscript in the following analysis. The retailer's profit maximization problem for a given week is then

$$\max_{p_1, p_2, \ldots, p_J} \quad \sum_{j=1}^{J} (p_j - w_j - c_j^r) \sum_{m \in R_j} M_m s_{jm}(\boldsymbol{p_m})$$

where product j is defined as above. R_j is the set of stores selling product j. In the example above, R_j would either include the 47 uniform price stores or either of the remaining stores. M_m is the market size, p_j is the retail price, w_j is the wholesale price, c_j^r is the unobserved component of the retailer's marginal cost, and s_{jm} is the product's market share.[13] I will omit the input, $\mathbf{p}_m$, for parsimony. Market shares are computed using the estimated demand system. I impose the assumption that c_j^r does not vary by market. The retail and wholesale prices do not vary by market due to the retailer's assumed product definitions.

This profit maximization problem yields the first order conditions

$$\sum_{m \in R_j} M_m s_{jm} + \sum_{k=1}^{J} (p_k - w_k - c_k^r) \sum_{m \in \{R_j \cap R_k\}} M_m \frac{\partial s_{km}}{\partial p_{jm}} = 0 \quad \forall\, j = 1, ..., J$$

which, following Villas-Boas (2007), can be represented together in matrix notation as

$$\boldsymbol{p} - \boldsymbol{w} - \boldsymbol{c^r} = -\Omega_r^{-1} Q$$

where the elements of Ω_r and Q are given by

$$\Omega_r(j,k) = \sum_{m \in \{R_j \cap R_k\}} M_m \frac{\partial s_{km}}{\partial p_j}$$

$$Q(j) = \sum_{m \in R_j} M_m s_{jm}$$

and $\Omega_r(j,k) = 0$ if $\{R_j \cap R_k\} = \emptyset$.[14] Because retail and wholesale prices are observed and the elements of the right-hand side of the first order condition can be computed from

[13] c_j^r is observed by the retailer but not by the econometrician.

[14] Note that although the partial derivatives are written in terms of p_{jm}, it is still true that $p_{jm} = p_j$.

the estimated demand system, the vector of unobserved retailer marginal costs, $\boldsymbol{c^r}$, can be backed out from the first order conditions. This is why I assume c_j^r does not vary across markets: each j only admits one first order condition equation, and I do not have data on market-level shifters of the retailer's unobserved marginal cost.

1.4.3 Upstream Manufacturer Pricing

Manufacturers set one wholesale price per product for the entire chain, and I assume that they set prices on a weekly basis. So, from the manufacturer's perspective a product is defined as a set of observations with the same brand, pack size, and week. This results in $N = 6,221$ manufacturer-level products across the sample, as opposed to the $19,728$ retailer-level products and $269,146$ consumer-level products. I model the wholesalers as setting wholesale prices without observing the ξ_{jmt}'s in each week. Results from this modeling choice may not be accurate if wholesalers (partially) observe ξ_{jmt} in the true data generating process. As a robustness check, I compare the results from this model to results from one in which the wholesaler observes the ξ_{jmt}'s then sets wholesale prices. Both modeling choices produce very similar counterfactual predictions.

I assume that manufacturers are Nash-Bertrand competitors and that they anticipate the retailer's response to changes in wholesale prices. Prices are set weekly, and I again omit the time subscript for parsimony. Manufacturer (brewer) b's profit maximization problem for a given week is

$$\max_{w_{1_b},\ldots,w_{F_b}} \quad \sum_{f\in B_b}(w_f - c_f^b)\sum_{m\in R_f} M_m s_{fm}(\boldsymbol{p_m}(\boldsymbol{w_m}))$$

where F_b is the number of products in B_b, the set of products produced by the manufacturer. The manufacturer's unobserved marginal cost for product f is c_f^b and is assumed to be constant across markets.[15] The rest of the notation is analogous to that found in the retailer's pricing problem, with the exception of the market shares. Here, the market shares are the expected shares because the wholesaler does not observe ξ_{jmt}.[16] The notation in the problem explicitly shows market shares as a composite function of retail and wholesale prices, although from here I will omit the input argument.

The first order conditions from this problem are given by

$$\sum_{m \in R_f} M_m s_{fm} + \sum_{h \in B_b} (w_h - c_h^b) \sum_{m \in \{R_f \cap R_h\}} M_m \frac{\partial s_{hm}}{\partial \boldsymbol{p_m}} \frac{\partial \boldsymbol{p_m}}{\partial w_{fm}} = 0 \quad \forall f \in B_b$$

and these conditions should hold for every manufacturer in the Nash equilibrium. Villas-Boas (2007) and Villas-Boas (2009) discuss the computational implementation of these terms, which are used to find c_h^b.

1.5 Estimation Results

1.5.1 Demand Results

I estimate the demand model under both simple logit and nested logit specifications. The estimates of demand parameters are presented in Table 1.3. Model 1 is the simple logit model with no instrument for price. Model 2 is the simple logit model with a price instrument. Model 3 is the preferred specification: a nested logit model with instruments for both

[15] This assumption may be too restrictive because it includes the distributor's marginal cost, and Asker (2015) finds evidence that there are marginal cost differences between distributors servicing different Dominick's stores.

[16] The model I use replaces the estimated ξ_{jmt}'s with $\xi_{jmt} = 0$ in the manufacturer's pricing problem. This is far less computationally burdensome than computing expected market shares by integrating over the joint distribution of ξ_{jmt}'s.

price and within-group share. Huber-White heteroskedasticity robust standard errors are presented in parentheses.

All three estimated models exhibit some of the expected patterns in the data. The estimated constant coefficient on price is negative for all three models. The interaction of price with income produces a positive coefficient, which implies that the magnitude of price sensitivity is lower in areas with higher median income levels. The coefficient on the interaction of price with the ability-to-shop index is negative, implying that persons who are more able to "shop around" are more price sensitive. The holiday dummy variable is positive and very precisely estimated.

Two other results arise in estimation across all three models, although these two are not predicted a priori. First, the coefficient on the interaction of price with age is positive. This suggests that elderly consumers are less price sensitive than younger consumers, although I do not speculate on the mechanism behind this result. In Model 3, this age interaction coefficient is the least well-identified of the price coefficients. Second, the coefficient on the 12-pack dummy is positive. This is likely an artifact of the definition of one "unit" as a 12oz serving in the demand system. Asker (2015) uses a different unit definition than this chapter - namely, an entire package - and finds that the sign on his size variable's coefficient is negative.

The coefficient estimates for the promotion dummy are potentially troubling. Promotional activity is often thought to increase demand for a product, but here the coefficient is estimated to be negative. This is explained by a closer examination of the promotion dummy. In the sample studied here, the promotions are overwhelmingly simple price discounts instead of feature or display promotions. A feature or display promotion should increase demand at any price (shifts the whole demand curve), but a simple price discount will only increase quantity demanded via the price decrease (movement along the demand curve). If Dominick's discounts prices in periods of particularly low demand, then the promotion dummy will be negatively correlated with the demand regression residual, resulting in a negative bias. When the true coefficient is zero, this can produce a negative estimate of the coefficient. Essentially, the coefficient estimate on this dummy variable could be picking up the retailer's expectation of low product demand in those periods, which triggers a price promotion. However, I have also estimated demand without the promotion dummy and found that the promotion dummy's inclusion in the regression appears to have little effect on the other coefficients' estimates.

A comparison of Model 1 to Model 2 provides a quick check of the validity of the price instrument. Both models include demographics interacted with price, so it is better to compare the implied own-price elasticities across the two models, rather than to compare the constant coefficient on price. The median own-price elasticities are -3.53 for Model 1 and -5.79 for Model 2. The lower numeric value in Model 2 is consistent with wholesale price being less correlated with the demand residual than is the retail price, indicating that the price instrument is working as intended.

The estimates for Model 3, the preferred specification, illustrate the importance of including the nesting structure in the demand system. The estimate of the nesting parameter is high at 0.748, suggesting a large degree of within-nest substitution in consumers' behavior. Miller and Weinberg (2016) also obtain a large estimate of the nesting parameter in the random coefficients nested logit models for beer, although their estimate is higher than mine. The final row of the table is of particular interest, as it captures the median cross-price elasticity of the outside option with respect to each product's price. This statistic provides a measure of the degree of substitution toward the outside option. As expected, the nested logit (Model

3) predicts far less substitution toward the outside option than the simple logit (Model 2), highlighting the need for the more flexible nested logit model - the simple logit predicts three times as much substitution toward the outside option. The median own-price elasticity is -7.44 for Model 3, with a standard deviation of 2.14 and an inter-quartile range of 2.19. Most of this variation in the is driven by the effect of ZIP code-level demographics on the price coefficient, although I do not present results from a model without these effects in the table. Such substantial variation in the elasticities across markets (stores) elicits the possibility of substantially increased profits if the firm fully price discriminates across locations.

1.5.2 Supply Results

Table 1.4 presents the results for the downstream retailer. The market shares used in the first order conditions are the true market shares in the data. Retailer profit maximization implies a mean markup above marginal cost (per 12oz) of 35.04¢, with a median of 33.53¢, and a standard deviation of 6.91¢. This is in contrast the the median markup above wholesale price observed in the data, which is only 10.47¢. The median implied unobserved marginal cost is -24.57¢. This result is surprising; unobserved marginal costs are expected to be positive. Instead, the implied unobserved component of downstream marginal cost is negative for nearly all observations. However, the total implied downstream marginal cost, which includes the both the wholesale price and the unobserved marginal cost, is negative in only 2.15% of observations.

The negative unobserved marginal cost estimates for the retailer are not necessarily an indictment of the demand-side model nor estimates in the beer category. There are several potential explanations for this surprising result. First, the negative unobserved marginal cost estimates could reflect cross-category incentives that are not modeled in the single-category demand system. If there are complementary goods to beer (ie, salty snack foods, disposable cups, and ping-pong balls) or if beer prices drive store traffic, then this would imply that Dominick's can gain additional cross-category profits from increasing its beer sales. In a 2013 interview with Bloomberg, a Walmart spokesperson claimed that "Grocers have long invested in [beer prices] to drive store traffic."[17] This explanation is consistent with the pricing of 24-packs as loss leaders in the Dominick's data. Second, the supply side of the model could be misspecified. For example, perhaps there is Nash bargaining between the retailer and the wholesalers, rather than the two-stage linear pricing analyzed here. Third, Dominick's may be under-reporting its markups above wholesale price.

I assume that cross-category incentives are the cause of these unexpectedly negatively signed estimates in my sample and proceed under the proposed two-stage linear pricing model. This assumption comes with the caveat that the retailer marginal costs derived from the first order conditions can no longer be interpreted as truly structural parameters. Instead, they must be interpreted as reduced form parameters from the solution to retailer's cross-category profit maximization problem. That is, the implied marginal costs computed here are unlikely to be truly constant in quantity (or price), although I will assume that they are for the remainder of the chapter.[18]

Table 1.5 presents results for the upstream manufacturers. The mean implied upstream markup (per 12oz) is 11.92¢, the median is 10.71¢, and the standard deviation is 3.53¢. The median implied upstream (manufacturer plus distributor) marginal cost is 34.48¢, the

[17]http://www.bloomberg.com/news/articles/2013-09-16/wal-mart-sells-coors-almost-at-cost-to-be-largest-beer-seller

[18]Smith and Thomassen (2015) present a model which illustrates how parameters recovered from the first order conditions can capture such cross-category incentives. I also provide a brief discussion of how cross-category incentives can be interpreted as reduced marginal costs in the Appendix B.

standard deviation is 7.58%, and none of the implied upstream marginal costs are negative. Upstream marginal cost represents a median of 60.59% of the retail price. If distributor marginal costs are small, then this is roughly consistent data from Tremblay and Tremblay (2005), who present a price-cost breakdown of mass-produced beers in 1996 which shows that manufacturer marginal costs account for approximately 58% of the retail price.

1.6 Counterfactual Simulations

1.6.1 Simulations and Results

In the counterfactual settings the retailer implements a pricing policy of price discrimination across stores. In all simulations, I assume that consumers are unable exploit the potential arbitrage opportunity present when some Dominick's stores charge lower prices than others. That is, consumers will not drive to a lower priced Dominick's store if their ZIP code's store has high prices. The retailer still sets prices weekly. The first order conditions for the retailer are the same as in Section 1.4.2 but now there are $N = 269,146$ distinct products, each of which appears in only one market. I first consider the case in which upstream firms do not respond to the retailer's policy change. This is analogous to holding the outcome of Stage 2 of the game described in Section 1.4 constant. I then consider the case in which upstream firms will respond to the retailer's policy change and compute a new Nash equilibrium of the two-stage linear pricing subgame. The upstream firms are assumed to continue charging a uniform wholesale price to all stores in the retail chain. The wholesale prices may still vary by product and over time. Mechanically, the upstream first order conditions do not change, except for the difference in retailer product definitions.

While it is straightforward to calculate changes in upstream and downstream profits, calculating the change in consumer welfare is more complex. With the nested logit model of demand, the change in consumer welfare within a store-week is computed as the compensating variation. For an individual, the compensating variation is computed as the change in expected maximum utility, scaled up to a dollar value by dividing by the marginal utility of income:

$$CV_{mt} = \frac{\ln\Big(\sum_g \big(\sum_{j \in g} \exp(V^{store}_{jmt}/(1-\sigma))\big)^{1-\sigma}\Big) - \ln\Big(\sum_g \big(\sum_{j \in g} \exp(V^{semi-chain}_{jmt}/(1-\sigma))\big)^{1-\sigma}\Big)}{\alpha_m}$$

The demand model assumes that all consumers within a market have the same marginal value of income (although it varies across markets), so the total change in consumer welfare within a market is computed by multiplying the compensating variation by the market size: $\Delta W_{mt} = M_m CV_{mt}$.

I first estimate the counterfactual setting without upstream adjustment (CF1). That is, the retailer switches to store-level pricing and the wholesale prices do not adjust. As expected, I find that prices decrease in some markets and increase in others. The welfare results are presented in Table 1.6. Values in the table are reported in real (1991) dollars per week across the 49 stores in the sample. Percentage changes are relative to when the retailer uses a semi-uniform pricing policy. The columns corresponding to CF1 present results without upstream adjustment to the retailer's new pricing policy. Downstream profits increase substantially by \$5,162 per week, or 11.36% over the sample. This estimate is most similar to one from Chintagunta et al. (2003), who apply a similar analysis in the categories of laundry detergent and refrigerated orange juice. They find an increase in retailer profit of 9.6% for the laundry detergent category. Interestingly, upstream profits in CF1 increase by \$624 per week, or

3.58%. This suggests that the upstream firms would desire the retailer to practice store-level pricing, assuming that the upstream firms do not incur additional costs due to the policy change. Together, these profit increases imply an increase of 9.2% for total profits within the vertical supply chain. Unlike both the upstream and downstream firms, consumers incur a loss in welfare. As a result of the policy change, consumer welfare decreases by $7,637, or 13.63%. So, while the policy change benefits firms, it is more harmful to consumers than it is beneficial to firms. The combined changes in vertical profits and consumer welfare imply a decrease in total welfare of 1.54%.

The second counterfactual experiment (CF2) allows upstream firms to adjust to the retailer's policy change. Wholesale prices generally increase. The median change in wholesale price across all products is an increase of 4%, with substantial heterogeneity across brewers. Miller sees the largest increases in wholesale prices, with a median of +7.5%. Anheuser-Bush sees the next largest change with a median of +2.1%, and Heilman has the smallest median change of +1.5%. It is, perhaps, unsurprising that the larger firms in the market exhibit larger adjustments to their wholesale prices. The columns in Table 1.6 corresponding to CF2 present welfare results for this counterfactual setting. Compared with CF1, the changes to all measures presented in Table 1.6 are fairly large. The downstream firm's profit gains are reduced to 3.5%, relative to the original pricing policy. Upstream profits increase as expected, and upstream firms now increase their profits by 11.45%, which is an additional 7.94%, relative to when they did not adjust wholesale prices. The pattern observed in wholesale price changes by brewer is also observed in the additional profit increases, with Miller benefitting the most from wholesale price adjustment and Heilman the least. In terms of dollar value, the upstream firms actually obtain a larger profit gain than the downstream firms in CF2. Vertical chain profits are also affected: relative to the initial equilibrium under the retailer's original pricing policy, the increase in vertical supply chain profits is more modest at 5.71%, as opposed to 9.20% in CF1. The decrease in consumer welfare is much larger in CF2 at 19.30%, and total welfare decreases by 6.23%.

1.6.2 Discussion

The large difference in counterfactual results between CF1 and CF2 highlights the importance of the researcher's choice of model for firm conduct along a vertical chain when performing counterfactual analyses. For example, the assumption of a highly competitive upstream and marginal cost pricing by upstream firms is often used in the industrial organization literature. This assumption is often made to validate demand instruments - such as wholesale prices - and also to simplify the numerical analysis in counterfactual settings. If this were the true data-generating model in my application, then the results from CF1 would be the appropriate focus for, say, a retailing consultant to supermarkets. However, if upstream firms have market power and the true data-generating model is two-stage linear pricing, then the results from CF2 would apply. In this particular context, the demand instruments are valid under either vertical pricing model for the reasons discussed in Section 1.4.1, but the counterfactual results are quite different. It is easy to conjecture that the researcher's choice of vertical pricing model can have a large effect on counterfactual predictions in other settings, such as analyses of downstream mergers or of changes to the regulatory environment of an industry. Ideally, the researcher would employ tests of firm conduct such as those developed by Villas-Boas (2007). When there is insufficient data to perform such tests, as is the case in this study, the choice of vertical structure should at least be well-motivated by institutional knowledge.

Some of the results from the counterfactual experiments in this study are an intuitive consequence of the increase in wholesale prices in CF2. The difference in counterfactuals' predictions for downstream profit arises because the higher wholesale prices in CF2 imply higher marginal costs for the retailer. The lower increase in total vertical profits in CF2 is also explained by the higher wholesale prices. Recall from Section 1.2 that in the double marginalization model, downstream profit maximization results in a retail price that is higher than the vertical profit-maximizing price. An increase in wholesale price will then result in a further reduction in vertical profits due to even higher retail prices. This same explanation applies to the additional consumer and total welfare decreases observed in CF2, relative to CF1. The higher retail prices resulting from higher wholesale prices harm the consumers, the downstream firm, and the vertical chain as a whole.

The difference in predicted welfare changes between the counterfactual experiments can provide some insight into why the retailer would set semi-uniform prices, as observed in the data. The managerial menu costs required to explain the retailer's observed semi-uniform pricing policy are much lower when upstream adjustment is taken into account (CF2) than when it is not (CF1). The difference is large: accounting for upstream adjustment reducing the required managerial menu costs by almost 70%. This large difference suggests the prospect of upstream adjustment as an important determinant of the retailer's choice of pricing policy. A possible caveat to this claim is that the upstream firms realize large profit increases in CF2, whereas their profit gains were more modest in CF1. This introduces a concern that the upstream firms will be willing to pay part or all of the menu costs associated with the more complex, store-level pricing policy. However, the strength of this concern is questionable for two reasons. First, the change in total vertical profits is relatively modest under CF2. So, while the upstream firms see a large percentage increase in their profits, the dollar value of that increase is not so large, relative to total vertical chain profits. Second, the manufacturers are competitors and this increase in profit is split amongst them. This introduces a coordination problem into a scheme to cover the retailer's managerial menu costs, as there would need to be a mechanism to determine how much of the increased profit would be paid for by each manufacturer.

The decrease to consumer welfare in both counterfactuals could also present an impediment to the retailer implementing a store-level pricing policy. As noted by Chintagunta et al. (2003), persistent losses in consumer welfare can cause store switching behavior in the long-run. This is in addition to the short-run cross-category effects captured in the negative unobserved retailer marginal cost estimates. The store-switching concern is present in the "naive" CF1 and is exacerbated in CF2 due to the much larger decrease in consumer welfare. The combined effect of the store-switching and managerial menu cost concerns could reasonably explain the retailer's decision to set uniform prices across most stores.

1.7 Conclusion

Retail chains often set uniform prices across stores, even though the consumers' preferences may vary substantially across locations. Using a dataset of weekly sales, retail prices, and wholesale prices, I observe one such retailer pricing uniformly across many locations. I explore the implications of upstream wholesale price adjustment for market outcomes when the retailer changes its pricing policy. I also evaluate the potential for upstream adjustment to explain the retailer's observed choice of semi-uniform pricing.

Using a structural model of demand and supply, I perform counterfactual experiments in which a retailer operating in multiple markets changes its policy from semi-uniform pricing to store-level pricing. I compute profit and welfare outcomes both with and without upstream wholesale price adjustment. I find that the upstream has significant incentive to adjust wholesale prices when the retailer's pricing policy changes and that this adjustment has a large effect on profits, consumer welfare, and total welfare. The upstream increases wholesale prices in response to the retailer's policy change, which decreases the retailer's profit gains by roughly $\frac{2}{3}$ and decreases total vertical profit gains by almost $\frac{1}{2}$. The upstream adjustment also decreases both consumer and total welfare by an additional 5% , relative to the case when wholesale prices are do not change.

Broadly, these results suggest that failure to account for upstream adjustment to policy changes in a downstream market can induce substantial inaccuracies in economic predictions, even when there is no change to the upstream competitive and regulatory environment. In the specific application to a retailer's pricing policy choice, the results suggest that upstream wholesale price adjustment can help explain the retailer's decision to charge prices with a large degree of uniformity across stores. The upstream adjustment limits the retailer's ability to increase profit through targeted, store-level pricing, suggesting that the retailer may not be willing to incur additional costs associated with a the more complicated pricing policy, such as hiring a consultant to determine optimal prices.

The supply-side estimates used in the counterfactual simulations suggest that the retailer may consider cross-category incentives in its optimal pricing decision for beer. A demand model that incorporates cross-category substitution/complementarity and a store choice decision could yield more accurate counterfactual predictions, but such an analysis would also require data which is not present in the Dominick's Database. Additionally, the analysis presented in this chapter does not rule out other potential explanations for the retailer's decision, such as additional demand-side reactions to a pricing policy change. Further research is needed to evaluate such explanations for uniform pricing.

Table 1.1: Products Included in Sample

Marketing Category	Brewer Miller	Anheuser-Busch	Heilman
Premium	Miller Genuine Draft, Miller High Life	Budweiser, Michelob Golden	Old Style, Old Style Classic
Super Premium	Lowenbrau	Michelob	Special Export
Light	Miller Lite	Bud Light	Old Style Light

Table 1.2: Summary Statistics

	Mean	Median	Std. Dev.	Min	Max
Product Variables:					
Retail Prices (cents per 12oz)	57.25	58.19	5.91	11.72	86.01
Wholesale Prices (cents per 12oz)	46.40	46.41	2.33	36.30	52.46
Product Market Share (%)	0.09%	0.04%	0.17%	0.001%	5.09%
Outside Option Share (%)	97.33%	97.96%	1.96%	80.89%	99.95%
$p - w < 0$ (%)	9.58%				
Promotion (%)	27.08%				
Demographic Variables:					
Income ($10,000)	4.23	4.21	1.15	2.13	7.56
Age$\geq$60 (%)	17.41%	17.55%	6.69%	5.81%	30.74%
Ability-to-Shop Index ($\times$100)	75.60	81.88	23.03	7.07	98.63

Product Variables: $N = 269,146$, Demographics: $N = 49$

Table 1.3: Results from Demand Estimation

Specification	Model 1 Logit	Model 2 Logit	Model 3 Nested Logit
Price (per 12oz)	-5.695	-5.264	-0.837
	(0.154)	(0.509)	(0.188)
Price$\times$Income/10000	0.336	0.536	0.263
	(0.036)	(0.081)	(0.035)
Price$\times$Age60	5.861	5.940	1.080
	(0.406)	(0.933)	(0.395)
Price$\times$ShopIndex	-3.802	-10.252	-4.726
	(0.184)	(0.417)	(0.199)
Within-Group Share (σ)			0.748
			(0.015)
Holiday	0.106	0.090	0.222
	(0.004)	(0.004)	(0.003)
12-pack dummy	0.388	0.285	0.057
	(0.004)	(0.012)	(0.007)
Promo	0.310	-0.047	-0.060
	(0.007)	(0.039)	(0.014)
Instruments	N	Y	Y
Median Own-Price Elasticity	-3.53	-5.79	-7.44
Std. Dev.	0.55	1.27	2.14
75th-25th %ile	0.77	1.45	2.19
Median $\frac{\partial \ln(s_{0m})}{\partial \ln(p_{jm})} \times 1000$	1.2	1.9	0.6

$N = 269,146$. Huber-White standard errors in parentheses

Table 1.4: Implied Downstream (Retailer) Markups, Margins, and Marginal Costs

	Mean	Median	Std. Dev.	Percent< 0	Median in Data
Markup (cents)	35.04	33.53	6.91		10.47
Margin (%)	62.95%	58.75%	15.03%		17.61%
Total Marginal Cost (cents)	21.40	23.70	9.22	2.15%	46.41
Unobserved MC (cents)	-24.57	-22.25	8.75	99.99%	

$N = 19,728$. "Median in Data" uses observed wholesale prices as the total marginal cost.

Table 1.5: Implied Upstream (Wholesaler) Markups, Margins, and Marginal Costs

	Mean	Median	Std. Dev.	Percent< 0	Median % of Retail Price
Markup (cents)	11.92	10.71	3.53		18.47%
Margin (%)	25.71%	22.94%	7.58%		
Marginal Cost (cents)	34.48	34.54	4.00	0%	60.59

$N = 6,221$

Table 1.6: Counterfactual Results

Counterfactual:	CF1 No Adjustment		CF2 Upstream Adjusts	
	Change ($)	Change(%)	Change ($)	Change (%)
Downstream Profit	$5,162	11.36%	$1,592	3.50%
Upstream profit	$624	3.58%	$2,000	11.46%
Vertical profit	$5,786	9.20%	$3,592	5.71%
Consumer Welfare	-$7,637	-13.63%	-$11,082	-19.30%
Total Welfare	-$1,851	-1.54%	-$7,490	-6.23%

Values are reported in real (1991) dollars per week, across all stores

Figure 1.1: Real Price (per 12oz) of Budweiser 12-Packs

Figure 1.2: Kernel Density Estimate of Median Income

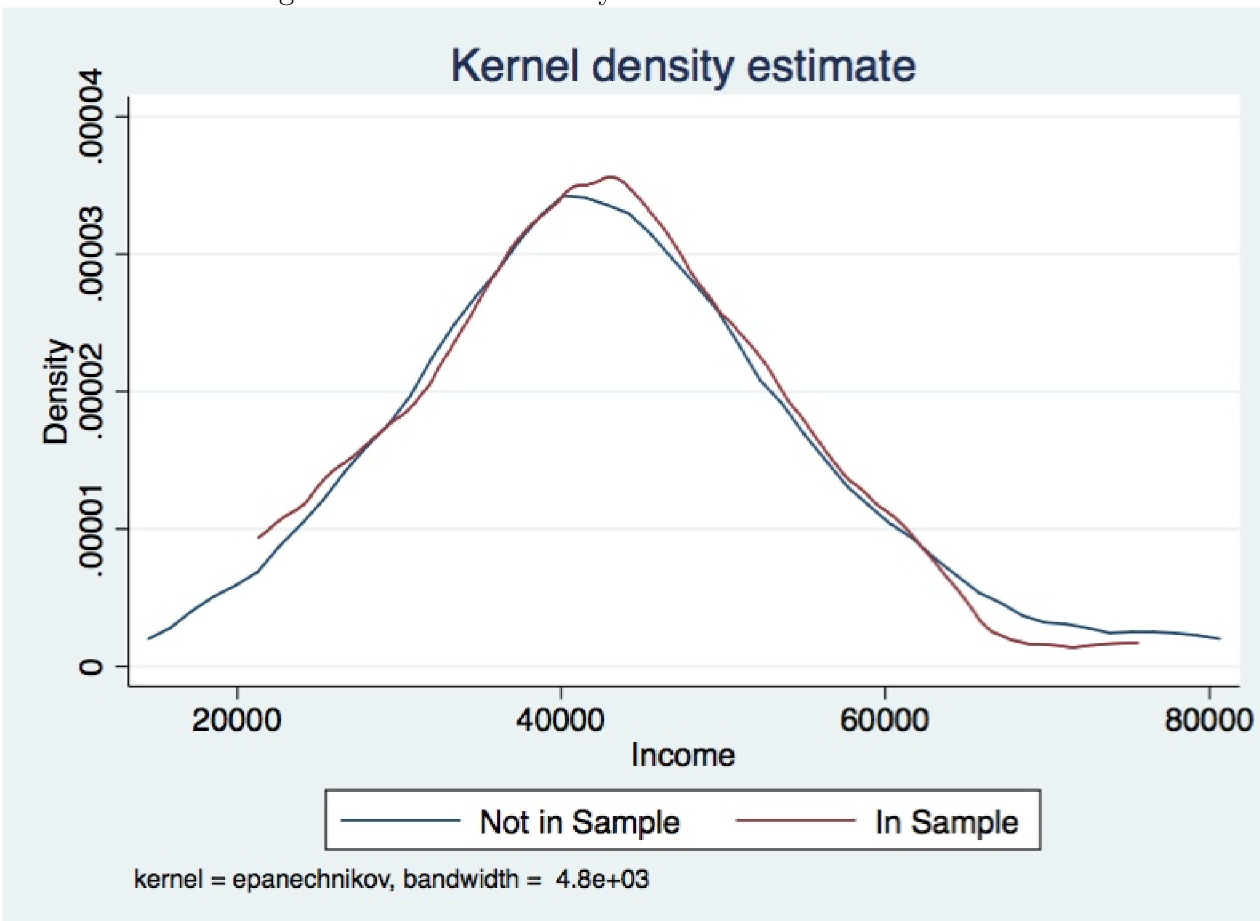
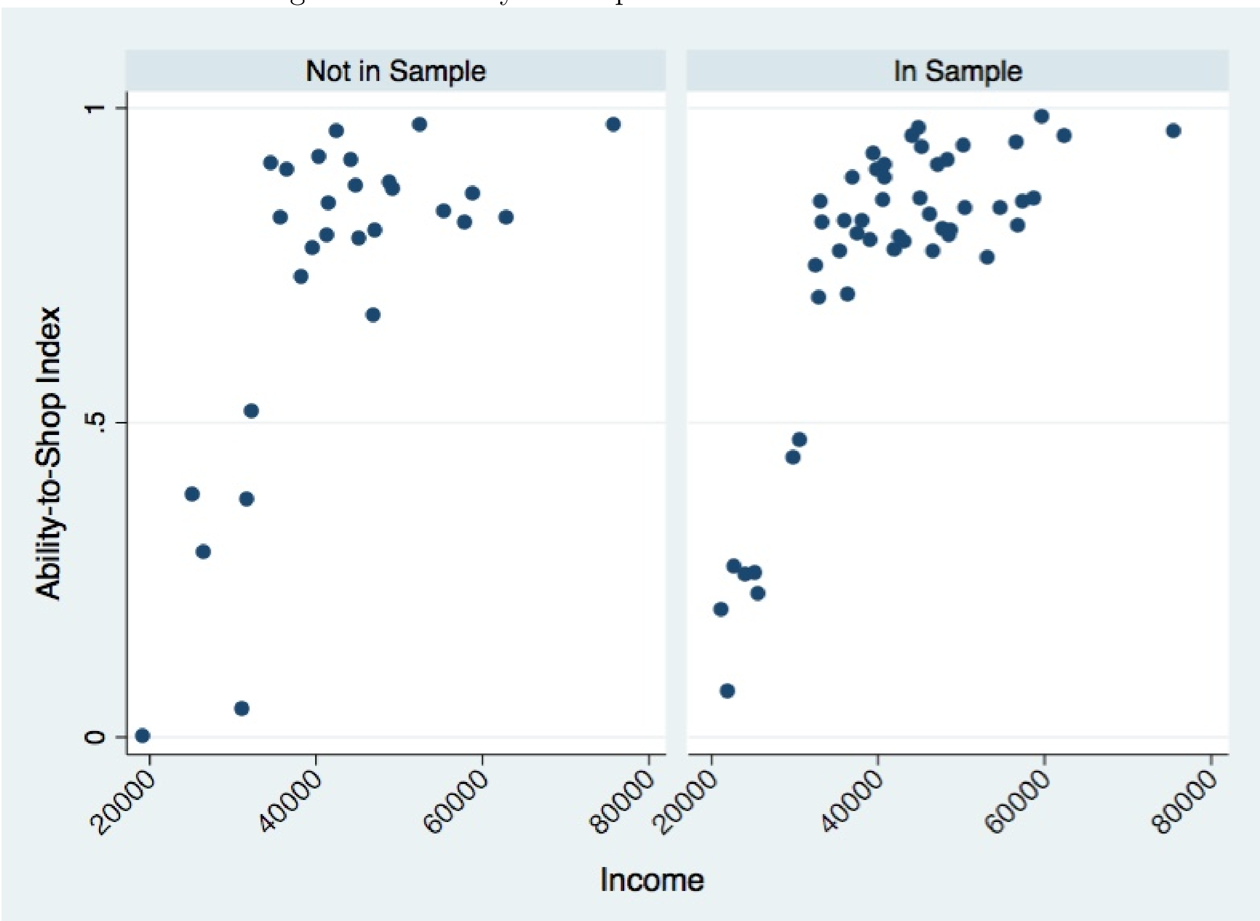


Figure 1.3: Ability-to-Shop Index vs. Median Income



CHAPTER II

Minimum Markups and Product Proliferation: Implications for Regulatory Policy

2.1 Introduction

The alcoholic beverage industry in the United States is heavily regulated, with federal, state, and sometimes local-level regulations imposed on both production and sales. In some states, these regulations include minimum markups imposed on retailers. For example, the state of Michigan requires the listed retail price of liquor to be marked up no less than approximately 17.9% higher than the wholesale price paid by the retailer.[19] In Michigan, the state also sets the wholesale prices paid by retailers, and so the minimum markup effectively imposes a price floor on all retailers in the state. These price floors often bind, with many liquor distributors advertising state minimum prices.[20]

Proponents of minimum markups primarily make a few - sometimes contradictory - claims

[19]Source: Price book from the Michigan Liquor Control Commission. http://www.michigan.gov/lara/0,4601,7-154-10570_14173—,00.html. The actual amount of the markup varies slightly but is generally within 0.1% of the stated amount.

[20]Other states with minimum retail markups for alcohol include Ohio, Connecticut, and Wisconsin. Washington required a minimum markup on liquor until 2012.

about the resulting outcomes. The first is lower prices via increased competition. That is, without the minimum markups in place, so-called “big box” stores will be able to temporarily engage in predatory pricing, drive their competitors out of business, and then use their increased market power to charge even higher prices. However, such predatory pricing strategies usually are not viable for raising profits (Viscusi et al., 2005). Many proponents of the policy agree that minimum markups result in higher prices and that this is beneficial from a social standpoint. There are significant negative externalities resulting from alcohol consumption, such as drunk driving-related motor vehicle collisions and increased healthcare costs. Higher prices can reduce the quantity of alcohol consumed, resulting in an increase in social welfare. Opponents of minimum markups tend to argue that such laws are harmful to consumers (due to increased prices) and are inefficient, inducing excess entry.[21][22]

Essentially, the economic arguments of both proponents and opponents of minimum markup laws rely on predictions about the laws' effects on prices, quantities, and entry, and on how these in turn affect consumer and total welfare. In a homogenous-goods market, these issues are straightforward to deal with using standard economic theory. If the minimum markup is binding, then it results in higher prices and reduces aggregate quantity, consumer welfare, and total welfare. In an oligopoly, the higher prices will also result in higher profits for firms, inducing entry. Examples of empirical studies of minimum markup laws in the market for a homogenous good (gasoline) are provided in Skidmore et al. (2005) and Anderson and Johnson (1999). However, such analysis may not be applicable to the market for alcohol because it is a differentiated products market with multi-product firms. It is therefore also

[21] Opponents also argue that these laws are paternalistic (teetotaling), although this argument is not economically relevant.

[22] Sullum (2016) provides several quotes from policymakers in Connecticut: https://www.forbes.com/sites/jacobsullum/2016/02/25/fearing-monsters-like-total-wine-liquor-merchants-want-connecticut-to-keep-prices-high/.

important to consider a minimum markup's effect on product proliferation, i.e. the retailers' selection of products to sell to consumers, because there is a "taste for variety" in the market for liquor, which can arise either from individual consumers desiring to consume from a diverse set of products or from heterogeneity in consumers' tastes for individual products.

This chapter examines the effects of minimum markup laws on prices, quantities, and welfare, taking into account firms' endogenous product variety decision.[23] The primary contribution is to present a model in which endogenous product variety has substantial implications for the economic impact of a minimum markup. The results from the model show that a minimum markup will induce firms to expand product variety, which in turn affects aggregate consumption, welfare, and the incentives for entry/exit (via its effect on firm profits). The purpose of the analysis here to show that the changes in product variety induced by minimum markups can indeed have a significant effect on the economic outcomes of interest, rather than to characterize the exact conditions that govern the size of the effects. Such an analysis is left to future work.

Section 2.2 presents the model of demand and supply. The demand model is the multinomial logit model, which is a member of a class of models commonly used in the industrial organization literature. Firms compete Nash-Bertrand on prices and product variety, and I characterize the equilibrium. Section 2.3 considers the effects of the introduction of a minimum markup with multinomial logit demand. Section 2.4 discusses the analysis and concludes.

[23] In this chapter, I use the term "variety" to refer to the number of products offered by each firm and abstract away from the issue of selection of products with specific characteristics.

2.2 Multinomial Logit Model

In this section, I present a version of the multinomial logit model of demand for differentiated products (McFadden, 1974; Berry, 1994). I fix the set of firms and allow firms to choose the number of products offered, along with the price of each of their products. I restrict the analysis to products that are "identical" in their non-price characteristics.

2.2.1 Demand

Suppose that there are $R \geq 1$ firms (retailers), indexed r, each selling $N_r \geq 1$ products. There are M consumers, each purchasing at most one unit. A consumer's utility from purchasing and consuming product j is

$$u_{ji} = \delta_j + \varepsilon_{ij} = \beta - \alpha p_j + \varepsilon_{ij}$$

and the utility of the outside option (no purchase) is $u_{io} = \varepsilon_{io}$. The products are "identical" in the sense that differences in their mean utilities, δ_j, can be exclusively attributed to differences in prices. Furthermore, I assume that each product available for purchase is sold by only one firm, although it is not yet necessary to index each product to its firm. Using this specification of utility, each product's market share can be expressed analytically as

$$s_j = \frac{\exp(\delta_j)}{1 + \sum_{k \neq 0} \exp(\delta_k)}$$

and the total quantity of product j consumed is then Ms_j.

It is useful at this point to consider a special case. Suppose that all firms produce the same number of products and set the same price for all products. That is, $N_r = N$ and $p_j = p$, so that $\delta_j = \delta$. Then in this case, all products will have identical shares equal to

$$s = \frac{\exp(\delta)}{1 + RN\exp(\delta)} = \frac{\exp(\delta)}{1 + N\exp(\delta) + (R-1)N\exp(\delta)}$$

In this symmetric case, each firm's share of the market is Ns, which is decreasing in R but increasing in N. Total market share of all firms is RNs, which is increasing in both R and N. It is clear that adding additional products for one firm (or all firms) will decrease the market share of each individual product, including that firm's own products. So, this model includes a cannibalization effect of variety expansion. Additionally, increasing the number of firms reduces the market shares for all existing firms' products. The fact that additional firms will capture some of every existing firm's demand differentiates this model from address models, such as Salop (1979), in which entrants will essentially only directly compete with firms that are locationally adjacent.

Finally, it is useful to define the expected utility for a consumer:

$$V = \ln(1 + RN\exp(\delta)) = \ln\left(\frac{1}{s_0}\right) = -\ln(1 - RNs).$$

This is an increasing function of the total market share of all firms, RNs. Because aggregate consumption is $MRNs$ and M is fixed, (expected) aggregate consumption is therefore a sufficient statistic for V.

2.2.2 Profit Maximization without Minimum Markups

I now consider an equilibrium in prices and the number of products sold. Here, there is no minimum markup and competition amongst firms is Nash-Bertrand. I assume that firms have a constant per-unit marginal cost, c, and a development (or shelving) cost for additional products, k. Because the per-unit marginal cost is the same for all products, I restrict attention to the firm setting a single price for all of its products. If Θ_r is the set of products sold by firm r, then the firm's profit maximization problem can be written as

$$\max_{p,N} \quad MNs_r(p-c) - Nk$$

$$s.t. \quad s_r = \frac{\exp(\beta - \alpha p)}{1 + N\exp(\beta - \alpha p) + \sum_{k \notin \Theta_r} \exp(\delta_k)}.$$

The firm treats δ_k as fixed for all $k \notin \Theta_r$. In a slight abuse of notation, the term s_r is the market share of each individual product produced by firm r. Although N should generally take a discrete value, I treat it as a continuous variable for the purposes of this analysis. One interpretation may be that N is large, so that adding (or subtracting) a product is a relatively small change and analysis based on continuous N can approximate the discrete case. The constraint can be substituted directly into the objective function, and the first order condition for pricing is therefore

$$MNs_r + MN\frac{\partial s_r}{\partial p}(p-c) = 0$$

and the first order condition for the number of products (variety) is

$$Ms_r(p-c) + MN\frac{\partial s_r}{\partial N}(p-c) - k = 0.$$

The intuition of the pricing equation is standard. The intuition of the variety equation is that marginal revenue minus marginal cost equals zero, and the middle term in the above equation represents the cannibalization effect of adding a new product. These two equations can be further simplified by defining $\phi = p - c$ and by noting that the analytic expression for the partial derivatives of s_r are

$$\frac{\partial s_r}{\partial p} = -\alpha s_r(1 - N s_r)$$

$$\frac{\partial s_r}{\partial N} = -s_r^2.$$

The two first order conditions then reduce to

$$\alpha(1 - N s_r)\phi = 1$$

$$s_r(1 - N s_r) - \frac{k}{M\phi} = 0$$

for price and product variety, respectively. Additionally, it is fairly straightforward to show that the second order condition is satisfied so long as $1 - 2Ns_r > 0$, i.e. no firm captures more than half the potential market.[24]

2.3 Effects of a Minimum Markup

I will now consider the effect of imposing a minimum markup on the equilibrium number of products, total consumption, and welfare. The analysis proceeds in reference to a symmetric

[24]This is a well-known condition for existence of an equilibrium in pricing (Aksoy-Pierson et al., 2013). However, it is somewhat surprising that this also ensures that marginal revenue as a function of variety to be downward sloping as well., which results in an equilibrium in both prices and the number of products.

Nash equilibrium in which all firms choose (p^*, N^*) to satisfy the first order conditions presented in the previous section. This can also be characterized as (ϕ^*, N^*) because ϕ is a monotonic function of p when c is fixed. I therefore also drop the subscript on the market share, so that $s_r = s$. I consider the effect of a small, binding increase in ϕ for all firms above the level ϕ^*. With c fixed, an increase in ϕ must be accomplished by increasing p. I will therefore, in a slight abuse of notation, write some derivatives with respect to p as derivatives with respect to ϕ instead.

In this section, I will consider *all* firms increasing markups identically (prices). So, it is necessary to define a derivative:

$$\frac{\partial s}{\partial \phi} = -\alpha s(1 - RNs)$$

which represents a decrease that is smaller in magnitude than $\frac{\partial s_r}{\partial p}$ because the other firms raising their prices will increase an individual firm's share, partially offsetting the effect of the firm increasing its own prices. I will also consider *all* firms adjusting their number of products identically, so that

$$\frac{\partial s}{\partial N} = -Rs^2$$

which is a larger decrease than $\frac{\partial s_r}{\partial N}$ in the previous section because the other firms adding varieties further decreases the share for an individual product.

2.3.1 Effect on Product Variety

As discussed in the introduction, the central questions of this chapter are focused on the impact of endogenous product variety on the model's predictions. Therefore, it is necessary to determine how the introduction of a minimum markup influences firms' choice of product variety. In particular, I am interested in the sign and magnitude of

$$\left.\frac{dN}{d\phi}\right|_{(\phi^*,N^*)}$$

which determines whether the small, binding minimum markup will increase or decrease the number of varieties available from each firm. I find that this term is positive, so that the minimum markup provides firms with an incentive to offer more varieties. The expression for the term of interest is provided in the following proposition.

Proposition 3. *Relative to the equilibrium with no minimum markup,* (ϕ^*, N^*) *the introduction of a binding minimum markup,* ϕ^{min}*, increases the number of products offered by firms when* $\phi^{min} - \phi^*$ *is small. Additionally, the derivative of the number of products with respect to the markup is*

$$\left.\frac{dN}{d\phi^{min}}\right|_{(\phi^*,N^*)} = \alpha N \left[\frac{R(1-2Ns)+Ns}{R(1-2Ns)+1}\right] > 0$$

Proof. Provided in Appendix A. □

2.3.2 Effect on Aggregate Consumption

The effect of a binding minimum markup on aggregate consumption is of particular importance for reasons discussed in the introduction. The minimum markup results in increased prices, which will decrease aggregate consumption if the number of products does not change. However, I showed in the previous section that the introduction of a minimum markup induces firms to offer more products in equilibrium. This has a countervailing effect of increasing aggregate quantity and motivates the question of interest in this section of the chapter: how much does the effect of increasing variety offset the effect of increasing prices on aggregate consumption? Specifically, I answer this question by comparing the predicted effect of introducing a minimum markup on Ns when the number of products is exogenous (fixed) to the size of the effect when the number of products is endogenous.[25]

The case in which the number of products is fixed is quite straightforward. In this case, the object of interest is

$$N\frac{\partial s}{\partial \phi} = -\alpha Ns(1 - RNs),$$

which is negative. However, the analysis is more complicated when the number of products is endogenous. In this case, the derivative of interest is

$$\left.\frac{d(Ns)}{d\phi}\right|_{(\phi^*,N^*)} = N\frac{ds}{d\phi} + \frac{dN}{d\phi}s.$$

Most important for this analysis are the overall sign of this derivative and its size, relative to $N\frac{\partial s}{\partial \phi}$. These are characterized in the next two propositions.

Proposition 4. *Relative to the equilibrium with no minimum markup, (ϕ^*, N^*) the introduction of a binding minimum markup, ϕ^{min}, decreases aggregate consumption when $\phi^{min} - \phi^*$ is small. Additionally,*

$$\left.\frac{d(Ns)}{d\phi^{min}}\right|_{(\phi^*,N^*)} = -\alpha Ns(1 - RNs)\left[\frac{1 - Ns}{R(1 - 2Ns) + 1}\right] < 0$$

Proof. Provided in Appendix A. □

Proposition 5. *When the number of products is endogenous, the derivative of Ns with*

[25] Aggregate consumption is $MRNs$. With M and R fixed, the relevant changes are in Ns.

respect to ϕ^{min} evaluated at (ϕ^*, N^*) is bounded by:

$$0 > \left.\frac{d(Ns)}{d\phi^{min}}\right|_{(\phi^*,N^*)} > \frac{1}{R}\left(N\left.\frac{\partial s}{\partial \phi^{min}}\right|_{(\phi^*,N^*)}\right).$$

Furthermore, if $R = 1$, then

$$\left.\frac{d(Ns)}{d\phi^{min}}\right|_{(\phi^*,N^*)} = \frac{1}{2}\left(N\left.\frac{\partial s}{\partial \phi^{min}}\right|_{(\phi^*,N^*)}\right).$$

Proof. Provided in Appendix A. □

Propositions 4 and 5 have some interesting implications for the effect of product line expansion on aggregate consumption. It would potentially be very concerning for policymakers if a minimum markup actually increased aggregate consumption, as one of the goals of such a policy is often to reduce consumption. Proposition 4 alleviates that particular concern by showing that the variety expansion effect on aggregate consumption does not completely offset effect of higher prices. However, Proposition 5 suggests that the variety expansion effect is still quite large - it undoes at least half of the effect of increased prices. This result highlights the importance of accounting for endogenous product variety. If there is a monopoly and product variety is mistakenly taken as exogenous, then the predicted effect of a minimum markup on aggregate consumption will be approximately twice as large as the true effect. For an oligopoly, the proportionate difference is at least as large as the number of firms.

2.3.3 Effect on Firm Profits

One of the arguments used in support of (and against) a minimum markup is that it can increase firms' profits. In this section, I confirm that this effect is present. I show this by evaluating the change in firm profits when a binding minimum markup is introduced. I still consider the symmetric equilibrium, so an individual firm's profits are

$$\Pi_r = MNs\phi - Nk.$$

The next proposition evaluates this derivative at the initial equilibrium.

Proposition 6. *Relative to the equilibrium with no minimum markup, (ϕ^*, N^*) the introduction of a binding minimum markup, ϕ^{min}, increases firm profits when $\phi^{min} - \phi^*$ is small and there is more than one firm operating in the market. Additionally,*

$$\left.\frac{d\Pi_r}{d\phi^{min}}\right|_{(\phi^*,N^*)} = M(Ns)^2 \left[\frac{R-1}{R(1-2Ns)+1}\right].$$

Proof. Provided in Appendix A. □

The derivative given by Proposition 6 is equal to zero when there is a monopoly ($R = 1$), which is to be expected. Proposition 4 shows that the small minimum markup raises the profits for firms in an oligopoly. This result is intuitive and arises from standard economic theory: in the unconstrained equilibrium, firms set prices at a level that is too low to maximize industry profits. Therefore, forcing all firms to raise prices will raise their profits. This also implies that the minimum markup is anticompetitive in the sense that it is as if the firms are forced to engage in a small amount of price collusion.

Although I fix the number of firms in this analysis, Proposition 6 does have some implications for a model with an endogenous number of firms (i.e., free entry/exit). Such a model would usually include an additional fixed cost of operation and a zero profit condition for equilibrium. Proposition 4 shows the minimum markup will increase firm profits, conditional on the number of firms. It is also reasonable to conjecture that per-firm profit should be decreasing in the number of firms. So, Proposition 4 suggests that introducing a minimum markup will induce entry, resulting in a greater number of firms operating in the market.

This analysis of the effect of minimum markups on firm profits has so far incorporated endogenous product variety. But how much of a difference does accounting for endogenous product variety make in the analysis? Alternatively, how large of an error is made if the researcher assumes that firms do not adjust their number of products? The next proposition will help answer this question.

Proposition 7. *When the number of products is endogenous and $R > 1$ (oligopoly), the derivative of Π_r with respect to ϕ^{min} evaluated at (ϕ^*, N^*) is bounded by:*

$$0 < \left. \frac{d\Pi_r}{d\phi^{min}} \right|_{(\phi^*, N^*)} < \frac{1}{R} \left(\left. \frac{\partial \Pi_r}{\partial \phi^{min}} \right|_{(\phi^*, N^*)} \right).$$

Furthermore, if $R = 1$, then

$$\left. \frac{d\Pi_r}{d\phi^{min}} \right|_{(\phi^*, N^*)} = \frac{1}{2} \left(\left. \frac{\partial \Pi_r}{\partial \phi^{min}} \right|_{(\phi^*, N^*)} \right) = 0.$$

Proof. The proof is similar to that of Proposition 5 and is omitted. □

Proposition 7 shows that failing to account for endogenous product variety in oligopoly will result in an over-prediction of the change in retailer profits and that the proportionate over-prediction is likely to grow with the number of firms. At a glance, this result is, perhaps, surprising. When oligopolistic firms are constrained to a fixed number of products, forcing them to increase their markups will intuitively increase profits for the reason previously discussed. It seems as if allowing them a margin of adjustment would increase profits, although this turns out to be false. This somewhat counterintuitive result arises because firms are already producing "too many" varieties. It is fairly straightforward to show that if $R > 1$, then

$$\left.\frac{\partial \Pi_r}{\partial N}\right|_{(\phi^*, N^*)} < 0.$$

Intuitively, when a firm considers expanding its product variety it does not consider the negative "business stealing" effect on its rivals' profits. This leads to excessive variety. When the minimum markup is imposed, all firms expand their product line, which further reduces industry profits and partially offsets the increase in profits from all firms raising prices.

Proposition 7 has further implications for the effect of introducing a minimum markup on the number of firms in model with free entry/exit. An analysis that omits endogenous product variety will instead over-predict the increase in profits, conditional on the number of firms. So, it is likely to also over-predict entry into the industry, instead of more limited entry, along with a larger number of varieties offered by each firm (both entrants and incumbents in this model).

2.3.4 Effect on Welfare

While the effect on aggregate consumption is important to policymakers, the effect of a policy change on welfare is also of great importance. Because there is a cost incurred from adding a variety, which is independent of how much of the variety is actually sold, the change in aggregate consumption is not a sufficient statistic for the change in total welfare.

Define total welfare in a symmetric equilibrium as follows:

$$W = M\frac{V}{\alpha} + R\Pi_r.$$

The first term is the total expected utility to consumers divided by the marginal utility value of a dollar. Changes in this term capture the compensating variation, which is used to measure changes in consumer welfare. The second term is industry profits. In this section, I examine the change in W when a small, binding minimum markup is introduced. The two main questions are: 1) how does the introduction of the minimum markup affect welfare when variety is endogenous; and 2) how does the prediction differ if product variety is fixed?

The previous sections of the chapter allow for a qualitative analysis of the impact of a minimum markup on each of the two terms in W separately. As discussed in Section 2.2.1, the change in aggregate consumption is a sufficient statistic for the change in V. So, the results from Section 2.3.2 are used to predict the change in V, which describes the change in consumer welfare. Introducing the minimum markup reduces aggregate consumption, even when product variety is endogenous. So, consumer welfare will fall. However, if the analysis ignores the induced increase in product variety, then it will over-predict the decrease in aggregate consumption and will also over-predict the decrease in consumer welfare as a result. For an oligopoly, Section 2.3.3 showed that the introduction of a minimum markup will increase the firms' profits and that the increase is larger if product variety is fixed.

Because imposing a minimum markup will decrease consumer welfare but increase firm profits, the effect of a minimum markup on total welfare depends on the relative sizes of these changes. The next proposition describes the result.

Proposition 8. *When the number of products is endogenous, the derivative of Ns with respect to ϕ^{min} evaluated at (ϕ^*, N^*) is bounded by:*

$$0 > \left.\frac{dW}{d\phi^{min}}\right|_{(\phi^*,N^*)} > \frac{1}{R}\left(\left.\frac{\partial W}{\partial \phi^{min}}\right|_{(\phi^*,N^*)}\right).$$

Furthermore, if $R = 1$, then

$$\left.\frac{dW}{d\phi^{min}}\right|_{(\phi^*,N^*)} = \frac{1}{2}\left(\left.\frac{\partial W}{\partial \phi^{min}}\right|_{(\phi^*,N^*)}\right).$$

Proof. Using results from previous propositions and analysis in the chapter, it is fairly straightforward to show that the derivatives of interest, evaluated at (ϕ^*, N^*) are

$$\frac{\partial W}{\partial \phi} = -MRNs(1 - RNs)\frac{1}{1 - Ns}$$

$$\frac{dW}{d\phi} = -MRNs(1 - RNs)\frac{1}{(1 - RNs) + R(1 - Ns)}$$

and the remainder of the proof proceeds as that of Proposition 3. □

Proposition 8 shows that the introduction of a minimum markup will decrease total welfare both with and without endogenous product variety. The decrease is smaller in magnitude once firms adjust their number of products. It is, perhaps, unsurprising that welfare decreases, even with endogenous product variety. With a fixed number of products, standard economic theory predicts that increasing prices above the oligopolistic equilibrium level will decrease welfare. When the number of products adjusts, the previous sections have generally shown that even if there is a countervailing effect from increasing the number of products, it will still not completely offset the effect of increased prices. However, it is less clear why firms' product variety adjustment countervails - rather than amplifies - the welfare effect of increased prices.

Consider an increase in the number of products offered by all firms around the initial equilibrium, (ϕ^*, N^*). It is straightforward to show that

$$\left.\frac{\partial W}{\partial N}\right|_{(\phi*, N^*)} = MRs(1 - RNs)\phi > 0.$$

So, the initial equilibrium features an inefficiently low number of products. The intuition behind this result is standard: quantities are inefficiently low if firms have market power. Here, the inefficiently low quantity manifests as both inefficiently high prices, conditional on the number of products, and also an inefficiently low number of products.

2.4 Discussion and Conclusion

This chapter presents a first effort to examine the effect of a minimum markup on firms' product variety. The results are derived for a stylized version of a non-address model of multi-product oligopoly competition with differentiated products commonly used in the industrial organization literature. I show that the introduction of a small, binding minimum markup induces firms to offer a greater number of products in equilibrium. I show that the minimum markup results in decreased aggregate consumption, increased firm profits (implying entry), and decreased welfare both with and without endogenous product variety but that all of these changes are attenuated when variety is endogenous, and that the attenuation is more severe with a greater number of firms. These results suggest that accounting for endogenous product variety is important in determining the short- and long-run economic impact of a minimum markup.

While the analysis in this chapter is informative, it is also subject to some caveats. First, the simple multinomial logit model presented here imposes several restrictions on demand and competition, which could be relaxed in future work. For example, this model implies that if a firm raises the prices of one of its products, then consumers are equally likely to buy a different product from that firm as they are to buy a different product from another firm. One way to ameliorate this particular issue would be to use a nested logit model (McFadden, 1978; Berry, 1994), with firm-level nests. Other well-known restrictions on substitution patterns also apply. Furthermore, products are only differentiated through prices and logit errors in this model. So, future work could consider heterogeneity in the mean utilities (β_j) or in costs. More broadly, alternative models of demand or competition, such as address models, may yield different predictions.[26]

Despite these concerns, this chapter presents an important first step by presenting a set of results that are relevant to an important and frequently-debated topic in the regulation of "sin" goods. Most importantly, it shows that simple price elasticities may not be sufficient to characterize the economic impact of a minimum markup imposed on firms if consumers have a taste for variety in the market. It also shows that analyses which fail to account for endogenous product variety may result in relatively large over-predictions for the effects of a minimum markup on the economic variables of interest to policymakers. An analysis that includes more flexible patterns for the impact of product variety on aggregate consumption, business stealing, and cannibalization would be a natural next step and is left to future work.

[26] Hamilton (2009) uses an address model of competition in his analysis of excise taxes in multi-product oligopoly.

APPENDICES

APPENDIX A
Proofs of Propositions

Proof of Proposition 1

It is first useful to first define some terms (with some abuse of notation):

$$\varepsilon_w = -\frac{dQ}{dw} \cdot \frac{w}{Q} > 0$$

$$s_i = \frac{q_i}{Q} > 0$$

$$\varepsilon_i = -\frac{dq_i}{dp_i} \cdot \frac{p_i}{q_i} > 0$$

$$\kappa_i = \frac{dp_i}{dw} \cdot \frac{w}{p_i} > 0$$

where $p_i = p_i^*(w)$ for $i \in \{1, 2\}$, the retailer's optimal price in market i, and $q_i = q_i(p_i^*(w))$. Also, $Q = q_1 + q_2$. So, optimal pricing by the retailer is assumed in these terms. Constant elasticity of demand implies that ε_i is constant in both markets, and without loss of generality I assume that $\varepsilon_1 < \varepsilon_2$ so that $p_1 > p_2$ when the retailer price discriminates and s_1 (s_2) is lower (higher) under price discrimination than under uniform pricing. Using these definitions and assumptions, I now provide some useful results.

Claim 1. Under price discrimination, $\kappa_1 = \kappa_2 = 1$.

Proof. This follows immediately from the first order condition of the retailer's optimal pricing decision in each market

$$p_i(1 - \frac{1}{\varepsilon_i}) = w$$

$$\Rightarrow \frac{p_i}{w} = 1 - \frac{1}{\varepsilon_i}$$

where ε_i is constant. □

Claim 2. Under uniform pricing, $\kappa_1 = \kappa_2 < 1$.

Proof. Under uniform pricing $p_1 = p_2 = p$, implying that $\kappa_1 = \kappa_2 = \kappa$. Consider the first order condition describing the retailer's optimal price under uniform pricing:

$$p(1 - \frac{1}{\tilde{\varepsilon}}) = w$$

where $\tilde{\varepsilon} = s_1\varepsilon_1 + s_2\varepsilon_2$ is the price elasticity of aggregate demand. Totally differentiating the first order condition yields

$$\frac{dp}{dw}(1 - \frac{1}{\tilde{\varepsilon}}) - \frac{p}{\tilde{\varepsilon}^2}\frac{d\tilde{\varepsilon}}{dw} = 1$$

$$\Rightarrow \kappa - \frac{p}{\tilde{\varepsilon}^2}\frac{d\tilde{\varepsilon}}{dp}\frac{dp}{dw} = 1$$

$$\Rightarrow \kappa - \frac{\frac{d\tilde{\varepsilon}}{dp}p^2\kappa}{w\tilde{\varepsilon}^2} = 1$$

$$\Rightarrow \kappa = \frac{w\tilde{\varepsilon}^2}{w\tilde{\varepsilon}^2 - \frac{d\tilde{\varepsilon}}{dp}p^2}$$

and the final step is then to show that $\frac{d\tilde{\varepsilon}}{dp} < 0$. Individual market elasticities are constant, so the expression of interest is given by

$$\frac{d\tilde{\varepsilon}}{dp} = \frac{ds_1}{dp}\varepsilon_1 + \frac{ds_2}{dp}\varepsilon_2 = \frac{ds_2}{dp}(\varepsilon_2 - \varepsilon_1)$$

where $\varepsilon_2 - \varepsilon_1 > 0$ by assumption and $\frac{ds_1}{dp} + \frac{ds_2}{dp} = 0$ because $s_1 + s_2 = 1$. It is straightforward to show that

$$\frac{ds_2}{dp} = \frac{1}{Q}(\frac{dq_2}{dp} - s_2(\frac{dq_1}{dp} + \frac{dq_2}{dp}))$$

$$= \frac{1}{pQ}(-\varepsilon_2 q_2 + s_2(\varepsilon_1 q_1 + \varepsilon_2 q_2))$$

$$= \frac{1}{pQ}(s_2 q_1 \varepsilon_1 - s_1 q_2 \varepsilon_2)$$

$$= \frac{s_1 s_2}{p}(\varepsilon_1 - \varepsilon_2) < 0$$

□

These claims imply that the elasticity of interest can be written as

$$\varepsilon_w = (s_1\varepsilon_1 + s_2\varepsilon_2)\kappa$$

where both κ and s_2 (s_1) are larger (smaller) under price discrimination than under uniform pricing. So, ε_w is larger under price discrimination than uniform pricing for any w. The upstream firm's optimal wholesale price will therefore be lower under price discrimination than under uniform pricing.

Proof of Proposition 3

The first statement follows from the second. For the second statement, the expression of interest is positive because $1 - 2Ns > 0$ is required for the firms' profit maximizing second order condition to be satisfied at (ϕ^*, N^*). Deriving the the actual expression requires total differentiation of the equilibrium first order condition on product variety with respect to ϕ. This yields

$$\left(\frac{\partial s}{\partial N}\frac{dN}{d\phi} + \frac{\partial s}{\partial \phi}\right)(1 - Ns) - s\left(\frac{dN}{d\phi}s + N\left(\frac{\partial s}{\partial N}\frac{dN}{d\phi} + \frac{\partial s}{\partial \phi}\right)\right) + \frac{k}{M\phi^2} = 0$$

which can be rearranged into

$$\frac{dN}{d\phi}\left[(1 - 2Ns)\frac{\partial s}{\partial N} - s^2\right] + \frac{\partial s}{\partial \phi}(1 - 2Ns) + \frac{k}{M\phi^2} = 0.$$

Substituting in the appropriate partial derivatives gives

$$-\frac{dN}{d\phi}s^2\left[(1 - 2Ns)R + 1\right] - \alpha s(1 - RNs)(1 - 2Ns) + \frac{k}{M\phi^2} = 0.$$

At (ϕ^*, N^*), the profit-maximizing first order conditions for the firm imply that $\frac{k}{M\phi^2} = \alpha s(1 - Ns)^2$. By substitution, we obtain

$$\alpha s\left[(1 - Ns)^2 - (1 - RNs)(1 - 2Ns)\right] = \frac{dN}{d\phi}s^2\left[R(1 - 2Ns) + 1\right].$$

The left-hand side can be simplified using basic algebra to obtain

$$\alpha Ns^2\left[R(1 - 2Ns) + Ns\right] = \frac{dN}{d\phi}s^2\left[R(1 - 2Ns) + 1\right]$$

and the result follows immediately.

Proof of Proposition 4

The first statement follows from the second. The derivative can by found by taking,

$$N\frac{ds}{d\phi}+\frac{dN}{d\phi}s=N\left(\frac{\partial s}{\partial\phi}+\frac{\partial s}{\partial N}\frac{dN}{d\phi}\right)+\frac{dN}{d\phi}s$$

$$=\frac{dN}{d\phi}\left(s+N\frac{\partial s}{\partial N}\right)+N\frac{\partial s}{\partial\phi}$$

$$=\frac{dN}{d\phi}\left(s-RNs^2\right)-\alpha Ns(1-RNs)$$

$$=s(1-RNs)\left(\frac{dN}{d\phi}-\alpha N\right).$$

This expression is negative at (ϕ^*, N^*) because the formula from Proposition ???? shows that $\frac{dN}{d\phi}<\alpha N$ at (ϕ^*, N^*). I invoke Proposition 1 to further derive that

$$\frac{dN}{d\phi}-\alpha N=\alpha N\left(\frac{R(1-2Ns)+Ns}{R(1-2Ns)+1}-1\right)$$

$$=-\alpha N\left(\frac{1-Ns}{R(1-2Ns)+1}\right)$$

and the result follows immediately.

Proof of Proposition 5

Proposition 2 and an earlier result imply that at (ϕ^*, N^*),

$$\frac{d(Ns)}{d\phi} = N\frac{\partial s}{\partial \phi}\left[\frac{1 - Ns}{R(1 - 2Ns) + 1}\right].$$

Focusing on the term in brackets, first note that it reduces to 1/2 when $R = 1$, which gives the equality in the proposition. For the case where $R > 1$, a bit of algebraic manipulation shows that

$$\frac{1 - Ns}{R(1 - 2Ns) + 1} = \frac{1 - Ns}{(1 - RNs) + R(1 - Ns)} < \frac{1 - Ns}{R(1 - Ns)} = \frac{1}{R}.$$

Combining this result with the fact that the two derivatives of interest are negative yields the bounds in the proposition.

Proof of Proposition 6

The first statement in the proposition follows directly from the formula. The firms' profit-maximizing first order conditions in the initial equilibrium (without a minimum markup) require that $\frac{1}{\alpha\phi} = 1 - Ns$ and $\frac{k}{\phi Ms} = 1 - Ns$. Taken together, these also imply that $\alpha k = Ms$. So, evaluating the relevant derivative at (ϕ^*, N^*),

$$\frac{d\Pi_r}{d\phi} = M\frac{d(Ns)}{d\phi}\phi + MNs - \frac{dN}{d\phi}k$$

$$= -MNs(1 - RNs)\frac{1 - Ns}{R(1 - 2Ns) + 1}\alpha\phi + MNs - N\frac{R(1 - 2Ns) + Ns}{R(1 - 2Ns) + 1}\alpha k$$

$$
\begin{aligned}
&= -MNs(1-RNs)\frac{1}{R(1-2Ns)+1} + MNs - N\frac{R(1-2Ns)+Ns}{R(1-2Ns)+1}Ms \\
&= MNs\left[1 - \frac{1-RNs}{R(1-2Ns)+1} - \frac{R(1-2Ns)+Ns}{R(1-2Ns)+1}\right] \\
&= MNs\left[\frac{RNs-Ns}{R(1-2Ns)+1}\right] \\
&= M(Ns)^2\left[\frac{R-1}{R(1-2Ns)+1}\right].
\end{aligned}
$$

APPENDIX B
Cross-Category Incentives and Unobserved Marginal Cost

Here I show a simple example to illustrate how cross-category incentives will affect the estimated marginal cost when first order conditions for only a subset of the products are used in estimation. Consider a firm selling two products, indexed $i \in \{1, 2\}$, and facing constant marginal costs. The firm's profit maximization problem is

$$\max_{p_1, p_2} \quad (p_1 - c_1)q_1(p_1, p_2) + (p_2 - c_2)q_2(p_1, p_2).$$

The first order conditions for this problem are

$$q_1 + (p_1 - c_1)\frac{\partial q_1}{\partial p_1} + (p_2 - c_2)\frac{\partial q_2}{\partial p_1} = 0$$

$$(p_1 - c_1)\frac{\partial q_1}{\partial p_2} + q_2 + (p_2 - c_2)\frac{\partial q_2}{\partial p_2} = 0.$$

If prices and quantities are observed, then the full demand system can be estimated, yielding the relevant partial derivatives. The researcher can then use the first order conditions to recover the unobserved marginal costs.

However, suppose that price and quantity are only observed for good 1. Suppose furthermore that the researcher is able to use these to obtain a close estimate of $\frac{\partial q_1}{\partial p_1}$. If the researcher tries to recover the marginal cost for good 1 by assuming that there is no cross-category effect, then the researcher assumes (incorrectly) that the firm's objective function is

$$\max_{p_1, p_2} \quad (p_1 - \tilde{c}_1) q_1(p_1)$$

The unknown term, $\tilde{c}_1$, is found using the first order condition for this problem:

$$\tilde{c}_1 = \frac{q_1}{\partial q_1 / \partial p_1} + p_1.$$

However, inspection of the first order condition for the price of good 1 in the firm's true profit maximization problem (over both prices) yields the following:

$$\frac{q_1}{\partial q_1 / \partial p_1} + p_1 = c_1 - (p_2 - c_2) \frac{\partial q_2 / \partial p_1}{\partial q_1 / \partial p_1}$$

The final term in the above equation is the cross-category incentive. So, the researcher's estimate of the marginal cost of good 1 will be

$$\tilde{c}_1 = c_1 - (p_2 - c_2) \frac{\partial q_2 / \partial p_1}{\partial q_1 / \partial p_1}.$$

Downward-sloping demand requires that $\frac{\partial q_1}{\partial p_1} < 0$. So, if the markup on good 2 is positive and good 1 is a gross complement for good 2 (i.e., $\frac{\partial q_2}{\partial p_1} < 0$), then $\tilde{c}_1 < c_1$; the cross-category incentive decreases the estimate of marginal cost for good 1. This is intuitive: decreasing the price for good 1 results in increased sales of both good 1 and good 2, so the cross-category effect increases the marginal revenue from selling an additional unit of good 1. This is equivalent to decreasing the marginal cost of good 1.

Two important questions arise from this analysis as it pertains to the marginal cost estimates in the body of the text. First, under what conditions can $\frac{\partial q_1}{\partial p_1}$ be closely estimated if prices and quantities are only observed for good 1? Second, when is the cross-category incentive constant? The first question is straightforward to deal with: a sufficient condition for accurate estimation $q_1(p_1, p_2)$ using traditional techniques is that $q_1(p_1, p_2) = q_1(p_1)$, so that changes in p_2 do not affect q_1. In other words, the cross-category incentive only goes one way. This may be a reasonable assumption if good 1 is a strong driver of store traffic but good 2 is not.

While this condition may be reasonable, it is not sufficient to ensure that the cross-category incentive will be constant. First, I clarify the meaning of a "constant" cross-category incentive. The retailer's first order condition for the price of good 2 should always hold, so that $p_2 = p_2^*(p_1)$. With this constraint imposed, the cross-category incentive can be thought of solely as a function of p_1 and is constant if

$$\frac{d}{dp_1}\left[(p_2 - c_2)\frac{\partial q_2/\partial p_1}{\partial q_1/\partial p_1}\right] = 0$$

for any feasible p_1. If the cross-category incentive is only one-way, then the retailer's first order condition on the price of good 2 reduces to

$$p_2 - c_2 = \frac{-q_2}{\partial q_2/\partial p_2}$$

and a constant one-way cross-category incentive therefore requires that

$$\frac{d}{dp_1}\left[\frac{-q_2}{\partial q_2/\partial p_2} \cdot \frac{\partial q_2/\partial p_1}{\partial q_1/\partial p_1}\right] = 0$$

for any feasible $(p_1, p_2^*(p_1))$. This appears to be a fairly strict condition, but it is indeed possible to satisfy with a reasonable demand system. Consider the model in which $q_1(p_1) = f(p_1)$ and $q_2(p_1, p_2) = g(p_1)h(p_2)$, where $f', g', h' < 0$, so that there is one-way complementarity. The first order condition on the price of good 2 is

$$p_2 - c_2 = \frac{-h(p_2)}{h'(p_2)}$$

which implies that $p_2^*(p_1) = \bar{p}_2$. So, at $(p_1, \bar{p}_2)$

$$\frac{-q_2}{\partial q_2/\partial p_2} \cdot \frac{\partial q_2/\partial p_1}{\partial q_1/\partial p_1} = \frac{-h(\bar{p}_2)^2}{h'(\bar{p}_2)} \cdot \frac{g'(p_1)}{f'(p_1)}$$

and a sufficient condition for this to be constant is that $g(\cdot)$ is an affine transformation of $f(\cdot)$.[27]

[27] It is unclear whether there is structural model of consumer behavior (i.e., utility function or search behavior) that would generate such a demand system. Exploration of this question is left to future work.

APPENDIX C
Nash-in-Nash Bargaining

In this appendix, I describe an alternative model of wholesale price setting. In this model, the wholesale prices are outcomes of bilateral negotiations. The model is one of Nash bargaining (Nash, 1950) where one side of the market has a single agent. The negotiation is over linear wholesale prices and does not include fixed fees. Once wholesale prices are determined, the retailer sets retail prices to maximize its own profits. The effect of a change in wholesale prices on the retail prices is internalized in the bilateral negotiation. Much of the notation in this appendix is similar to that of Crawford and Yurukoglu (2012), who use a similar model of bilateral negotiations, and is somewhat different from that used in Chapter 1.4.3.

Let $\Psi = \{w_f\}$ be the set of wholesale prices, each for a manufacturer's product, f. The vector of manufacturer (brewer) b's wholesale prices is denoted as w_B, where B is the set of products sold by manufacturer b. The bilateral negotiation between manufacturer b and the retailer, r, results in a vector of wholesale prices, w_B, that maximizes the Nash product:

$$NP_{b,r}(w_B, \Psi_{-B}) = [\Pi_b(w_B, \Psi_{-B}) - \Pi_b(\Psi_{-B})]^\eta \, [\Pi_r(w_B, \Psi_{-B}) - \Pi_r(\Psi_{-B})]^{1-\eta}$$

where the profits for both the retailer and manufacturer are summed across all markets. The bargaining weight, η, is common across negotiations. The baseline specification in Chapter 1.4.3 is obtained by setting $\eta = 1$. The disagreement payoffs, $\Pi_b(\Psi_{-B})$ and $\Pi_r(\Psi_{-B})$, assume that manufacturer b will not sell any of its products to the retailer if there is disagreement on the wholesale prices for any subset.[28] This assumption may be reasonable in the context of macrobrews, as stores are rarely seen carrying one but not all of a brewer's products. That is, stores that carry Budweiser also carry Bud Light, etc. One potential alternative assumption would be that disagreement on product f only causes the retailer to drop that product. I also follow Crawford and Yurukoglu (2012) in assuming that disagreement between manufacturer b and the retailer does not change the other manufacturers' wholesale prices, Ψ_{-B}. This assumption is in the spirit of the Nash equilibrium, although more details about the industry would be needed to comment further on its validity. Additionally, this assumption has a significant computational advantages, as there is no need to compute separate vectors of wholesale prices under disagreement.

The equilibrium of this model is "Nash-in-Nash," so that each w_B is chosen to maximize $NP_{b,r}(w_B, \Psi_{-B})$. An implicit assumption in this equilibrium is that the retailer sends separate representatives to negotiate with each manufacturer and that these representatives do not coordinate with each other during negotiations. I also assume that the manufacturers do not use a common representative, which is likely the case in this industry.

An issue arises from the fact that the upstream marginal costs are unobserved and enter the manufacturers' profit functions. If wholesale prices are high on average, this could be due either to high upstream marginal costs or to a high bargaining parameter. So, it is difficult to separately identify the bargaining parameter from upstream marginal costs using the data at hand.[29] Instead of separately identifying them, I propose fixing the value of the bargaining parameter and then using the first order conditions from the Nash-in-Nash equilibrium to

[28] Because there is only one retailer, $\Pi_b(\Psi_{-B}) = 0$.

[29] If each manufacturer has a separate bargaining parameter, then it is impossible to separately identify these from manufacturer marginal costs. See Crawford and Yurukoglu (2012) for a discussion of this issue.

back out the upstream marginal costs. As a robustness check, the counterfactual simulations can then be repeated using the Nash bargaining structure and the backed-out marginal costs. This procedure can then be repeated for several different values of the bargaining parameter.

Although I leave the actual implementation of this procedure to future work, it is useful to comment on the likely effects of decreasing the bargaining parameter from $\eta = 1$. First, decreasing the bargaining parameter essentially decreases the market power of the upstream firms. So, it is likely that lower values of η will result in smaller wholesale price increases in the counterfactuals. I expect for this change to be "roughly" continuous in η. Second, a loose check of the validity of each value of the bargaining parameter is to compare the median (or average) backed-out upstream marginal cost as a percentage of the retail price to the 58% value from Tremblay and Tremblay (2005). As noted in Chapter 1.4.3, the manufacturer marginal costs implied by $\eta = 1$ represent a median of 60.59% of the retail price. As the value of the bargaining parameter falls, the median backed-out upstream marginal cost will rise and therefore diverge further from 58% of the retail price. This provides some evidence that the baseline specification ($\eta = 1$) may be the most reasonable choice.

Bibliography

Aguirre, Inakai, Simon Cowan, and John Vickers, "Monopoly Price Discrimination and Demand Curvature," *American Economic Review*, September 2010, *100*, 1601–1615.

Aksoy-Pierson, Margaret, Gad Allon, and Awi Federgruen, "Price Competition Under Mixed Multinomial Logit Demand Functions," *Management Science*, August 2013, *59* (8), 1817–1835.

Anderson, Rod W. and Ronald N. Johnson, "Antitrust and Sales-Below-Cost Laws: The Case of Retail Gasoline," *Review of Industrial Organization*, May 1999, *14* (3), 189–204.

Asker, John, "Diagnosing Foreclosure due to Exclusive Dealing," Working Paper 2015.

Berry, Steven, "Estimating Discrete-Choice Models of Product Differentiation," *The RAND Journal of Economics*, 1994, *25* (2), 242–262.

Cardell, N. Scott, "Variance Components Structures for the Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity," *Econometric Theory*, April 1997, *13* (2), 185–213.

Chintagunta, Pradeep, "Investigating Category Pricing Behavior at a Retail Chain," *Journal of Marketing Research*, May 2002, *39* (2), 141–154.

_ , Jean-Pierre Dubé, and Vishal Singh, "Balancing Profitability and Consumer Welfare in a Supermarket Chain," *Quantitative Marketing and Economics*, 2003, *1*, 111–147.

Cowan, Simon, "Third-Degree Price Discrimination and Consumer Surplus," *The Journal of Industrial Economics*, 2012, *60* (2), 333–345.

Crawford, Gregory S. and Ali Yurukoglu, "The Welfare Effects of Bundling in Multichannel Television Markets," *American Economic Review*, April 2012, *102* (2), 643–685.

Goldberg, Pinelopi and Rebecca Hellerstein, "A Structural Approach to Identifying the Sources of Local-Currency Price Stability," Working Paper 2012.

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

Hamilton, Stephen F., "Excise Taxes with Multiproduct Transactions," *The American Economic Review*, March 2009, *99* (1), 458–471.

Hellerstein, Rebecca, "Empirical Essays on Vertical Contracts, Exchange Rates, and Monetary Policy." PhD dissertation, University of California, Berkeley 2004.

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

Hendel, Igal and Aviv Nevo, "Sales and Consumer Inventory," *The RAND Journal of Economics*, 2006, *37* (3), 531–561.

Hoch, Stephen, Byung-Do Kim, Alan Montgomery, and Peter Rossi, "Determinants of Store-Level Price Elasticity," *Journal of Marketing Research*, February 1995, *32* (1), 17–29.

Kahneman, Daniel, Jack Knetsch, and Richard Thaler, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, September 1986, *76* (4), 728–741.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior," in P. Zarembka, ed., *Frontiers in Econometrics*, New York: Academic Press, 1974, pp. 105–142.

_ , "Modeling the Choice of Residential Location," in A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, eds., *Spatial Interaction Theory and Planning Models*, Amsterdam: North-Holland, 1978, pp. 75–96.

McMillan, Rob, "Different Flavor, Same Price: The Puzzle of Uniform Pricing for Differentiated Products," Working Paper 2007.

Meza, Sergio and K. Sudhir, "Pass-Through Timing," *Quantitative Marketing and Economics*, 2006, *4*, 351–382.

Miller, Nathan and Matthew Weinberg, "The Market Power Effects of a Merger: Evidence from the U.S. Brewing Industry," Working Paper 2016.

Montgomery, Alan, "Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data," *Marketing Science*, 1997, *16* (4), 315–337.

Nash, John F., "The Bargaining Problem," *Econometrica*, 1950, *18* (2), 155–162.

Pigou, Arthur C., *The Economics of Welfare*, London: Macmillan, 1920.

Rey, Patrick and Jean Tirole, "The Logic of Vertical Restraints," *American Economic Review*, 1986, *76*, 921–939.

Salop, Steven C., "Monopolistic Competition with Outside Goods," *The Bell Journal of Economics*, Spring 1979, *10* (1), 141–156.

Schmalensee, Richard, "Output and Welfare Implications of Monopolistic Third-Degree Price Discrimination," *American Economic Review*, 1981, *71* (1), 242–247.

Skidmore, Mark, James Peltier, and James Alm, "Do State Motor Fuel Sales-Below-Cost Laws Lower Prices?," *Journal of Urban Economics*, January 2005, *57* (1), 189–211.

Smith, Howard and Øyvind Thomassen, "Multi-Product Competiton, Shopping Costs, and Market Power: A Model of Supermarket Pricing," Working Paper 2015.

Sullum, Jacob, "Fearing 'Monsters' Like Total Wine, Liquor Merchants Want Connecticut To Keep Prices High," 2016.

Tremblay, Victor J. and Carol Horton Tremblay, *The U.S. Brewing Industry: Data and Economic Analysis*, Cambridge, MA: The MIT Press, 2005.

Varian, Hal R., "Price Discrimination and Social Welfare," *American Economic Review*, 1985, *75* (4), 870–875.

Villas-Boas, Sofia Berto, "Vertical Relationships between Manufacturers and Retailers: Inference with Limited Data," *The Review of Economic Studies*, April 2007, *74* (2), 625–652.

_ , "An Empirical Investigation of the Welfare Effects of Banning Wholesale Price Discrimination," *The RAND Journal of Economics*, 2009, *40* (1), 20–46.

Viscusi, W. Kip, John M. Vernon, and Joseph E. Harrington, *Economics of Regulation and Antitrust*, Cambridge, MA: MIT Press, 2005.