2019

# ESSAYS ON PRICE DISCRIMINATION AND DEMAND LEARNING

Benjamin E. Wallace
*University of Kentucky*, bewa225@g.uky.edu
Author ORCID Identifier:
https://orcid.org/0000-0002-3962-7172
Digital Object Identifier: https://doi.org/10.13023/etd.2019.304

Benjamin E. Wallace, Student

Dr. Adib Bagh, Major Professor

Dr. William Ederington, Director of Graduate Studies

# ESSAYS ON PRICE DISCRIMINATION AND DEMAND LEARNING

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Business and Economics at the University of Kentucky

By

Benjamin E. Wallace

Lexington, Kentucky

Co-Directors: Dr. Adib Bagh, Professor of Economics
and Dr. Anthony Creane, Professor of Economics

Lexington, Kentucky

# ABSTRACT OF DISSERTATION

## ESSAYS ON PRICE DISCRIMINATION AND DEMAND LEARNING

This dissertation consists of three essays examining how and why firms set prices in markets. In particular, this dissertation shows how firms may utilize nonlinear pricing to price discriminate, how firms may experiment with the prices they set to learn about the demand function in the market they serve in later periods and the effects of these pricing strategies on consumer welfare.

In Essay 1, I show how firms in the milk market use nonlinear price schedules – quantity discounts – to price discriminate and increase profits. I find that firms have a greater ability to price discriminate on their own "private label" products rather than regional branded that they sell alongside their own. Though some consumers benefit from a lower price as a result of the price discrimination, total consumer surplus is lower than if the store had to offer a fixed price per unit. Additionally, I compare my structural demand estimates, which using the Nielsen household panel data include consumer demographic information and actual household choices, to the standard approach in the literature on price discrimination that uses only market level data. By doing so I find that ignoring demographic information and actual consumer choices leads to biased parameter estimates. In the case of the milk market, the biased parameter estimates due to ignoring household demographic information and actual consumer choices lead to underestimating welfare harm to consumers on average.

After finding that price discrimination harms consumers overall in this market, I quantify which consumer demographic are better off and which are worse off. I find that households with children and low income households with children are the only households to benefit from the price discriminatory practices of firms in this market. Since these groups are particularly vulnerable, I suggest that policymakers take no action to correct this market, as any action will directly hurt these consumer groups.

In Essay 2, I study how firms learn about the demand in a new market by exploiting a significant change in Washington's state's liquor laws. In 2012, the state of Washington switched from a price-controlled state-store system of selling liquor to one in which private sellers could sell liquor with minimal restrictions on price and range of products. As a result, a heterogeneous group of firms entered the liquor market across the state with little knowledge of the regional demand for alcohol in the state of Washington across heterogeneous localities. Using the Nielsen retail scanner data

I am able to observe the variation in pricing and offerings seasonally and over time to see if there is convergence in offerings and prices, and how quickly that convergence occurs across different localities depending on local demographics and competition. I also investigate the extent to which the variation is "experimentation" by the firms, i.e., the firms purposely experimenting to learn more about demand and the extent that local demographics and competition can affect the experimentation and whether there are spill-overs from local competition (i.e. do firms learn from each other and does this effect how much they experiment and how quickly they learn).

My main findings are that over time, firms within this market have learned better how to price discriminate over the holiday season; firms experiment more with prices for the pint sized products than the larger sizes; and that menu of options that firms have offered has been expanding but at a slower rate, suggesting that they are approaching a long-run steady state for the optimal menu of options.

KEYWORDS: Price Discrimination, Nonlinear Pricing, Demand Learning, Industrial Organization, Structural Modeling

Author's signature: Benjamin E. Wallace

Date: July 12, 2019

ESSAYS ON PRICE DISCRIMINATION AND DEMAND LEARNING

By

Benjamin E. Wallace

Co-Director of Dissertation: Dr. Adib Bagh

Co-Director of Dissertation: Dr. Anthony Creane

Director of Graduate Studies: Dr. William Ederington

Date: July 12, 2019

For my mother, Hannah, and my wife, Megan

## ACKNOWLEDGMENTS

I am extremely grateful to my Dissertation Co-Chairs Dr. Adib Bagh and Dr. Anthony Creane, and to Dr. Frank Scott, the pseudo third Co-Chair of my dissertation committee. As far as I am concerned, all three of these men have gone above and beyond what was required of them and I am very appreciative of the countless hours that each has invested in me during my time here at the University of Kentucky. Without their comments and guidance, this dissertation would not exist. I am particularly grateful for the patience and advice I have received from each of these individuals, and I truly feel that they have always represented my best interest with the help they have provided, and for that I am thankful. I cannot imagine having a better set of advisors.

Next I would like to thank the other members of my dissertation committee: Dr. Lala Ma and Dr. Yuqing Zheng. They both have been very active and helpful in providing feedback and help in my work in this dissertation, and I have benefited greatly from their expertise. Finally, I thank my wife, Megan, and my mother and father, Hannah and Eric, for their support and encouragement during this time.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

# Chapter 1 Nonlinear Pricing: Evidence of Price Discrimination in the Fluid Milk Market.[1]

## 1.1 Introduction

Nonlinear pricing schedules are a pricing strategy in which firms vary the marginal price per unit of their goods over several sizing options. Generally, firms use nonlinear pricing to provide quantity discounts to high demand consumers, but it is also possible that firms may utilize pricing menus that provide a quantity premia.[2] Dolan (1987) states that firms use nonlinear pricing to either price discriminate, incentivize consumers to purchase the options which have a lower marginal cost, to remain competitive, or some combination of the previous three. By varying marginal price per unit over size, firms are able to charge consumers with lower preferences for the good in question higher marginal prices per unit than those who prefer the good more and extract additional surplus from the market. Firms may also use nonlinear pricing schedules to incentivize consumers to choose the larger option, which is cheaper for the firm to produce per unit. Additionally, if a firm's competitors are offering cheaper prices per unit for larger sizing options, then they may have to follow suit or lose business. This paper will focus only on firms using nonlinear pricing as a way to price discrimination.

With price discrimination as the reason for nonlinear pricing in mind, we may see firms manipulate either the price[3] or the size of their products as to increase the marginal price/marginal quantity per unit ratio of a small version above the efficient ratio. Doing so would allow firms to extract additional surplus from the lower demand consumers who purchase the smaller sizes. The theoretical literature on price discrimination suggests that this is not possible for the largest sizes of a product offered, as high demand-type consumers have relatively more elastic demand functions than low demand-type consumers.[4] This finding from the theoretical literature suggests that the largest sizes of a product must be priced and sized at the efficient marginal price/marginal quantity per unit ratio. However, since low demand-type consumers have a relatively more inelastic demand function than the high demand-type consumers, there is a concern that the smallest sizes, which are constructed by the firm to capture the low demand-type consumer market, could be inefficient in terms of their marginal price/marginal cost per unit ratios. Since the practice

[1] *Calculated (or Derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.*

[2] Since firms within the fluid milk industry exclusively utilize nonlinear pricing schedules that provide quantity discounts, this paper will focus entirely on nonlinear pricing in terms of quantity discounts.

[3] From here on forward, when I refer to price, I am referring to price per unit.

[4] Assuming the Spence-Mirrelees condition holds.

of second-degree price discrimination may be to the detriment of consumers from a welfare perspective, this concern leads me to analyze the welfare implications of these nonlinear pricing schedules. Which is a question that the previous empirical literature has found differing answers for.

The main empirical issue for estimating whether or not the theory holds in the case of nonlinear pricing schedules in the past has been due to data availability. Many other papers have estimated demand models using only product characteristics, prices and aggregate sales data. One of the strengths of this paper is that with the Nielsen Consumer Panel, I am able to observe household shopping decisions, as well as the demographic characteristics of the households, for every shopping trip they make for every year they participate in the panel. The Nielsen data allow me to use not only product characteristics to estimate the preferences of the households, but characteristics of these households to obtain a better fit to my discrete choice model by allowing for additional heterogeneity in tastes.

The market in question for this paper is fluid milk.[5] I choose to focus on fluid milk market because it has several convenient properties for the estimation of consumer preferences on the consumer side of the model, as well as marginal costs on the firm side. First, the USDA regulates and publishes raw fluid grade milk prices in the United States which provides a reasonable starting point for estimating the marginal cost per ounce of fluid milk. Estimating costs has been difficult for several other empirical studies and these regulations allow me to relax some of the assumptions that other papers have needed when modeling the firm's side of the model. Second, fluid milk sizing options in their current state are exogenous in that they follow the Imperial measurement system in the United States. This market norm allows me to focus this paper only on pricing and not have to worry about firms changing their menus through size. This is a potential weakness of the previous literature, which chooses products for analysis where no such market norms exist. The main reason why the endogeneity of size may be a concern is that firms likely have the ability to choose both prices and quantities in the majority of other markets, a fact that has been mostly ignored by the majority of the previous literature and may lead to biased parameter estimates. Due the exogeneity of sizes in the fluid milk market, this potential source of bias is no longer a concern. Another strength of fluid milk for this type of analysis is that the outside good in question, or substitute good, can be well defined. I will assume that the outside good for fluid milk is any other consumable liquid product.[6]

There are a couple necessary conditions required for firms to price discriminate. These necessary conditions include: firms must face a downward sloping demand curve, firms must be able to prevent resale of their product, and consumers must have heterogeneous tastes for the product. One may question that these necessary conditions are met in this market, since fluid milk appears to be a fairly homogeneous

[5]Which I define to be fluid milk only from cows. This definition excludes non-fluid milks, i.e. evaporated or condensed milk, milks that come from non-cow animals and milks from plants. This is done for simplicity.

[6]The outside good includes all substitutes for fluid milk such as: pop, juice, water, etc.

good. However, several simple observations of this market make it clear that these conditions are met.

There are many different types of fluid milk that are offered by many different firms. Some examples of the ways that firms vary their fluid milk's product characteristics are through a variety of sizes,[7] butterfat contents,[8] flavors[9] and whether or not the product has the USDA organic seal. I even observe some producers of fluid milk produce several different brands targeted at different segments of the market. Since all of these different product characteristics in all types of combinations exist for fluid milk on stores shelves, consumers must have heterogeneous tastes for fluid milk, otherwise there would be no reason to have this much variety within the market.

Additionally, it seems reasonable to assume that resale is not an issue in this market for a couple of reasons. Fluid milk is a relatively small portion of the average consumer's budget in that generally one to two gallons per week are plenty for households with even strongest of preferences for milk and the average price of a gallon of fluid milk is a couple of dollars. In addition, if resale were a concern in this market, we would observe consumers purchasing the largest size of one type of fluid milk at the efficient marginal price/marginal quantity per ounce bundle and then pour the portion that they do not plan to consume out into another container and attempt to resell it on some secondary market. I am not aware of such a market. Another reason this is a reasonable assumption is that resale would require a large amount of coordination between households, which would seem to have a much higher marginal cost in terms of effort than the marginal gain from saving a couple of pennies. Additionally, the perishability of milk further prevents its resale.

One may think that firms within the fluid milk market may not face a downward sloping demand curve and have no market power thus would be unable to price discriminate due to fluid milk being a fairly homogeneous good. For this, I provide a couple of observations as anecdotal evidence of price discrimination within the market. On one shopping trip at a Kroger store in October 2016, I observed that the price for private label[10] gallon containers of 2% fluid milk was listed at \$1.88 whereas the price of private label 2% quart containers was listed at \$1.89. The packaging of both containers were very similar. This observation is hard to explain through a story of channel efficiency alone, as it is not likely that the absolute cost of producing a gallon container of fluid milk is cheaper than producing a quart of milk. With this pricing menu, the large grocery chain is actually paying consumers a penny to take away an additional three quarts of fluid milk.

One of the largest sources of market power that firms likely have in this market are search costs. Since fluid milk is such a small portion of the budget constraint, household on their weekly shopping trip are highly unlikely to make a trip to a different store to purchase fluid milk, even if they know that they can purchase the product for \$0.50 less per gallon at a different location. Firms must have a decent idea as to what these search costs are for consumers and set their prices accordingly.

[7] One and a Half Gallon, Gallon, Half Gallon, Quart, Pint, Cup, and a variety of multipacks.

[8] Skim, .5%, 1%, 1.5%, 2% and Whole fluid milk.

[9] Regular, chocolate, strawberry, etc.

[10] Private label meaning store brand product.

Even more evidence of firms within this market having market power was the difference in price between the non-organic regional brand of milk and the private label brand. A gallon of regional brand 2% fluid milk at this grocery store was listed at $4.99 a gallon, more than twice the price of the private label of $1.88. Unless this regional brand producer uses a wildly different production process, this more than two and a half times price differential is hard to explain through differential costs alone. There are two different explanations for this price differential. Regional firms could be exerting market power to charge a higher price for their "different, higher quality product." One concern with this explanation is that, overall, the retailers have the final say in what price is seen by consumers, though many of these regional firms are known to provide incentives to retailers to set the price they prefer. An alternative explanation for this observation, based upon the fact that retailers have final say in prices, could be some profit maximizing behavior by retailers attempting to induce consumers to purchase their private label brand by having the shelf price of the regional brand good be significantly higher.

There are two main contributions of this paper. First, I show that firms use nonlinear pricing schedules to price discriminate among heterogeneous consumers with consumer choice micro data rather than aggregate sales and simulated consumer behavior. I do so by estimating a structural model of household preferences for fluid milk, I then use these parameter estimates, which allow for additional heterogeneity in tastes among households through the micro data, to compute the welfare implications of this form of price discrimination from the consumer's perspective. To estimate the welfare implications of price discrimination within this marketplace, I compare the total welfare in a hypothetical world where price discrimination via nonlinear pricing schedules is not allowed to the total welfare of the market place as it exists currently. This has been done in the past[11] but as stated previously, has yet to have been agreed upon within the empirical literature. Since this paper uses actual consumer choices during shopping trips, I find my result more compelling as the previous papers use only aggregate sales data and simulated consumer decisions.

Second, I compare my estimates using household demographics and actual decisions to the status quo approach using aggregate data. Doing so, I find that ignoring household demographics and choice data leads to underestimating the welfare harm to consumers that is due to this pricing strategy.

This chapter is organized as follows. Section 1.2 presents a layout of the composition of the market for fluid milk. Section 1.3 provides a review of the existing literature for both the theoretical and empirical work on price discrimination as it relates to this paper. Section 1.4 discusses the datasets used for this paper as well as the processes and algorithms used for their cleaning. Section 1.5 provides preliminary evidence of nonlinear pricing being used to price discriminate in this market. Section 1.6 introduces and estimates a structural model for consumer demand as well as a strategy for estimating the marginal cost of fluid milk firms face serving this market. In section 1.7, I provide a counterfactual analysis to determine whether or not the nonlinear price schedules utilized by firms in this market is to the determent or ben-

[11] See Leslie 2004 and others.

efit to the households they serve, both in aggregate and by household demographic groupings. Section 1.8 concludes this chapter.

## 1.2 Market Description: Fluid Milk in USA

With many steps of the production process, the market structure, and regulation of bulk milk prices, fluid milk retail pricing can be somewhat complex. This section of the paper will discuss all of these in three subsections. The first subsection will discuss the production process and market structure for fluid milk, whereas the second will discuss the regulations that are in place for the fluid milk market, and lastly the third will discuss the implications of the market structure and regulations.

### 1.2.1 Production Process and Market Structure

The first part of the production process starts at a farm where farmers raise cows that produce bulk milk. There are also seasonal supply fluctuations for production, as is true with almost all agricultural products. In the spring and early summer, cows produce more milk, whereas production of milk by cows decreases in the fall and winter. Once the cows produce the bulk milk, it is sent to a processor that separates, produces and packages the various dairy products that come from the bulk milk. Once the production process is complete, the finished fluid milk product is distributed to a retailer where, according to Manchester and Blayney (USDA 2001) [MB01], the retailer then sets the price of the final good and consumers make their consumption choices.

Dairy farmers may also have an incentive to form cooperatives to help gain market power when dealing with processors. These cooperatives control the supply of milk sent to processors and help eliminate some of the risk that a dairy farmer may face due to daily, season and other types of market fluctuations.

The production process for fluid milk has led to an interesting market outcome in this industry. Many of the large grocery chains are fully integrated and have their own processing plants. Unfortunately, this is not the case for all large grocery chains within the industry. If it were, a possible way to identify which UPCs come from a vertically integrated production process would be via private label packaging vs regional brand. However, Nielsen masks the identity of each store's parent company by just assigning a common identifier among stores who are owned by the same parent company due to confidentiality reasons. As such, information on which large chains are vertically integrated is not useful, as it is not possible to line up the parent companies with the stores within the data.

Another difficulty with identifying which goods are produced via a vertically integrated production process is that some firms are vertically integrated in some parts of the country but not in others. For example, consider the following quote from Dean's, the largest regional brand fluid milk firm found in my data, website:

> The Company (Dean's) is one of the nation's largest processors and direct-to-store distributors of fluid milk marketed under more than 50 local and

regional dairy brands and private labels.

This quote illustrates another issue with identifying which products come from a vertically integrated production process or not, as some of the "private label" brand goods, which are usually what is thought of when discussing a product that is produced within a vertically integrated firm are not necessarily from a vertically integrated production process.

Additionally, which processing plant that grocery stores are receiving their fluid milk from is not readily available due to confidentiality as well, however the several potential processing plants that supply the stores are known.

All smaller convenience stores deal with independent milk processors. Many of these stores only carry one brand of fluid milk in only a few sizes. Due to this, as well as the generally higher prices, it is likely more appropriate to treat these convenience stores as a completely different market.

According to MB01, in the past many large grocery chains had several[12] regional brands on their shelves that were fully serviced by a representative from each brand respectively. In this regard, fully serviced means that each of the brands would have a representative bring the fluid milk to the store and have that representative stock, rotate and care for their brands space. In recent years,[13] there has been a shift from grocery chains to only have one or no regional brands in their floor plan. In addition to this, these regional brands no longer fully service these stores, instead they ship their product to the store and it is handled by the retailer. Once the product hits the loading dock of a retailer it becomes their property, although many of these regional brands will refund the retailer for damaged and outdated goods.

### 1.2.2 Regulations

Throughout this production process, there are many regulations set in place by the government. These pricing regulations are all placed on the processors and aim to benefit domestic dairy farmers. According to MB01, these regulations include: Federal milk price support, Federal milk marketing orders, import restrictions, export subsidies, domestic and international food aid programs, state level milk marketing programs and multi-state milk pricing organizations. Each of these regulatory practices has its own goal, some of which will be discussed below.

Federal milk price support was first put into place in 1949. These price supports provide a price floor for the price of bulk milk, though the price floor often does not bind. Federal milk marketing orders (1937) also provide a similar function in that they set pricing minimums for bulk milk, regulate the quality of fluid milk and limit the monopsony power of dairy processors.

The federal government also attempts to protect domestic dairy farmers through import restrictions and export subsidies. These two processes in conjunction give domestic dairy farmers the upper hand against foreign competition.

[12] They say two to six.

[13] MB01 states the 1990s onward.

In addition to Federal pricing support, many states have their own policies, which vary from state to state.

Overall, the goal of these regulations are to protect domestic dairy farmers from both international competition and dairy processors. These regulatory policies are stated to aim to help domestic dairy farmers by ensuring a minimum price in the market for bulk milk as well as ensure the quality of the fluid milk that comes from the processor. It also seems that the regulations that exist aim to lower risk for dairy farmers, though this is not explicitly stated by the USDA.

These data on bulk milk prices are available from the USDA's National Agricultural Statistics Service. The butterfat, bulk and skim milk prices are reported monthly at the state level.

### 1.2.3 Implications

According to MB01, even with the complex market structure and regulations in place in the fluid milk industry, ultimately the retailer is allowed to set any price they see fit in regards to the price that consumers face.[14] Even though each may set whichever price they see fit in terms of the prices that consumers face, many of the regional brands provides have a strong incentive to nudge retailers to set prices that the regional brands would prefer.[15] There is no federal regulation that requires a retailer to set a certain price for each good,[16] thus retail prices are determined by some combination of consumer demand, competition between retailers, and wholesale price.[17]

## 1.3 Literature Review

In this section, I summarize the current literature that exists on nonlinear pricing and price discrimination, starting with the theoretical literature then the empirical literature. The theoretical literature on price discrimination is quite extensive, as is the case with the majority of topics within the industrial organization field, whereas the current empirical literature is not nearly as extensive, mainly due to data limitations.

The story for price discrimination starts with Dupuit (1849). Dupuit laid out the necessary conditions for price discrimination by a profit maximizing monopolist as well as provided an example. In his paper, Dupuit discusses the several classes of tickets offered by trains. These different classes were used to charge higher demand consumers higher prices as well as used to induce the marginal consumer to purchase

[14] Though the range of prices that stores may set depends on which state they are located in. It is my understanding that in states such as California, grocery stores are allowed very little freedom to set their own prices due to the state level price regulations implemented by the California legislature, but in other states, no such pricing regulations exist at the state level at all. I am currently working on finding this information out for the states in my sample, but this information has been difficult to find.

[15] One such way of doing so, which I have heard is the case in the soda industry, is to provide large cost reductions to stores if they set the price at $X during a specific period of time.

[16] Though there may be state level regulations that impact this decision.

[17] Which is a function of the Federal regulation on butterfat, skim and bulk milk prices.

a higher class ticket then they might otherwise purchase. Pigou (1920) and Robinson (1933) both more formally define the degrees to which a monopolist can price discriminate as well as named them.

The first couple of modern papers to discuss second-degree price discrimination were Mussa and Rosen (1978) and Maskin and Riley (1984). These papers consider a monopolist's design of a price and sizing menu when consumers' preferences towards the product are unknown to the firm, but have some distribution that is known to the firm. The equilibrium that is found by both papers is that the highest demand consumer is offered a product that is efficiently sized and all other consumers self select into a smaller option that is tailored to their preferences though inefficiently sized.

Many papers have used this result and attempted to explain why firms may use nonlinear pricing in such a way. Dolan (1987) provides three possible stories as to why firms may use nonlinear pricing as a mechanism for price discrimination: heterogeneous consumers, channel efficiency and competitive bidding.

A quote from Buchanen (1953) sums up the heterogeneous consumers' reason best "the demand schedule of small buyers is more inelastic over the relevant price range than that of large buyers". This variation in consumer type allows firms to exploit small demand consumers while having to somewhat cater to those with high demand in lower pricing per unit.

The channel efficiency argument for price discrimination is that firms may have cost savings if consumers buy specific quantities of their goods. In an anecdotal example of this Dolan shows why Sealed Air Corporation (SAC) provides lower marginal prices to those who purchase larger quantities. If SAC's product is sold in quantities that are smaller than a truckload, they must send one truck to multiple locations to serve many consumers, but if one consumer purchases an entire truckload of their product, then they only have to send the truck to one location. Having a truck serve a single location only results in considerable shipping cost savings for the firm. To induce consumers to purchase entire truckloads of their products, SAC lowers marginal price as the quantity that consumers purchase increases.

The last reason for price discrimination discussed by Dolan is competition over market share. Dolan states that if only a portion of firms within a market place utilize nonlinear price schedules to provide quantity discounts, then the firms who do not may not be able to compete with their competitors for consumers whose demand is relatively more elastic. The example he uses is electronic components suppliers. When interviewed as to why they offer lower marginal prices to large buyers of their products, they stated that it was because the other firms do it too.

Another paper that looks at the welfare implications of price discrimination is Sharkey and Sibley (1993) [SS93]. In this paper, SS93 model two types of consumers and test the standard finding within the literature that marginal price per unit is equal to marginal cost per unit for the highest demand users. In their model, SS93 find that the distribution of consumer types determines if marginal price per unit is equal to marginal cost per unit for the highest demand users of the goods. Depending on how types are distributed, equilibriums of marginal price per unit higher than marginal cost per unit and marginal cost per unit higher than marginal price per unit can be supported. SS93 also test to see if price discrimination can exist under competition. With their model, they find that price discrimination can exist under duopoly, but only if there is a regulator that favors high demand users of the good over low demand users.

The final theoretical paper that I will discuss that takes the welfare implications of price discrimination into account is Reiss and White (2006) [RW06]. In their paper they provide a method for welfare analysis that estimates consumer surplus in the face of menu pricing. They too find that the distribution of demand determines the welfare implications of nonlinear pricing.

Due to recent advances in empirical techniques as well as increases in data availability, the empirical literature on price discrimination has been growing in recent years. Some of the earlier work uses reduced-form estimation techniques, as opposed to the newer empirical papers on nonlinear pricing which use structural models to estimate demand for the goods in question. One downfall of the current literature is that almost all of these papers use aggregate sales data rather than actual consumer choice. The first of these types of papers I present Friebel et. al (2015) [FOG15]. FOG15's finds evidence for price discrimination in Russia's wheat market, a fairly homogeneous good similar to fluid milk. Their paper deviates from mine in that they focus on price discrimination in the export market rather than in the consumer goods market. This allows them to use a reduced-form approach that is inappropriate for the questions that this paper attempts to answer. FOG15 finds evidence for price discrimination by Russian firms in 25 out of 61 destination countries within their data set over the 2002-2011 period.

The more closely related literature to my paper are papers that use structural models of demand. These papers include: Leslie (2004), Cohen (2008), Liu and Shen (2012) [LS12], Miller and Osborne (2014) [MO14], and McManus (2007).

Leslie (2004) uses data from the Broadway show *Seven Guitars* to estimate the welfare and profit implications of price discrimination. In this paper, Leslie uses a structural model that allows him to use counterfactual analysis that allows Leslie to estimate the welfare implications for price discrimination. Leslie finds that the price discrimination observed by the firm increases profits by 5% but has little to no effect on consumer welfare, thus price discrimination in this case seems to increase total surplus.

Cohen (2008) estimates how packaging size is used in the paper towel market as a vehicle for price discrimination. In this paper, Cohen uses a discrete choice model with product characteristics and aggregate sales data to estimate model parameters for demand. This allows Cohen to perform counterfactual analyses similar to Leslie 2004. Cohen finds quantity discounts that are consistent with second-degree price discrimination rather than what could be attributed to cost differences across the sizes of the goods. Cohen also finds that consumers are better off with price discrimination as there is more competition amongst firms in the multi-roll package size segment of the market due to the pricing strategy.

In a working paper, Liu and Shen (2012) [LS12] estimate the degree to which firms in the carbonated soda market are able to price discriminate based upon firm type; i.e. private label brand or national brand products. LS12 estimates a discrete choice model using a subset of the data that this paper uses, the Nielsen Scanner dataset, but do not use the Nielsen Panel data and thus, must simulate individual consumer preferences. LS12 also uses the assumption that firms in this market compete in a Bertrand-Nash equilibrium setting that allows them to estimate the marginal cost of the sodas in their model, though this assumption is hard to believe, as these firms clearly do not compete in such a way. This assumption allows LS12 to identify to what extent both private labels and national labels are able to price discriminate. Their main findings are that private labels are able to price discriminate just like national brands within this market, which they claim at least somewhat explains the growth of market shares of private label sodas.

Miller and Osborne 2014 [MO14] use data of the Portland cement industry in California, Nevada and Arizona to estimate a structural model of spatial differentiation and price discrimination. MO14 finds that if price discrimination were not allowed, consumer surplus in the cement industry would increase by approximately $30 million dollars per year, thus implying that price discrimination is welfare decreasing from the consumer's perspective,[18] a result not commonly found within the previous literature.

McManus (2007) uses a structural discrete choice model to estimate whether or not specialty coffee shops are able to price discriminate using sizing for their good. McManus also looks at how nonlinear pricing may not only lead to different marginal prices per ounce along sizing options, but also how it may distort product characteristics away from their efficient levels. He finds that nonlinear pricing allows firms to make size of the smallest cups of specialty coffee "inefficiently small" for the price that is set, which McManus concludes that firms do so in part to incentivize consumers to either buy the next size up or have to pay additional rents to the firm. He also finds that these distortions do not exist for the highest demand consumers, a finding that is consistent with the theoretical literature.

The model that I estimate in this chapter is based on Cohen 2008 model, but deviates in two regards. First, I observe consumer demographics for each purchase, thus I do not have to use Cohen's noisy measure of the population demographics within physical market to deal with potential sources of consumer heterogeneity. As such, my second deviation from Cohen 2008 is that I do not utilize the random coefficients framework for estimating my model, which Cohen uses to address issues of substitution patterns that exist within multinomial logit modeling, the interdependence of irrelevant alternatives (IIA). My reason for not addressing this within this paper is that by utilizing household characteristics, the IIA issue is less of a concern. For example, suppose a regional brand firm introduces a new skim milk product to the marketplace. Household types that that are more likely to consume products similar this type of product, such as households that are more likely to a different regional brand skim milk, are far more likely to switch than households who are more likely to consume a far different product, such as private label whole milk. Since I have

[18]The change in total welfare is unknown to Miller and Osborne, as they do not model firm profits, but can infer that price discrimination must be profit increasing as otherwise firms would not have an incentive to utilize the practice.

household characteristics, I identify which type of household prefers each product type, making the IIA issue far less of a concern.

In summary, the theoretical literature on price discrimination defines price discrimination and has attempted to rationalize this pricing strategy commonly used by firms, whereas the newer empirical literature on price discrimination uses structural models to estimate demand for goods, then discusses the price discriminatory implications of their parameter estimates via counterfactual analyses. The main theme from the theoretical literature is that since the demand for the good is relatively inelastic for low demand users as compared to high demand users, firms are able to set higher marginal prices per unit for low demand users and it largely agrees that the highest demand user will always have an option where marginal price is equal to marginal cost.

The empirical literature largely answers two main questions: are nonlinear pricing schedules used as a way for firms to price discriminate and if so, what are the welfare implications within the market place. The main deviation within the empirical literature are how the supply side/marginal costs are dealt with. Some papers (Liu and Shen (2012) and Cohen (2008)) model and estimate marginal costs of a representative firm, whereas other papers either do not have to worry about marginal costs (the Miravete papers, Ayral (2014) and Leslie (2004)) as the marginal costs are sufficiently small in their industry, or they have discussions with firms to find what the costs of the products are (McManus (2007)). Estimates for marginal cost are important in this literature as many of the findings could also be explained as costs savings by the firm without them.

### 1.4 Data

Data for this paper come from the Nielsen Panel and Scanner Datasets for the years 2008 and 2009.[19] In the subsections below, I discuss the characteristics of these datasets, as well as any other supplementary datasets that have been used.

#### 1.4.1 The Nielsen Panel Dataset

The Nielsen Panel dataset includes 40,000 households from 2004-2006 and was expanded to 60,000 households starting in 2007 and thereafter. These data include each and every item purchased by a household within the panel year at the household-shopping trip level. Since the observation level of these data is the household-shopping trip, the Nielsen Panel dataset includes household demographics for each participating household that participates, as well as the Universal Product Codes (UPC) of all items purchased by the household purchases during the particular shopping trip

[19]The Nielsen Panel itself goes from 2004-2016, but due to a variety of reasons including: potential changes in preferences that may occur over the long-run, macroeconomic factors, and most importantly, the quality of these data are much better over this two year period than any other two year period in the panel. After removing inactive panelists and trimming down to an individual region, 13,352 households with 22,154 household years are left in my sample. As such, I omit all other years from the analysis.

to the store in question. Due to potential differences in preferences that may vary within regions,[20] I use households from the sample that live in FIPS Division 3: Midwest East North Central.[21] This FIPS division has some convenient qualities that are discussed further in a later section of this chapter.

The strength of using the Nielsen Panel data as opposed to datasets used in previous studies is the observation of actual purchasing choices by households rather than aggregate sales by firms. With these data, I am able to use real household characteristics[22] coupled with actual choices rather than simulate consumer heterogeneity as done previously by Liu and Shen (2012), Cohen (2008), Leslie (2004) and McManus (2007). One potential drawback of this panel is the high attrition rate of the panelists. This may be due to the somewhat weak incentives for participation, which are discussed in more detail later in this chapter. The Nielsen company states that it generally retains 80% of their panel from year to year.[23] This is not particularly important to this paper as I am more interested in firm behavior, so I will be treating this panel as a repeated cross section from year to year, though I use the panel aspect of the data to construct a pseudo-purchasing history variable, the time (days) since the household last purchased fluid milk, to include how the stock of fluid milk the household currently has may affect future or current purchasing decisions.

Nielsen uses a stratified, proportional sampling technique to create a representative sample of the continental United States. Households are randomly selected and invited to join the panel either via mail or email invitation. Nielsen does not directly pay the households selected to participate in the panel, but it does provide some other incentives. These incentives include: monthly prize drawings, gift points awarded for weekly transmission of data and a sweepstakes. Nielsen states that they also try to encourage participation and create enthusiasm through ongoing communication. They use the following methods to do so: a monthly newsletter, telecommunications, Q & A section with helpful tips and reminders, personalized computer tips and reminders after transmitting, notice of monthly sweepstakes winner, personalized letters for reporting problems and questions, letters from the president, gift point statement, help desk, an 800 number for panelists to call and exit interviews.

In addition to providing rich household characteristics, this dataset also has a variety of product, brand, and retailer demographics as well as the overall purchases by households.[24] Sales of UPCs sold that are in sizes that do not utilize the Imperial

[20] And also in part due to the size of these data.

[21] The states included in this region are: Illinois, Indiana, Michigan, Ohio and Wisconsin.

[22] Which include: household income, household size, type of residence, household composition, age and presence of children, male and female head employment status, male and female head education, male and female head occupation, marital status of the head of household, race, location of the panelist, a dummy variable for WIC program participation, and the age of the male and female heads of household. See section 1.8 for tables with summary statistics for the panelists.

[23] In my subsample of these data, I find that 2,306 households only participate in the year 2008, 2,244 household only participate in 2009 and 8,802 households participate in both years.

[24] Observable with this dataset are: the date of each purchase, a store code and location which provide information on the type of retailer, the amount spent by the consumer on each shopping trip, a product UPC for each item purchased, a retailer code which provides information on the type of retailer (grocery store, mini mart, etc.), the price paid for each item, the packaging size for each

measurement system are discarded from this sample for estimation purposes, but kept for the construction of other variables.[25] With that being said, the discarded portion of the data are less than 1% of the overall sample. Additionally, the removal of these sizes allows me to clearly define the top of the menu, as previously it was somewhat muddled. Summary statistics for this data cleaning procedure are included in section 1.8.

### 1.4.2 The Nielsen Scanner Dataset

Ideally, I would observe not only what households choose to purchase during a shopping trip, but all of the choices the household did not choose as well. The Nielsen Panel dataset only provides information on the choices that households make on a particular trip, rather than the full menu of options that household faced. Since this is the case, I utilize the Nielsen Scanner dataset to simulate the remaining portion of the choice set the households did not choose. The Nielsen Scanner dataset contains the weekly sales of each UPC sold at a store as well as product characteristics similar to what are provided within the Panel Dataset. More information on how the choice sets were constructed are found in section 1.8.

### 1.4.3 Costs of Raw Fluid Grade Milk

The USDA regulates the price that farmers sell butterfat, bulk and skim milk to dairy processors. The prices of these goods are all regulated by the USDA are published in the National Agricultural Statistics Service's (NASS) monthly report. I use these sales prices as a way to control for the input cost of fluid milk.

## 1.5 Preliminary Evidence of Price Discrimination

This section provides simple reduced-form evidence that is consistent with the idea that firms within the fluid milk market are using nonlinear pricing as a way to price discriminate. First, I provide several graphs and tables that are consistent with price discrimination, then I provide regression results from a simple hedonic pricing model. The reduced-form regression results presented in this section are consistent with price discrimination as well.

Table 1.1 and table 1.2 present unit-price statistics and figure 1.1 provides graphs of the mean price, price-cost differential, and price-cost ratio for units sold by the firms in these data. The graphs break down the price statistics into several categories, purchases made with and without coupons[26] and purchases of regional brand vs private label products. Since I assume that fluid milk is only sold in containers

item, an organic claim dummy variable for each product, various product demographics, and the brand of each product.

[25] UPCs with containers of the following sizes remain: gallon, half gallon, and quarts.

[26] Nielsen does not provide a clean definition as to what a "coupon" is beyond a piece of paper that is given at the register to provides a discount.

that are sized using the Imperial measurement system, all observations fall into one of three categories: gallon, half gallon or quart.

Since the average price-cost differential and price-cost ratio generally decrease as the size of a product increases, figure 1.1 is consistent with firms using nonlinear pricing to price discriminate. Additionally, for gallon containers, which are assumed to be the largest sized offered by firms, the price-cost differential is very close to zero, suggesting that these products are close to the efficient price-quantity bundle, an observation that is consistent with the current literature. It is interesting to note that in the graphs on the right, it appears that firms are not using nonlinear pricing to price discriminate between cup and pint sized units. These means are very close however and as a whole, there is a downward trend. Additionally, these are just means and do not control for differences other than size.

Another interesting portion of the data is the pricing behavior of private label goods. It is interesting to note that not only are these graphs consistent with a story of firms using nonlinear sizes to price discriminate, but that it holds for the smallest sized containers as well. This suggests that firms are not only able to use nonlinear pricing to price discriminate with their private label brand products, but they are able to do so to a greater degree than regional brands. This observation agrees with the main finding of Liu and Shen (2012) in their study of the soda industry.

After observing these simple trends in price statistics along sizes, I estimate the following price hedonic model:

$$PPO_{ist} = \beta_1 Size_{is} + \beta_2 CPO_{it} + X'_{ist}\gamma + \epsilon_{ist} \tag{1.1}$$

where $PPO_{ist}$ is the average price per ounce of UPC $i$, of size $s$ at time $t$. $Size_{is}$ are a vector of size dummies for product $i$ of size $s$.[27] $CPO_{it}$ is the price of raw fluid grade milk per ounce of UPC $i$ at time $t$. $X'_{ist}$ are a vector of various product characteristics, year and market controls that are included in the table.

Additionally, price may vary from brand to brand, thus I also estimate the following model with brand fixed effects:

$$PPO_{ist} = \beta_1 Size_{is} + \beta_2 CPO_{it} + X'_{ist}\gamma + \delta_i + \epsilon_{ist} \tag{1.2}$$

where $\delta_i$ is a brand fixed effect for UPC $i$. The results of these regressions are presented in table 1.3. The first two columns of table 1.3 are the simple price hedonic models,[28] and third and fourth column of table 1.3 are price hedonic models with brand fixed effects.[29]. All of the coefficient estimates for size in ounces negative and statistically significant at the 1% level of significance, except in the case of models with controls for 1.5% fluid milk.[30] Since all of these regressions include the published raw fluid grade milk prices, these findings are consistent with firms using nonlinear

[27] In these data, size for each UPC does not vary over time.

[28] The first column without any controls and the second column with controls included.

[29] The third column without any controls and the fourth column with controls included.

[30] May be due to sample size since only 2,487 observations of 1.5% fluid milk in my sample of approximately 2.365 billion observations.

pricing schedules to price discriminate in this market. This finding is not only robust to model specification,[31] but also to the different butterfat contents of fluid milk.

These findings do not come without limitations however. First, all of these models implicitly assume that container size is continuous and that the product set is dense in ounces. As such, the interpretation of these coefficient estimates are equivalent to estimating the consumers average marginal willingness to pay for an additional ounce of fluid milk. Since these data are measured using the Imperial measurement system, the variation in size that these models are using to identify these coefficients is in large discrete changes in size, rather than small changes in size. Since this is the case, it is clear that these coefficient estimates are not properly identified.

Following an argument originally presented by McManus (2007), this econometric strategy would be appropriate if firms offered a menu of sizing options that were continuous in size, or it were possible to purchase any amount in ounces at the very least. This is the limitation of reduced-form econometric identification strategies in this scenario, as the model is attempting to identify an effect that is implicitly continuous, with variation that is in large unit changes. Another downfall of this identification strategy is for sales of the largest container size, as I do not observe a size that the household is unwilling to pay. Since I will never see a price they are unwilling to pay for an additional ounce of fluid milk it is impossible to identify their marginal willingness to pay for an additional ounce of milk.

With these two main points in mind, these results should be taken as a signal that firms may be using nonlinear pricing schedules to price discrimination in this market, and a more appropriate approach should be taken to show that this is the case. The structural approach proposed in the next section of this paper does not suffer from these weaknesses, and the marginal willingness to pay for an additional ounce of fluid milk is properly identified even with these large discrete changes in size.

According to Berry (1984), there are several other strengths to estimating structural models instead of reduced-form models for demand estimation in addition to previous concerns. He states that it is possible to estimate a reduced-form model for demand, but due to the endogeneity of price that I discuss later, one must find a set of valid instruments to estimate the price elasticities of each good. The problem with just finding instruments and using reduced-form techniques in this case is that in a market with $N$ goods there are $N^2$ elasticities to estimate. Parameterizing the consumer utility function allows one to estimate the $N^2$ cross-price elasticities from far fewer parameters.

Another strength to structural estimation that Berry points out in his paper is the ability to perform counterfactuals. This strength of structural estimation is particularly crucial to this paper, as I intend to estimate the welfare effects of price discrimination within this market as well as check to see if firms could be better off without following the Imperial measurement system for their menu of options.

The last strength of structural estimation that Berry discusses is the model's ability to allow one to move easily between statements about aggregate demand and

[31] Controls or no controls, and OLS versus brand fixed effects

statements about consumer utility. This will be very important for this paper in the welfare analysis section.

With both Berry and McManus' comments in mind, it seems appropriate to estimate a structural model for demand rather than a reduced-form model in this context, as I will be able to cleanly identify the marginal willingness to pay for an additional ounce of fluid milk for each consumer type. Additionally, I will also have the ability to perform the interesting counterfactuals that are not possible in a reduced-form model.

## 1.6 Structural Model

In the first subsection of this section of the paper, I estimate the demand for fluid milk using a discrete choice model. In the subsection that follows the demand portion of the model, I introduce a model for estimating firm's marginal costs within the fluid milk market. Both the consumer and firm sides of the model follow closely to Cohen (2008). It is important to estimate the demand side of the market separately from the supply side of the market due to the endogeneity of prices.[32] Following a reduced-form approach in estimating this type of model will cause estimates to be biased due to simultaneity of the demand and supply portions of the market.

### 1.6.1 Demand for Fluid Milk

I assume that each household, $i$, chooses to purchase one or zero units of a fluid milk UPC during each shopping trip.[33] For simplicity, my model does not consider the household's decision of whether or not to go on a shopping trip, as well as which store to patron during a trip if they choose to go on one.[34] The decision to purchase zero units of a fluid milk UPC is described as choosing the outside good.[35] Purchasing options are indexed by $j \in J$. I assume that all stores that households choose to patron for a shopping trip offer all possible choices within the choice set $J$.[36]

[32] In the model that is presented in this section, I will assume that there are some characteristics of each product that are observed by both firms and consumers but unobserved to the econometrician. It is fair to assume that a profit maximizing firm will take these characteristics into account when determining what price to set when constructing each price-quantity bundle. Thus by construction, price is correlated with part of the error term of the model and is endogenous.

[33] I define a shopping trip to be going to a store that offers a menu of products that includes fluid milk

[34] This is important to point out as Thomassen (2017) states that a source of potential bias in cost-side parameter estimates exists from ignoring the cross-category pricing effects of other goods the firm sells. A portion of the motivation for this type of bias is that households consider the whole basket of goods they intend to purchase when choosing which store to patron. By removing the household's choice of when and where to go on a shopping trip, and considering how small of the portion of the average household's budget is in terms of fluid milk expenditures, this potential source of bias is far less of a concern for my model.

[35] Which as previously defined, includes all consumable liquids other than fluid milk from a cow.

[36] As described in Section 1.4.2, the Nielsen Panel dataset does not allow me to observe the entire menu of goods available for households to choose from, but only those actually chosen by the household. As is the case, I use the Nielsen Scanner dataset to simulate the menu of products the

Each fluid milk UPC is grouped into a choice, $j$, within the choice set, $J$ to reduce the vast number of UPCs of fluid milk that the household could potentially choose from. Each choice, $j$, is characterized by: (1) $\mathrm{BFC}_j$, which is a dummy variable that is equal to one if the UPC has a butterfat content of 2% or greater; (2) $\mathrm{size}_j$, which is a dummy variable that is equal to one if the UPC is sold in a gallon sized container and zero if the UPC is sold in a half gallon or quart sized container;[37] (3) $\mathrm{PL}_j$, which is a dummy variable that is equal to one if the UPC is a private label branded[38] product; (4) $\mathrm{org}_j$, which is a dummy variable that is equal to one if the UPC is an organic product and zero otherwise; (5) $\xi_j$, product characteristics of the UPC that are observed by economic agents but unobserved to the econometrician;[39] (6) the unit-price, $p_j$.

First, I consider the model for demand assuming without controlling for household demographic characteristics, which I will refer to as the standard model for the remainder of this paper. I do so because the vast majority of the existing empirical literature on price discrimination has been done with aggregate market level data which implicitly some sort simulation for household decisions. The existing literature does so because household level decision data were not available to authors of those papers. Since I have the household level decisions and demographic characteristics available via the Nielsen Consumer Panel and Scanner dataset, my preferred specification will include household demographic information. However, by specifying and estimating a structural model of the demand for fluid milk without including household demographics, and comparing those estimates to my preferred specification which includes household demographic characteristics, I am able to observe the potential bias of the parameter estimates that is caused by the aggregation in using market level data in a somewhat back-of-the-envelope manner.

I then assume that the indirect utility associated with purchasing product $j$ to be linear in product characteristics in both models. The specification of the model without household demographics is presented in the equation below:

$$U_{jt} = \beta BFC_j + \tau size_j + \zeta org_j + \alpha p_{jt} + \xi_j + \epsilon_{jt} \tag{1.3}$$

Where $\epsilon_{jt}$ represents the standard model's deviations from the mean preference for good $j$, which I assume to be distributed iid with a type I extreme value distribution (TIEV). Under this assumption for the distribution of the error term, the standard model's level indirect utility becomes a multinomial logit model (MNL) where the

store had available on the date of the shopping trip as outlined in that section.

[37]Note: I discard all shopping trips where households purchase fluid milk of sizes that are not gallon, half gallon are quart sized, but I do include them when calculating the measure of how long it has been since they last purchased fluid milk.

[38]Or "store" brand.

[39]I utilize a fixed effects approach to control for the product characteristics that are unobservable to the econometrician but observed by economic agents. In particular, I include a private label dummy as well as the private label dummy interacted with various channel type indicator variables. Using a fixed effects approach with these particular fixed effects is valid if $\xi_j$ only varies at the private label vs. regional brand and channel type levels.

market shares for each good $j \in J$ are defined as:

$$S_j = \frac{exp(\delta_j)}{\sum_{k=1}^{J} exp(\delta_k)} \tag{1.4}$$

where $\delta_j = \beta BFC_j + \tau size_j + \zeta org_j + \alpha p_j + \xi_j$ is the population mean utility for product $j$. I then solve for these parameters via maximum likelihood.

To introduce household demographic characteristics into the model, I assume that household $i$ receives indirect utility from purchasing product $j$ as indicated below:

$$U_{ijt} = \beta BFC_j + \tau size_j + \zeta org_j + \alpha p_{jt} + \pi_{ij} + \xi_j + \epsilon_{ijt} \tag{1.5}$$

where the main difference between this specification and the standard specification is the $\pi_{ij}$, which is defined below:

$$\pi_{ij} = (H_i \times size_j)\rho + (H_i \times BFC_j)\gamma + (H_i \times org_j)\eta \tag{1.6}$$

where $H_i$ is a vector of household characteristics.[40] Similar to the standard specification, I assume that $\epsilon_{ijt}$ represents household $i$'s deviation from the mean preference for good $j$, is distributed iid with a type I extreme value distribution (TIEV).

One deviation from the standard specification is that I no longer predict market shares, but instead the probability that each household $i$ chooses good $j$ during a shopping trip. Since I make similar assumptions on the household specific taste shock, this specification is also a MNL model and the probability that household $i$ chooses product $j$ is defined as:

$$S_j = \frac{exp(\delta_j + \bar{\pi}_j)}{\sum_{k=1}^{J} exp(\delta_k + \bar{\pi}_k)} \tag{1.7}$$

where $\delta_j = \beta BFC_j + \tau size_j + \zeta org_j + \alpha p_j + \xi_j$ is the population mean utility for product $j$ and $\bar{\pi}_j$ represents the mean household characteristic's contribution to the utility of product $j$. I then solve for these parameters via maximum likelihood as well.

A well-known issue of the MNL model is known as the "independence from irrelevant alternatives" (IIA)[41] IIA, caused by the iid TIEV assumption of the MNL error term, is the unintuitive substitutional patterns that predicts that households will respond to an increase in the unit-price of product $j$ by substituting to the most popular products rather than substituting to products that have characteristics that are similar to product $j$. This is far less of a concern for my model's specification however, as the $\bar{\pi}_j$ term, which includes $H_i$, constructs heterogeneous preferences specific to each household based upon their demographic characteristics. In turn, substitutional patterns within my MNL model are not only a function of the mean

[40] Which includes: the number of days since the household last purchased fluid milk, income, education, the average age of the household heads, whether or not children are present in the household and the household head's marital status.

[41] Which has lead many of the papers previous in the literature to utilize a either a random coefficients logit model, an ordered logit model or some other model specification of a similar nature.

utility of the products within the choice set, but also a function of each household $i$'s preferences, which are constructed and estimated by including the $\pi_{ij}$ term. As a result, I find that it is both unnecessary and inefficient to implement an approach that would essentially constructs heterogeneous preferences among households via simulation[42] or via an a priori grouping of choices within the choice set,[43] as I am able to estimate these heterogeneous preferences using the household demographic information instead.[44]

### 1.6.2 Firm Behavior

On the firm's side of the model, I assume that firms take the menu of products sold by the firm as given.[45] Additionally, I assume that firms participate in price competition. This allows me to infer marginal costs without firm-level cost data if: (1) unit marginal costs are constant for each product $j$, (2) unit marginal costs are non-constant, but a function of container size over the product set $J$. Together these two assumptions imply that the cost of producing the first half gallon container of private label whole milk is the same as the cost of producing the second half gallon container of private label whole milk for a particular firm $F$. However, the cost of producing one gallon container of private label whole milk is not restricted to be equivalent to the cost of producing two half gallon containers of private label whole milk for the same firm $F$.

From here, I define marginal costs for product $j$ sold by firm $f$ in state $s$ during week $t$ as a function of the vector $X^{Cost}$:

$$mc_{fsjt} = X^{Cost}_{fsjt}\gamma + \omega_{fsjt} \tag{1.8}$$

where $X^{Cost}_{fsjt}$ consists of the butterfat cost for each choice $j$ that varies at the firm/state level, the skim cost for each choice $j$ that varies at the firm/state level, the total number of firms within the three digit zip code and channel dummy variables interacted with a gallon sized dummy variable.

I can then use the demand system/parameter estimates that determine consumer choices to infer marginal costs. For simplicity, I refer to the market shares that are constructed by the demand parameters as:

$$S(\alpha, \theta; p, X, H; \xi) \quad or \quad S(\Theta) \tag{1.9}$$

[42] As would be the case in a random coefficients logit specification.

[43] As would be the case in an ordered logit model specification.

[44] Note: This approach assumes that household preference heterogeneity can be exactly modeled via household demographic characteristics. As such, if household preference heterogeneity varies within demographic characteristics, then this approach still lead to odd substitutional patterns. For example, perhaps households choose to consume organic products because they are "going green" rather than for some other reason which would be captured via demographic information. However, the within demographic household heterogeneity would have to be the main determinant for households for this approach to be invalid, which I find unlikely.

[45] Cohen (2008) states that if the decision whether or not to offer multiple sizes of products is treated as a choice variable in the model that this choice is likely correlated with the product unobservables $\xi$ and would lead to biased parameter estimates.

where $\theta$ are the utility parameters estimated in the demand system (excluding the disutility of price parameter $\alpha$). $p$, $X$ and $H$ refer to the prices in the observed prices, observed product characteristics and household demographics respectively.

Before continuing, two key points must be made: (1) due to a lack of data on the menu of products that each firm offers, I use the simulated menu described section 1.4.2 and assume each firm offers all choices within the set $J$ at some time during my sample period. However, if store has not sold a unit of a particular choice within a month of the current week, I assign them a zero for in the "product ownership" matrix for that time period. Additionally, I aggregate to the store-week level. Assuming that unit marginal costs are constant over output, firms' profits are proportional to:

$$\sum_{h \in J} \left[p_{fsht} - mc_{fsht}\right] \times S_{fsht}(\Theta) \tag{1.10}$$

where $h \in J | h \neq j$. Then the first order condition for $p_j$ where $j \in J$ is given by:

$$\left[\sum_{h \in J} (p_{fsht} - mc_{fsht}) \frac{\partial S_h(\Theta)}{\partial p_j}\right] + S_j(\Theta) = 0 \tag{1.11}$$

Which expressed in vector form is:

$$(-\Delta)\left[p - mc\right] + S = 0 \tag{1.12}$$

where $\Delta_{jr} = -\Omega^*_{jr} \frac{\partial S_j}{\partial p_r}$. $\frac{\partial S_j}{\partial p_r}$[46] is the $(J-1) - by - (J-1)$ matrix of own- and cross-derivatives and $\Omega^*_{jr}$ is the firm ownership matrix which elements are equal to 1 if firm $i$ sells products $j$ and $r$ and zero otherwise.

From here, I obtain marginal costs to be:

$$mc = p - \left[\Delta^{-1} S\right] \tag{1.13}$$

where $\left[\Delta^{-1} S\right]$ is the markup term: $M(\alpha, \theta; p, X, H; \xi)$ or, $M(\Theta)$.

The strategy I use for estimating the cost side of the model is to find the parameters that minimize the prediction errors (i.e. the distance between the vector of observed and predicted unit-prices) from the following:

$$p - M(\Sigma, \delta) = MC(\gamma; X^{Cost}; \omega) \tag{1.14}$$

After estimating these markups, I find that some of the time they are negative.[47] However, there is a strong correlation[48] between negative markups and whether or not the product is on sale that week. From this, I draw the conclusion that when firms advertise/put fluid milk on sale they treat the product as a loss leader, whereas during normal weeks they price discriminate.

[46] Which following Choi et. al (2013) is computed as $E_{jr} * \frac{s_{rj}}{p_r}$

[47] Which varies based upon the product characteristics.

[48] A sample correlation coefficient of .87.

### 1.6.3 Demand Results

Note that the parameter estimates presented in table 1.4 cannot be directly interpreted as marginal effects, but finding the marginal effects are quite simple as doing so requires dividing these estimates by the absolute value of $\alpha$, the price sensitivity parameter. These marginal effects are shown in table 1.5. The first set of columns in tables 1.4 and 1.5 present the results from the standard model whereas the next set of column are my preferred specification with includes household demographic characteristics. I will frequently compare the estimates obtained from each model throughout the remainder of my discussion of these estimates. Before comparing the results of the standard model to my preferred specification, it may be useful to discuss the interpretations of several of the estimated parameters. Therefore, I begin with a discussion of the estimated marginal benefits of my preferred specification.

First, consider the estimated marginal effect associated with the gallon sized dummy variable, which is 1.33417 in my preferred specification. This parameter estimate implies that, on average, households are willing to pay approximate $1.33 more to purchase goods sold in gallon sized containers rather than the smaller sizes. The estimate of this marginal effect seems quite reasonable both in terms of sign and magnitude, as the average difference in prices between gallon sized containers and is similar in magnitude for goods that share an organic status.

The marginal effect of 0.17433 associated with the butterfat content dummy variable seems reasonable in terms of magnitude and since it is positive, is consistent with the consumer theory which states that on average, consumers prefer more product characteristics to less. The parameter estimate associated with the organic product characteristic may appear somewhat odd, as ex-ante one may believe that it should be positive. The reason for this is that this parameter estimate has little to no meaning alone and must be interpreted in conjunction with the household characteristics parameters associated with organic milk as well. This topic is discussed in greater detail in the subsection that follows.

The marginal effects with the "UFE" label are fixed effects that control for product characteristics that are observed by economic agents within the model, but unobserved to the econometrician.[49] One of the major issues with demand estimation of markets is endogeneity caused by not controlling for the unobserved characteristics, with the major concern of endogeneity being shown via biased estimates of the price sensitivity parameter, $\alpha$. The reason for this concern is that often times these unobserved product characteristics are correlated with some notion of quality that the econometrician cannot properly control for via standard estimation procedures. If economic agents participating within the marketplace do in fact associate the unobserved product characteristics with quality, then it is likely that those characteristics will increase the prices of goods where they are present. This is because both households and firms can observe the unobservable and respond accordingly. As such, this source of endogeneity is said to bias the estimates of the price sensitivity parameter towards zero on average, and could potentially cause the parameter estimate to re-

[49]Note: The "UFE: (Channel)" fixed effects are all interacted with a private label indicator variable.

turn as positive. A positive parameter estimate for $\alpha$ would be particularly alarming as it would suggest that households prefer to spend more money on goods within this market, as opposed to spending less, ceteris paribus. I report nearly identical parameter estimates for $\alpha$ of approximately -1.96 in the model both of the estimated models, which I find to be quite reasonable in terms of magnitude and are clearly the correct sign.

However, utilizing this method to control for potential product unobservables, rather than an IV style approach, does have come with the caveat that my models implicitly assume, and are only valid if, households and firms perceptions of the unobservable product characteristics differ only between store channels and brand type.[50] Thus, the fixed effects approach I implement to control for the endogeity of prices is not valid if a subset of the consumer population has a particular taste for the unobservable characteristics, which if true would lead to biased parameter estimates. However, the price sensitivity parameter estimates in both specifications of the model have the proper sign and are reasonable in magnitude. As such, my fixed effects approach that controls for unobservable product characteristics appears to be appropriate for treating the this potential source of endogeneity.

Additionally, the parameter estimated associated with the set of fixed effects are negative in sign, which implies that consumers prefer to purchase fluid milk at stores who operate with the base channel, in this case grocery stores, which is what one would expect ex-ante.

Marginal effects with the label "GAL Char" are household demographic characteristics that have been interacted with the gallon sized container indicator variable, whereas "BFC Char" marginal effects are the same household demographic characteristics interacted with the butterfat content $\geq 2\%$ indicator variable instead. I do not have strong ex-ante beliefs for the expected sign of these marginal effects for the majority of these demographic characteristics, thus I will only highlight selected characteristics. One would anticipate that "Days Since Last Milk Purchase" would have a positive marginal effect for the gallon size interaction. This is because a positive marginal effect associated with this temporal measure's interaction with the gallon sized container indicator variable would imply that as a household's home stock of fluid milk diminishes, they would have stronger preferences for fluid milk sold in gallon sized containers. However, my results report this marginal effect to be negative and very small in terms of magnitude (-0.00975), suggesting that the households home stock is largely unimportant in comparison to the other interacted paramater estimates.

Table 1.6 reports for the own- and cross-price elasticities implied by the marginal effects reported in table 1.5. The own-price elasticities found in table 1.6 are somewhat larger in magnitude than what has been previously found in the literature, but are consistent economic theory in terms of their relative magnitudes.

The consistency of the relative magnitudes of the own-price elasticities can be

[50] i.e. I am assuming that all firms with the channel have the same exact unobservable characteristics for their goods and that any potential differences within channel must vary along the private label vs. regional brand dimension.

observed by comparing the organic choices within the choice set to the non-organic choices that are otherwise identical. Economic theory would suggest that households would be more price sensitive to the changes in price of organic fluid milk UPCs than they would be to non-organic fluid milk UPCs. Since there are strict regulations on how organic dairy cows are to be taken care of and how organic fluid milk is to be processed in comparison to their non-organic counterparts, it is reasonable to believe that households view organic milk as more of a luxury-type good when compared to a non-organic fluid milk product. To highlight this consistency, consider the example of comparing the own-price elasticities of choice 5 to choice 4 found in table 1.6.[51] Since choice 5 and choice 4 are identical to households in all product characteristics except organic status, economic theory would suggest that households would more elastically demand choice 5 than choice 4. Table 1.6 confirms this suspicion since choice 5's own-price elasticity is -6.965, whereas choice 4's own-price elasticity is -4.431, which suggests that consumers are relativity more sensitive to price changes for choice 5 than they are to changes in price of choice 4. When the remaining three organic choices in the choice set's own-price elasticities are compared to their non-organic counterpart,[52] the organic but otherwise identical choice's own-elasticity is larger in terms of relative magnitudes to their non-organic counterpart. This suggests that holding other product characteristics constant, households more elastically demand organic fluid milk products than non-organic fluid milk products.

As a result, even though own-price elasticities I compute are larger in magnitude than those previously reported in the current economic literature, they are consistent with economic theory of how price sensitivity should vary across the choice set for households. As such, I find my estimates of these own-price elasticities for fluid milk to be more compelling than those previously estimated for fluid milk within the current economic literature, since my for estimates these elasticities are constructed with actual consumer choices rather than aggregate data commonly used in the other work.

One concern is that a nested and/or mixed logit could be a superior approach and obtain theoretically more plausible estimates for the elasticities. However, there are several ameliorating differences in my approach. First, I have the micro level consumer choice data which allow me to utilize actual household choice data when estimating my discrete choice model and, second, and more importantly, the elasticities are not central to the analysis here. Finally, on the flip side of having the micro level consumer choice, I have 70,000 observations in my 5% random sample of the data, while the majority of work with, e.g., nested logit, deals with about an order of magnitude fewer observations, reducing significantly the number of nodes in the decision tree. As a result, computational time expands exponentially. Finally, unlike, e.g. in McFadden (1981), where the ordering is natural, here it would seem to be arbitrary as to how to construct the decision tree.

[51] Choice 5 represents a regional branded, gallon sized, low-fat, organic fluid milk UPC, whereas choice 4 represents a regional branded, gallon sized, low-fat, non-organic fluid milk UPC.

[52] I.e. by comparing choice 3 to choice 2; choice 9 to choice 8; and choice 10 to choice 11's own-price elasticities.

### 1.6.4 Comparison of Demand Parameter Estimates

In this subsection of the paper, I compare and contrast the results of the standard model to my preferred specification to show why it is important to utilize household level data when feasible.

To begin, consider the parameter estimate for the gallon sized container dummy variable between the models. This estimate increases drastically[53] after introducing the household demographic characteristics to the model. This suggests that a market level analysis that does not consider household demographics and micro level decisions may lead to estimates that are biased towards zero.

Additionally, consider the difference between the parameter estimates of households tastes for butterfat. In the specification that does not include household demographic characteristics, I find that households prefer to have less butterfat to more by approximately \$.0044.[54] Whereas in the model that includes households demographic characteristics, I find that households have a preference for more butterfat rather than less and are willing to pay approximately \$0.18 more for it. Which again provides evidence that ignoring household level micro choice data can cause parameter estimates to be biased, and in this case the bias causes the parameter estimates to disagree in sign.

Next, consider the difference between the parameter estimates of households tastes for organic products. In the standard specification, I find that households prefer non-organic products to organic products and are willing to pay approximately \$.32 more for the non-organic goods on average. However, in the model that includes household demographic characteristics, I find that households have an even stronger preference towards non-organic goods on average and are willing to pay approximately double, \$0.64, for non-organic goods on average.

Both results appear troubling at first glance, however the results associated the standard model are more troubling, as it is not possible to consider how different types of households may sort amongst the goods available within the choice set since they only consider an overall average marginal effect. However, in the case of my preferred specification, which includes household demographic characteristic interaction terms labeled "ORG Char:", I am able to piece together an explanation this seemingly counter-intuitive result. By observing the relative magnitudes of the marginal effects associated with the interaction terms for organic products, it appears that the main household demographic characteristics that shape household preferences towards organic goods are education,[55] income[56] and marital status. The marginal effects in the household level model indicate that households that have a yearly income greater

[53]From a marginal effect for a preference for gallon sized containers over the smaller sizes of \$0.97 to \$1.33 after introducing the demographic characteristics to the model.

[54]Though one could argue that this coefficient estimate is economically insignificant even though it is statistically significant.

[55]As indicated by the coefficient estimate associated with the dummy variable titled "At least 1 College" which equals 1 if one of the household heads have attended college in some capacity and is zero otherwise

[56]Which is measured by a dummy variable that equals 1 if the household's income is greater than the U.S. median income during the time period.

than the U.S. median and have at least one household head that has attended college develop prefer organic fluid milk goods to non-organic products, whereas the opposite is true for households where the household heads are married. I am agnostic as to how these household preferences are developed within these demographic groups, but it is important to note that I am only able to make these types of statements due to the introduction of household demographic characteristics to the model. Much of the previous literature on price discrimination either is unable to include this type of information on households due to data availability or has determined that household demographic characteristics are mostly unimportant, though my results clearly indicate otherwise.

Lastly, I will consider differences between the price parameter estimates and the channel/private label fixed effects. As previously discussed, these fixed effects are included to control for product characteristics that are unobserved by the econometrician but observed by all economic agents within the model. The price parameter estimates are almost identical between the standard specification and my preferred specification, a finding that was not anticipated as it is easy to think of a story where not including the household demographic characteristics may bias price sensitivity estimates. The unobserved product characteristic fixed effects (UFE) are similarly almost identical as well. This allows me to infer that the UFEs must be almost completely uncorrelated with household demographic characteristics, as strong correlation between the UFEs and household demographic characteristics would lead to endogenous parameter estimates in the standard specification, which would vary greatly from those presented in my preferred specification. If my assumptions regarding the treatment of the UFEs are valid in my preferred specification, the results as presented in table 1.5 for the UFE parameters estimated in standard model would suggest that the endogeniety has been dealt with in this model as well.[57]

### 1.6.5 Supply Results

Table 1.7 presents the decomposition of the marginal costs that I estimate via equation (1.13) and are used during the counterfactual analysis. The coefficient estimates associated with the gallon sized container dummy, butterfat dummy and organic dummy are each positive and the relative magnitudes of the coefficient estimates associated with the gallon sized container and organic dummies are significantly larger in terms of magnitude than the coefficient estimate associated with butterfat. These results suggest that container size and organic status have a much larger effect in determining the marginal cost of a fluid milk UPC than butterfat status.

The sign of the parameter estimate associated with butterfat content suggests that whole and 2% fluid milk products have higher marginal costs to firms on average than skim and 1% goods. This parameter estimate is consistent with the idea that whole and 2% fluid milk products have a higher opportunity cost of production than skim and 1% products, because whole and 2% products require more butterfat than

[57]Which is an assumption that I confident in, since the price parameter estimate is reasonable in magnitude and has the correct sign.

the latter, which cannot be used by the firm to produce other dairy products once used to produce the higher butterfat content goods.

Furthermore, private label dummy variable is negative, implying that it is cheaper for firms to produce and sell their own private-label branded products in comparison to those sold by a regional brand. Similarly, both the raw milk cost and total number of firms in a 3-digit zip code are positive as expected ex-ante.

The only somewhat odd finding in Table 1.7 is that some of the channel dummy variables that are included in the regression have coefficient estimates that are negative in sign, implying that these channels have lower marginal cost than the omitted case which are firms within the grocery channel. However, the magnitudes for the warehouse store and other store channels are very small and imply that firms within these channels may have lower marginal costs on average than grocery store firms, but if they do it is by a margin of approximate \$0.01-\$.04 per unit produced. Even though these estimates are statistically different from one another, they are very small in terms of magnitude, suggesting that there is not meaningful economic difference in marginal costs of production of fluid milk UPCs between the grocery and warehouse store channels.

The coefficient estimate associated with the drug store channel however is not only statistically significant, but fairly large in magnitude in comparison to the parameter estimates of the other channel types. This clearly implies that drug stores have significantly lower marginal costs for producing fluid milk than grocery stores. I found this to be odd at first, but after looking at the differences in prices between these channels this is no longer troubling, as on average, drug stores sell their products for much lower prices than grocery stores.

### 1.6.6 Measure of Price Discrimination

Following Cohen (2008), I compute unit-markups and unit-costs for each size for each trip. The unit-price discount is equal to the difference in marginal costs between the two size groupings is ($MC_{smaller\ size} - MC_{larger\ size}$) plus the difference in markups ($Markup_{smaller\ size} - Markup_{larger\ size}$), all of which are measured in quantities per-unit. By dividing the difference in unit-markups, by the difference in unit-prices (outlined previously), I obtain a measure of the extent to which price discrimination is a determinant in unit-price differentials across sizes. I compute this measure for private label and regional brand products separately. In doing so I find that 7.2% of the markup for private label products can by explained by price discrimination, whereas only 0.34% of the markup for regional brand products can be explained by price discrimination. This implies that firms are only able to price discriminate on their own private-label branded goods, a result consistent with Lui and Shen (2012).

### 1.6.7 Welfare Counterfactual

One of the motivating reasons behind structural estimation techniques are the ability to perform counterfactual experiments once the parameters of the model have been estimated, this section of the paper includes these counterfactuals computed in aggregate. Since there has yet to be a consensus within either the empirical or theoretical literature on whether or not price discrimination is welfare increasing from the consumer's perspective.[58] I test to see the effect of price discrimination from the consumer's perspective via counterfactual analysis. As a brief highlight of the current literature, Leslie (2004) and Cohen (2008) found that price discrimination within their market of focus to be welfare increasing from the consumer's perspective, whereas McManus (2007) and Miller and Osborne (2014) found it to be welfare decreasing.

To test the welfare implications of this pricing strategy from the consumer's perspective, consider the following thought experiment: suppose there is a world where firms who participate within the fluid milk market are no longer allowed to utilize nonlinear pricing schedules, or that firms within this market must charge the same price per ounce for every size container they sell.[59] One approach to finding the overall affect of welfare would be to compare the current world to the fictional world outlined above in terms of total welfare. I attempt to do so by following Ben Akiva (1972), McFadden (1973) and Domencich and McFadden (1975)'s "log sum formula" for finding consumer surplus, a process similarly used by Choi et. al (2013):

$$Emax\ V_i = \frac{1}{\alpha}\left[log\left(\sum_{j\in J} exp\big(V_j^1\big)\right) - log\left(\sum_{j\in J}\big(V_j^0\big)\right)\right] \tag{1.15}$$

where the superscript "1" denotes the fictional world with linear price schedules and the superscript "0" denotes the nonlinear price schedule.

Without solving for the optimal price schedule, I need to make assumptions on which linear price the firm sets in the fictional world. Since it is not immediately clear, table 1.8 presents welfare calculations using the linear price per ounce of the listed size for all options given the parameter estimates presented in section 1.8. Since it is not immediately clear to me what prices a government agency may implement in this hypothetical world, I consider several different linear price sets, from an unconditional mean price per ounce based upon the size of the products to a mean price per ounce that is conditional on both organic status and brand type.

In each of the counterfactual analyses, the units for the welfare change in the first column are in dollars per shopping trip, which implies that under the various Gallon price based pricing regimes, each household is on average between \$.107 to \$.166 better off per shopping trip than under current market conditions according the counterfactuals which utilize household demographic information. This result is to be expected as firms use nonlinear pricing schedules that provide quantity discounts to those who purchase the gallon sized container. Contrastingly, the households in my sample are on average between \$0.057-\$0.0669 worse off per shopping trip under a half gallon/quart pricing regime, which is again to be expected as in order to provide a quantity discount to consumers who purchase gallon sized containers, firms must charge more per ounce for the smaller sized half gallon and quart sized units.

[58]Note: It must be profit increasing, as otherwise firms would not utilize the practice.

[59]Perhaps the government disallows nonlinear pricing in this world.

Of these sets of counterfactuals, I find the average price for all sizes condition an scenario to be most compelling. Under the average price for all sizes, I find that households are on average between \$0.0284-\$.0609 better off per shopping trip. Since the results of the counterfactuals are somewhat split, being that households are better off under gallon linear prices and worse off under half gallon/quart linear prices, I find the counterfactuals that utilize a mean price per ounce between the two to be most compelling. As such, I feel comfortable concluding that households are worse off due to the nonlinear price schedules that firms within the fluid milk industry utilize in the marketplace.

The last column of table 1.8 presents a back-of-the-envelop estimate of the total welfare changes as a sum of the total change of all households who live in my region[60] during a year. To calculate these changes, I assume that each household takes a number of trips to stores that could sell fluid milk that is equal to the average found in the Nielsen dataset over my sample period of 2008 and 2009. I then use the total number of households that live in the region, as provided by the Census Bureau, of 18,592,941 total households and multiply the total number of households by the numbers in the per trip column and by the average number of trips per household. I find that the total potential welfare benefits for all households under a gallon linear pricing regime is between approximately \$263 and \$399 million per year, whereas households are between \$148 million and \$178 million worse off a year in aggregate under a half gallon/quart designed linear pricing schedule.

In the average of all prices counterfactuals, I compute that consumers would be between approximately \$63 million to \$140 million better off per year in aggregate. Note that these estimates represent an upper bound of the possible yearly welfare change, as these calculations implicitly assume that all households living in this region participate in the market for fluid milk, whereas only households who do actually participate in this market would be effected by the nonlinear pricing in this market. These estimates may seem large, but according to the USDA, approximately 109,425 million pounds[61] of fluid milk were sold nation wide during the time period, thus these dollar amounts seem fairly reasonable as these estimates would represent a small fraction of the market in terms of sales revenue, yet a large welfare change to consumers who participate in the market.

Since I have looked into the effect of including household demographic characteristics on the demand parameter estimates in detail, the next logical step is to compare the results of my counterfactual analysis under both sets of assumptions. I also look into both specifications of the model, i.e. the "Standard Model" and "My Model", to quantify the consequences of ignoring the micro level consumption decisions that households make. Observing the differences in welfare levels per trip between the two specification, I show that the standard specification appears to systematically overstate the welfare loss to consumers if a linear price schedule based upon any of the current average price per ounce were introduced to the market place. Though overstatement is seemingly small in per trip terms, differences in yearly aggregates

[60] Illinois, Indiana, Michigan, Ohio and Wisconsin.

[61] Or 941,055 million gallons.

are measured in millions of dollars. One potential implication from this result is that if the household level choices/demographic characteristics are ignored when estimating demand parameters, then welfare counterfactuals may be understated in terms of their magnitudes. This result may also help explain why my estimated elasticities are somewhat larger in magnitude than what has been estimated previously within the literature.

I find the counterfactual using the average price per ounce of all product sizes to yield the most compelling results of the three linear price sets as it seems like the most likely of the three that would be used by firms within the market if they were forced to use linear price schedules tomorrow. As such, and for the other reasons presented above I find that the nonlinear price schedules used by the firms within the market are welfare increasing from the consumer's perspective.

However, this methodology is clearly limited, as firms have not been given a chance to respond and reoptimize in terms of setting their linear price. Additionally, firms within the fluid milk industry may find that it is no longer profitable to sell the smaller sized containers, which is ignored by this methodology. I find the price computations with the average price per ounce of all sizes to be most compelling of the three measures as it would represent a world that firms would seem mostly likely to choose if not given the chance to reoptimize. Thus, I find that nonlinear pricing is welfare I cannot be entirely certain of the effect of mandating a linear price schedule from the consumer's perspective until these ideas are taken into account.

With all that being said, I find these calculations to be consistent with a story that nonlinear price schedules in this market are welfare decreasing, however an approach where firms are allowed to reoptimize their prices and the menu of products would provide more compelling evidence.

## 1.7 Heterogeneous Welfare Implications of Nonlinear Pricing by Demographic Groupings in the Fluid Milk Market

There have been a great number of papers studying the welfare effects of price discrimination on consumers and welfare. However, to the best of my knowledge, there has not been an analysis that studies which particular types of consumers are harmed and which benefit from price discrimination. This question could be quite important for policymakers considering whether to step in and take action in markets where nonlinear pricing occurs. This is because regardless of the overall effect of price discrimination in a market in terms of consumer welfare, there are likely some winners and some losers due to the pricing strategy set by firms. Due to this, this paper investigates which consumers are worse off, and which are better off due to price discrimination in fluid milk market as examined in chapter 1 of this dissertation.

To answer this question, I use the demographic information provided by the Nielsen household panel dataset and explore how particular groups are affected. In particular, I am interested in how families close to the WIC threshold, but do not receive WIC benefits, are affected by the price discriminatory practices of firms in the fluid milk market. I do so by following an approach very similar chapter 1 of this dissertation, where I estimate and compare consumer surplus if firms were not allowed to price discriminate to the current world, then consider these outcomes by demographic grouping. The other demographic grouping margins I consider are: households with and without children, income, and education.

This question is particularly relevant due to the findings of the negative overall impact of nonlinear pricing on consumer welfare found previously in this chapter. However, even though the overall effect is negative in aggregate, it is still an empirical question as to how different groups of households fair in this market. Since households in this market receive a quantity discount when they purchase gallon sized containers, high-demand type households may gain from the nonlinear pricing since they would only purchase the gallon sized containers and will always receive a discount in terms of price per ounce. It is reasonable to believe that particular household characteristics, such as the presence of child, may drive the sorting into low- and high-demand type groupings, and if that were to be the case, household with children could benefit from the price discriminatory practices of firms in this market, even though the overall effect for all households still results in a reduction of consumer surplus.

The findings of nonlinear pricing leading to an overall reduction in consumer surplus found in the previous section of this chapter may suggest that policymakers should ban nonlinear pricing if taken at face value. However, this would prove to be unwise if the high-demand group consists of families with small children, as these households are often directed consume more fluid milk by pediatricians to help the children present in the household fend off malnourishment. Since this is the case, the banning of nonlinear pricing could lead to lower fluid milk consumption by children in non-WIC households. If these children are of lower income, where malnutrition is a concern, then such a banning would be unwise for policymakers to implement.

### 1.7.1 Literature

Choi et. al (2013) examines the effects of the introduction of organic milk to the fluid milk market and considers the implications by demographic grouping using the same Nielsen datasets that I use in this chapter as well as the first. Overall Choi et al. (2013) finds that the introduction of organic milk increases consumer welfare and it's findings suggest that this is true for each and every consumer demographic grouping. However, this paper finds that consumer welfare is more dramatically increased for wealthier, more educated households. Using the framework provided in Choi et. al (2013)'s counterfactual analysis will allow me to see which groups are not only most affected by the price discriminatory practices of firms in the fluid milk market, but also which are better and which are worse off due to the pricing strategy.

### 1.7.2 Results

Table 1.9 presents the results of my counterfactual analysis considering different demographic group margins. As stated in previously in this chapter, a positive result suggests that the counterfactual world is better than the current world, suggesting that nonlinear pricing leading to price discrimination lowers consumer welfare, whereas a negative result suggests that the opposite is true. The table is presented in dollars better off or worse off per trip and each counterfactual is computed using a linear price schedule constructed by the average per ounce price conditional on brand type and organic status for each size chosen compared to the current world's pricing structure that utilizes price discrimination. Table 1.9 is separated by different demographic margins and the values that each takes.

The results for each counterfactual are quite reasonable in terms of size and magnitude. The household income margin suggests that households below the US median have a greater reduction of consumer surplus than those above the US median. This suggests that on average, households below the US median are more likely to purchase smaller sizes and become victim of the pricing strategy by firms. Since lower income households have tighter budgets, they may purchase the smaller sized containers because they are cheaper than the gallon sized containers, even though by doing so firms charge them a higher price per ounce for the product. However since both are positive and similar in terms of magnitude, this does not appear to be a meaningful different in terms of economic significance.

Education also does not appear to play a major role in who this pricing strategy affects, however those with more education are on average less adversely affected than those without any college education. This result is consistent with the idea that those with more education are likely to observe the difference in the per ounce prices across the pricing menu those without. These results are very closely related to income, which is expected since education and income are strongly correlated.

The counterfactuals relating to the presence of children are signed as expected: households with children benefit from nonlinear pricing leading to price discrimination, whereas households without children are worse off. This is consistent with the idea that households with children have a higher demand for fluid milk than those without and thus are far more likely to purchase the gallon sized containers and receive their associated quantity discount.

The counterfactuals regarding whether or not the household heads are married both suggest that regardless of the martial status of the heads-of-household, nonlinear pricing due to price discrimination reduces surplus for households. However, I find that married heads of household a less adversely affected than unmarried ones. Marital status here serves somewhat as a proxy for single headed households, as the majority of households where the household head is not married are single headed.

Finally, the welfare counterfactual regarding households that are low income and have children is relatively large in terms of magnitude and negative. This suggests that low income families with children significantly benefit from nonlinear pricing due to price discrimination. Due to this finding, it would seem that banning nonlinear pricing may not behoove policymakers as this could be directly to the detriment of low income household with children and may reduce fluid milk consumption in households where malnutrition seems to be a concern.

## 1.8 Conclusion

This paper finds evidence of second-degree price discrimination in the oligopoly market of fluid milk. Using consumer level data provided by the Nielsen Company, I find that firms are only able to price discriminate on their own private label products, of which 7.2% of the markup can be explained by price discrimination, but not the regional brand.

In addition, through counterfactual analysis, I find that price discrimination is welfare decreasing from the consumer's perspective and if firms were forced into linear price schedules, households would gain anywhere between $63 million to $140 per year in aggregate in consumer surplus in my sample region yearly. Additionally, by estimating a model with and without household demographic characteristics and comparing the models, I see that ignoring consumer decisions as well as the demographic make up of the market place being study leads to biased parameter estimates. Households who purchase the smaller sizes, quart and half gallon containers, yield some amount of utility from more convenient sized packages, whereas consumers who purchase gallon sized containers enjoy lower prices as a result of higher competition in this segment of the market. I also find that though overall consumers are worse off due to the pricing strategy, those in vulnerable groups benefit from it, suggesting that policymakers should not act to stop firms from utilizing this strategy. These findings are in line with what the previous literature has found and are possibly more compelling than what previous papers have done given the richness of the data that I utilize as provided by the Nielsen company.

## 1.9 Tables and Figures

Table 1.1: Unit Price Summary Statistics by Size and Coupon

| | No Coupon | | | Coupon | | |
|---|---|---|---|---|---|---|
| | Quart | Half Gallon | Gallon | Quart | Half Gallon | Gallon |
| Mean Price | 1.387 | 1.704 | 2.375 | 0.896 | 1.056 | 1.502 |
| St Dev | 0.354 | 0.688 | 0.509 | 0.444 | 0.775 | 0.784 |
| Observations | 417,805 | 161,650 | 16,439 | 22,235 | 8,780 | 492 |

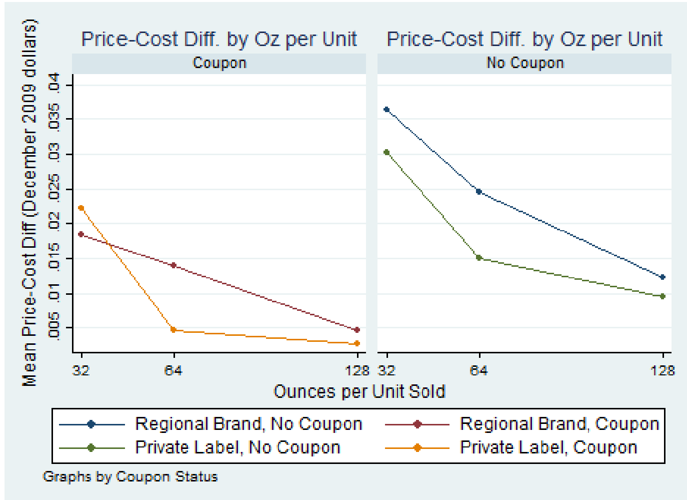
Note: An observation is a UPC purchased by a panelist in the years 2008-2009. Prices are in December 2009 dollars. Unit price is defined to be the final price paid by the panelist, minus any coupons, for one unit (which in most cases is a single container) of the UPC in question.

Table 1.2: Price Statistics by Size and Coupon

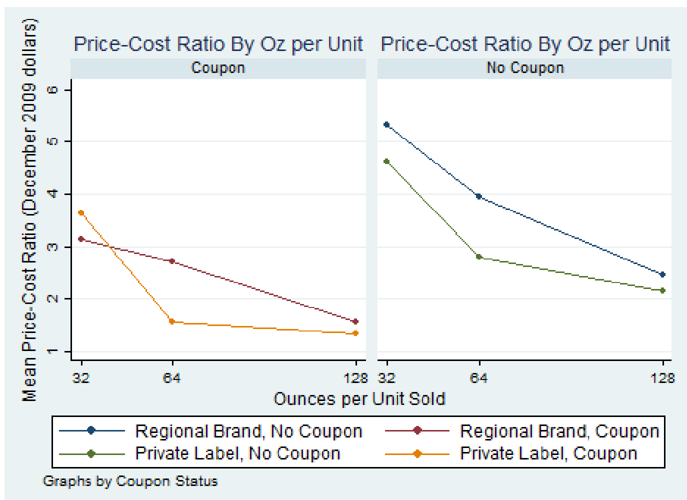
| | No Coupon | | Coupon | |
|---|---|---|---|---|
| | Mean | St Dev | Mean | St Dev |
| ***Quart*** | | | | |
| Raw Milk Price/oz | 0.009 | 0.001 | 0.009 | 0.001 |
| Price/oz | 0.043 | 0.011 | 0.028 | 0.014 |
| Price-Cost Ratio | 5.155 | 1.530 | 3.240 | 1.685 |
| Price-Cost Differential | 0.03480 | 0.01123 | 0.01927 | 0.01392 |
| ***Half Gallon*** | | | | |
| Raw Milk Price/oz | 0.009 | 0.001 | 0.008 | 0.001 |
| Price/oz | 0.027 | 0.011 | 0.017 | 0.012 |
| Price-Cost Ratio | 3.175 | 1.378 | 1.991 | 1.518 |
| Price-Cost Differential | 0.01813 | 0.01082 | 0.00808 | 0.01217 |
| ***Gallon*** | | | | |
| Raw Milk Price/oz | 0.008 | 0.001 | 0.008 | 0.001 |
| Price/oz | 0.019 | 0.004 | 0.012 | 0.006 |
| Price-Cost Ratio | 2.206 | 0.531 | 1.401 | 0.747 |
| Price-Cost Differential | 0.01006 | 0.00403 | 0.00331 | 0.00609 |
| ***Observations*** | | | | |
| Quart | 16,439 | | 492 | |
| Half Gallon | 161,650 | | 8,780 | |
| Gallon | 417,805 | | 22,235 | |

Note: An observation is a UPC purchased by a panelist in the years 2008-2009. Prices are in December 2009 dollars. Raw Milk Price/oz is the raw milk price in ounces at the time of purchase. Price/oz is the unit price per ounce at the time of purchase. Price-Cost Ratio is computed by dividing the unit price per ounce by the raw milk price per ounce. Price-Cost Differential is computed by subtracting the raw milk price per ounce from the unit price per ounce.

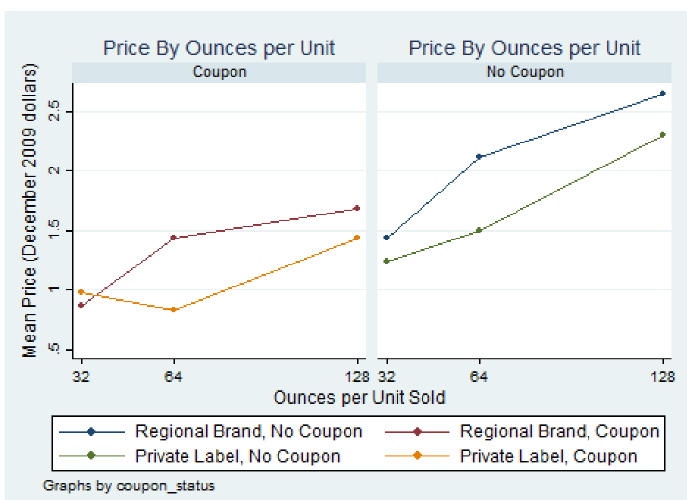
Figure 1.1: Graphical Evidence of Price Discrimination




(a) Price-Cost Differential




(b) Price-Cost Ratio




(c) Price by Ounces per Unit

Table 1.3: Reduced Form Regression Results: Size Dummies

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Gallon Sized Container | $-0.024677^{***}$ | $-0.021815^{***}$ | $-0.021376^{***}$ | $-0.020873^{***}$ |
| | (0.00039) | (0.00058) | (0.00019) | (0.00019) |
| Half Gallon Sized Container | $-0.016787^{***}$ | $-0.015495^{***}$ | $-0.015701^{***}$ | $-0.015703^{***}$ |
| | (0.00040) | (0.00029) | (0.00026) | (0.00025) |
| Raw Milk cost/oz | $0.057766^{**}$ | 0.020579 | $0.053478^{***}$ | $0.036165^{**}$ |
| | (0.02945) | (0.02467) | (0.01746) | (0.01811) |
| Private Label Dummy | | $-0.004687^{***}$ | | $-0.002624$ |
| | | (0.00088) | | (0.00178) |
| Organic Dummy | | $0.021756^{***}$ | | $0.020875^{***}$ |
| | | (0.00050) | | (0.00033) |
| Flavor Dummy | | 0.000491 | | 0.000475 |
| | | (0.00033) | | (0.00037) |
| Coupon Value | | $-0.005929^{***}$ | | $-0.006116^{***}$ |
| | | (0.00012) | | (0.00013) |
| Controls | N | Y | N | Y |
| Brand Fixed Effects | N | N | Y | Y |
| Obs. | 627,401 | 627,401 | 627,401 | 627,401 |
| Brands | . | . | 195 | 195 |

Note: There are 627,401 sales of fluid milk from a cow sized using this paper's definition of the Imperial measurement system. Controls include channel type indicator variables as well as state fixed effects. Standard errors are clustered at the store level for all regressions. Standard errors are shown in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 1.4: Demand Parameter Estimates

| | Standard Model | | My Model | |
|---|---|---|---|---|
| Parameter | Estimate | Std Err | Estimate | Std Err |
| Gallon Sized Container | 1.903 | 0.0188 | 2.617 | 0.087 |
| Butterfat Content $\geq 2\%$ | $-0.009$ | 0.0145 | 0.342 | 0.084 |
| Organic | $-0.631$ | 0.0660 | $-1.262$ | 0.531 |
| $\alpha$ (Price) | $-1.960$ | 0.0094 | $-1.961$ | 0.010 |
| UFE: (Private Label) | 0.413 | 0.0157 | 0.424 | 0.016 |
| UFE: (Discount or Hypermarket) | $-0.775$ | 0.0326 | $-0.795$ | 0.033 |
| UFE: (Warehouse Club) | $-1.764$ | 0.1019 | $-1.746$ | 0.102 |
| UFE: (Drug Store) | $-1.605$ | 0.0684 | $-1.605$ | 0.069 |
| UFE: (Gas Station/Convenience) | $-2.493$ | 0.1112 | $-2.493$ | 0.108 |
| UFE: (Other) | $-0.882$ | 0.1281 | $-0.936$ | 0.131 |
| GAL Chars: Days Since Last Milk | — | — | $-0.019$ | 0.000 |
| GAL Chars: HH Inc > U.S. Med | — | — | $-0.048$ | 0.026 |
| GAL Chars: At least 1 College | — | — | 0.113 | 0.032 |
| GAL Chars: Avg Age of HH Heads | — | — | $-0.012$ | 0.001 |
| GAL Chars: Kids Dum | — | — | 0.219 | 0.033 |
| GAL Chars: Married Dum | — | — | 0.388 | 0.030 |
| BFC Chars: Days Since Last Milk | — | — | $-0.011$ | 0.000 |
| BFC Chars: HH Inc > U.S. Med | — | — | $-0.083$ | 0.028 |
| BFC Chars: At least 1 College | — | — | $-0.195$ | 0.031 |
| BFC Chars: Avg Age of HH Heads | — | — | 0.004 | 0.001 |
| BFC Chars: Kids Dum | — | — | 0.032 | 0.035 |
| BFC Chars: Married Dum | — | — | $-0.189$ | 0.031 |
| ORG Chars: Days Since Last Milk | — | — | $-0.083$ | 0.022 |
| ORG Chars: HH Inc > U.S. Med | — | — | 0.330 | 0.141 |
| ORG Chars: At least 1 College | — | — | 1.461 | 0.296 |
| ORG Chars: Avg Age of HH Heads | — | — | 0.025 | 0.005 |
| ORG Chars: Kids Dum | — | — | 0.047 | 0.206 |
| ORG Chars: Married Dum | — | — | $-2.193$ | 0.147 |
| Observations | 66,767 | | 66,767 | |

Note: Estimates obtained via 5% random sample of households. Standard errors bootstrapped with 500 repetitions.

Table 1.5: Demand: Marginal Effects

| Parameter | Standard Model | My Model |
|---|---|---|
| Gallon Sized Container | 0.97107 | 1.33417 |
| Butterfat Content $\geq$ 2% | $-0.00440$ | 0.17433 |
| Organic | $-0.32192$ | $-0.64346$ |
| UFE: (Private Label) | 0.21079 | 0.21597 |
| UFE: (Discount or Hypermarket) | $-0.39536$ | $-0.40538$ |
| UFE: (Warehouse Club) | $-0.89998$ | $-0.89002$ |
| UFE: (Drug Store) | $-0.81864$ | $-0.81851$ |
| UFE: (Gas Station/Convenience) | $-1.27191$ | $-1.27109$ |
| UFE: (Other) | $-0.45013$ | $-0.47723$ |
| GAL Chars: Days Since Last Milk | — | $-0.00975$ |
| GAL Chars: HH Inc > U.S. Med | — | $-0.02432$ |
| GAL Chars: At least 1 College | — | 0.05736 |
| GAL Chars: Avg Age of HH Heads | — | $-0.00605$ |
| GAL Chars: Kids Dum | — | 0.11149 |
| GAL Chars: Married Dum | — | 0.19783 |
| BFC Chars: Days Since Last Milk | — | $-0.00557$ |
| BFC Chars: HH Inc > U.S. Med | — | $-0.04225$ |
| BFC Chars: At least 1 College | — | $-0.09948$ |
| BFC Chars: Avg Age of HH Heads | — | 0.00207 |
| BFC Chars: Kids Dum | — | 0.01649 |
| BFC Chars: Married Dum | — | $-0.09650$ |
| ORG Chars: Days Since Last Milk | — | $-0.04214$ |
| ORG Chars: HH Inc > U.S. Med | — | 0.16839 |
| ORG Chars: At least 1 College | — | 0.74487 |
| ORG Chars: Avg Age of HH Heads | — | 0.01281 |
| ORG Chars: Kids Dum | — | 0.02393 |
| ORG Chars: Married Dum | — | $-1.11824$ |

Note: Marginal effects obtained by dividing each parameter estimate by the absolute value of the price parameter $\alpha$.

Table 1.6: Estimated Elasticities

| Choice | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | C13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Choice Characteristic Definitions** | | | | | | | | | | | | |
| Brand Type | RB | RB | RB | RB | PL | PL | RB | RB | RB | RB | PL | PL |
| Size | SM | SM | GA | GA | SM | GA | SM | SM | GA | GA | SM | GA |
| BFC | L | L | L | L | L | L | W | W | W | W | W | W |
| Org Status | N | O | N | O | N | N | N | O | N | O | N | N |
| **Elasticity Matrix** | | | | | | | | | | | | |
| | −3.370 | 0.003 | 0.072 | 0.006 | 0.101 | 0.187 | 0.055 | 0.004 | 0.083 | 0.001 | 0.103 | 0.180 |
| | 0.086 | −5.011 | 0.116 | 0.007 | 0.139 | 0.265 | 0.094 | 0.005 | 0.124 | 0.001 | 0.139 | 0.252 |
| | 0.073 | 0.004 | −4.431 | 0.007 | 0.130 | 0.241 | 0.081 | 0.005 | 0.104 | 0.001 | 0.132 | 0.232 |
| | 0.122 | 0.006 | 0.165 | −6.965 | 0.192 | 0.366 | 0.135 | 0.008 | 0.174 | 0.002 | 0.192 | 0.352 |
| | 0.047 | 0.002 | 0.064 | 0.004 | −2.706 | 0.146 | 0.052 | 0.003 | 0.068 | 0.001 | 0.074 | 0.140 |
| | 0.064 | 0.003 | 0.085 | 0.006 | 0.105 | −3.541 | 0.071 | 0.004 | 0.093 | 0.001 | 0.106 | 0.178 |
| | 0.047 | 0.003 | 0.070 | 0.006 | 0.097 | 0.179 | −3.235 | 0.004 | 0.080 | 0.001 | 0.099 | 0.173 |
| | 0.087 | 0.004 | 0.117 | 0.007 | 0.140 | 0.266 | 0.096 | −5.042 | 0.125 | 0.001 | 0.140 | 0.254 |
| | 0.074 | 0.004 | 0.096 | 0.007 | 0.129 | 0.240 | 0.083 | 0.005 | −4.417 | 0.001 | 0.131 | 0.231 |
| | 0.134 | 0.007 | 0.182 | 0.011 | 0.219 | 0.412 | 0.148 | 0.008 | 0.196 | −7.809 | 0.221 | 0.395 |
| | 0.048 | 0.002 | 0.064 | 0.004 | 0.072 | 0.144 | 0.052 | 0.003 | 0.068 | 0.001 | −2.695 | 0.138 |
| | 0.066 | 0.004 | 0.088 | 0.006 | 0.109 | 0.192 | 0.073 | 0.004 | 0.096 | 0.001 | 0.110 | −3.675 |

Table 1.7: Firm Side Model Parameters Estimates

**Marginal Cost Regression Estimates**

| Parameter | Estimate | Std Err |
|---|---|---|
| **Product Characteristics** | | |
| Gallon Sized Container | 1.10833 | 0.00172 |
| Butterfat Content $\geq 2\%$ | 0.21635 | 0.00172 |
| Organic | 1.67761 | 0.00209 |
| Private Label | $-0.27917$ | 0.00209 |
| **Input Costs** | | |
| Raw Milk Cost | 0.09164 | 0.00012 |
| Competition[62] | 0.01739 | 0.00010 |
| **Channel Type**[63] | | |
| Discount or Hypermarket | 0.08566 | 0.00242 |
| Warehouse Store | $-0.04678$ | 0.00485 |
| Drug Store | $-0.13852$ | 0.00323 |
| Gas Station/Convenience | 0.11077 | 0.00371 |
| Other | $-0.02043$ | 0.00749 |
| Total Number of Firm-Weeks | 51,547 | |

[58] Proxy for transportation costs.

[59] Grocery is the omitted channel.

Table 1.8: Welfare Counterfactuals

**Unconditional Mean Price/oz by Size**

| | My Model | | Standard Model | |
|---|---|---|---|---|
| Linear PS | \$/Trip | \$/yr Aggregate | \$/Trip | \$/yr Aggregate |
| Gallon | 0.1589 | \$399MM/yr | 0.1640 | \$412MM/yr |
| Average | 0.0555 | \$140MM/yr | 0.0583 | \$147MM/yr |
| Smaller Sizes | −0.0687 | −\$173MM/yr | −0.0669 | −\$168MM/yr |

**Mean Price/oz by Size Conditional on Organic Status**

| | My Model | | Standard Model | |
|---|---|---|---|---|
| Linear PS | \$/Trip | \$/yr Aggregate | \$/Trip | \$/yr Aggregate |
| Gallon | 0.1225 | \$308MM/yr | 0.1236 | \$311MM/yr |
| Average | 0.0371 | \$93MM/yr1 | 0.0385 | \$97MM/yr |
| Smaller Sizes | −0.0710 | −\$178MM/yr | −0.0687 | −\$173MM/yr |

**Mean Price/oz by Size Conditional on Brand Type**

| | My Model | | Standard Model | |
|---|---|---|---|---|
| Linear PS | \$/Trip | \$/yr Aggregate | \$/Trip | \$/yr Aggregate |
| Gallon | 0.1296 | \$326MM/yr | 0.1326 | \$333MM/yr |
| Average | 0.0306 | \$77MM/yr | 0.0313 | \$79MM/yr |
| Smaller Sizes | −0.0578 | −\$145MM/yr | −0.0565 | −\$142MM/yr |

**Mean Price/oz by Size Conditional on Brand and Organic Status**

| | My Model | | Standard Model | |
|---|---|---|---|---|
| Linear PS | \$/Trip | \$/yr Aggregate | \$/Trip | \$/yr Aggregate |
| Gallon | 0.1046 | \$263MM/yr | 0.1055 | \$265MM/yr |
| Average | 0.0249 | \$63MM/yr | 0.0258 | \$65MM/yr |
| Smaller Sizes | −0.0588 | −\$148MM/yr | −0.0572 | −\$144MM/yr |

Note: In computing yearly aggregate I assume that each household takes the average number of trips per year found in my data, 135.12 trips/year, and based upon census data, there were a total of 18,592,941 households in Nielsen's Midwest Region in the years 2008 & 2009.

Table 1.9: Welfare Counterfactuals by Demographic Grouping

**Household Income**

| Margin | \$/Trip |
|---|---|
| Below US Median | 0.0306 |
| Above US Median | 0.0239 |

**Education**

| Margin | \$/Trip |
|---|---|
| At Least 1 Year of College | 0.0236 |
| No College | 0.0305 |

**Children**

| Margin | \$/Trip |
|---|---|
| Household with Children | −0.0099 |
| Household Without Children | 0.0307 |

**Married**

| Margin | \$/Trip |
|---|---|
| Married Heads of Household | 0.0184 |
| Unmarried Heads of Household | 0.0396 |

**Low Income with Children**

| Margin | \$/Trip |
|---|---|
| Low Income with Children | −0.0238 |

Note: All counterfactuals use a linear price schedule of mean price/oz by size that is conditional on brand, and organic status and are in terms of dollars per trip.

Table 1.10: Choice Set Simulation: Percent of Observations with Missing Values per Step

| Level of Aggregation | % of Obs with Missing Choices |
|---|---|
| Store-Week | 0.7796 |
| Store-Month | 0.7653 |
| Store-All | 0.7259 |
| Retailer-Week | 0.7012 |
| Retailer-Month | 0.6756 |
| Retailer-All | 0.4425 |
| Channel-Week | 0.2043 |
| Channel-Month | 0.1952 |
| Channel-All | 0.1662 |
| Mean Choice Set | 0.0000 |
| Observations | 66,889 |

Note: This table outlines how the simulated choice set was constructed. As stated in in section 1.4.2, only the choice that the household makes on a shopping trip is observable as data, but the entire menu of choices is necessary to estimate demand. As such, I construct the household's choice set by starting with the characteristics of the choice they chose on the trip, then use the various aggregates of data listed in the table to update information about the choice set the household faces using data from purchases made by other households on other products as well as through the Nielsen Scanner data during that week.

Table 1.11: Panelist Years

| Year | Freq. | Percent |
|---|---|---|
| 2008 | 7,754 | 50.84 |
| 2009 | 7,498 | 49.16 |

Note: An observation is an active panelist during a panel year period. There is a total of 15,252 active panel years between 2008 and 2009.

Table 1.12: Household Income Measures

| Income Measure | Mean | St Dev |
|---|---|---|
| HH Income > US Median | 0.54 | 0.50 |
| HH Income < 150% Poverty Line | 0.09 | 0.29 |
| Observations | 15,252 | |

Note: An observation is a household year. The first measure of income, labeled HH Income > US Median" is a dummy variable that equals 1 if the household's income is above the US median during a particular panel year and zero otherwise. The second measure of income, labeled HH Income < 150% Poverty Line is a dummy variable that equals 1 if the household's total income is below 150% of the Federal poverty line.

Table 1.13: Household Size

| Size of Household | Freq. | Percent |
|---|---|---|
| 1 | 3,077 | 30.58 |
| 2 | 6,986 | 69.42 |
| 3 | 2,276 | 22.62 |
| 4 | 1,847 | 18.35 |
| 5 | 720 | 7.15 |
| 6 | 235 | 2.34 |
| 7 | 74 | 0.74 |
| 8 | 24 | 0.24 |
| 9 | 13 | 0.13 |

Note: An observation is an active household year.

Table 1.14: Household State of Residence

| State | Freq. | Percent |
|---|---|---|
| Illinois | 3,578 | 23.46 |
| Indiana | 2,023 | 13.26 |
| Michigan | 3,068 | 20.12 |
| Ohio | 4,514 | 29.60 |
| Wisconsin | 2,069 | 13.57 |

Note: An observation is an active household year.

Table 1.15: Marital Status of Household

| | Freq. | Percent |
|---|---|---|
| Married | 14,102 | 63.65 |
| Widowed | 1,697 | 7.66 |
| Divorced/Separated | 3,215 | 14.51 |
| Single | 3,140 | 14.17 |

Note: An observation is a panelist year.

Table 1.16: Age and Presence of Children

| | Freq. | Percent |
|---|---|---|
| Under 6 only | 780 | 3.52 |
| 6-12 only | 1,327 | 5.99 |
| 13-17 only | 1,638 | 7.39 |
| Under 6 & 6-12 | 727 | 3.28 |
| Under 6 & 13-17 | 123 | 0.56 |
| 6-12 & 13-17 | 899 | 4.06 |
| Under 6 & 6-12 & 13-17 | 148 | 0.67 |
| Under 6 & 6-12 & 13-17 | 16,512 | 74.53 |

Note: An observation is a panelist year.

Table 1.17: Age of Household Head

| Household Head Age | Obs | Mean | Median | St Dev |
|---|---|---|---|---|
| Male Head Age | 12,164 | 44.54 | 51.28 | 25.24 |
| Female Head Age | 14,071 | 50.88 | 53.29 | 19.33 |
| Observations | 15,252 | | | |

Note: An observation is an active household year.

## Chapter 2 Firm Demand Learning in Washington State's New Liquor Market.[1]

### 2.1 Introduction

Demand learning has been empirically researched extensively, but almost exclusively from the point of view of the consumers. There is little current empirical research on how firms learn about demand for their product. This chapter is one of the first to show empirical evidence for learning behavior by firms.

It is important to understand how firms learn about the demand for their products since without this understanding we may observe behaviors from firms that appear irrational. Many of these behaviors that may appear to be non-profit maximizing may be explained by firms learning about the demand for their products. For example, consider a firm that competes in a market with their competitors a la Bertrand. We may observe this firm set a price that is not the Nash Equilibrium price when competing in prices with their competitors in order to learn, either about the demand function of their consumers or how their competitors compete. Though there are many reasons why firms may not play Nash-Bertrand strategies, demand learning is one of the most compelling and is the main focus of this chapter of the dissertation. This chapter seeks to observe and identify demand learning behavior exhibited by firms, as prices and market conditions contain signals of information. It may be fine to ignore these signals when discussing a market that is clearly in long-run equilibrium, but for markets that are not, it is clearly inappropriate.

To my knowledge of the literature, there are four papers that discuss and identify demand learning by firms. These papers are: Hitsch (2006), Doraszelski, Lewis and Pakes (2016) [DLP16], Escobari (2012) and Huang et. al (2018). Hitsch (2006) uses the ready-to-eat (RTE) cereal market to show how firms learn about entry and exit decisions in the market. [DLP16] looks at how firms in the UK electricity market learn about the demand for electricity as well as how they compete against one another. Escobari (2012) uses airlines data to see how firms learn about demand for their product in a perishable goods market. McGoldrick and Mark (1985), Hitsch (2006), [DLP16], Escobari (2012) and Huang et. al (2018) is discussed at length in a later section of this chapter.

This chapter uses the Washington state spirits market after Initiative 1183 was enacted in June 2012. This initiative privatized spirits sales in Washington state. Previous to this change in the law, spirits were exclusively sold by state owned and licensed liquor stores. At these stores, the state government limited the options that the stores were allowed to carry and the prices were set at the state capital. After

[1] *Calculated (or Derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.*

the law change, stores are allowed to set any price above the state minimum. In a somewhat related paper, Seo (2019) uses Initiative 1183 to estimate the value of firm scope and one-stop shopping. The details of Initiative 1183 and Seo (2019) are discussed in greater detail later in this chapter as well.

The Washington state spirits market is particularly appealing for this type of analysis for several reasons, most important of which are the implications of Initiative 1183. Since Initiative 1183 allows grocery stores to start selling liquor starting in June 2012, the landscape of the market for liquor drastically changed. With these grocery stores entering the market and several liquor stores[2] exiting the market, demand conditions for spirits within the state were uncertain. The uncertainty of demand conditions created by Initiative 1183 provides incentive for these firms to learn about the true demand parameters within the spirits market in order to maximize profits.

Another appealing feature of this market is the heterogeneity of population densities within the state of Washington. There are some areas of the state, such as the Seattle, Spokane and Tacoma areas, that are very densely populated. Conversely, other regions of the state are sparsely populated, such as the areas surrounding the mountains in the middle of the state and the eastern portion of the state. In the densely populated areas of the state, there are very competitive markets for liquor sales, whereas in the less densely populated areas of the state there are many instances of grocery stores operating as a local monopoly for spirits sales. For evidence of this distribution of the population density, as well as competitiveness of these regional markets, see the maps and tables in section 2.10.

The heterogeneity in population densities of the state is relevant to this paper due to the priors of the demand parameters that these firms that are entering the market may have. Since many of the firms that are entering this market are a part of a national grocery chain, they might have priors on the parameters of the demand function for spirits that are very close to the true demand parameters in regions of the state that are similar in composition to other markets they currently operate in. Thus, there is a reasonable probability that little learning will occur in the densely populated areas of the state. However, there is no reason to believe that firms who operate in areas with low population densities will have priors that are close to the true demand parameters, as the demand parameters in these regions depend greatly on local tastes.

One way that I observe firms learning about the demand for spirits is through changes in prices. This type of learning behavior will be observed as firms changing prices of their products over time in response to consumer behavior. For example, suppose a grocery store on June 1st, 2012 sets the price for a fifth of Johnnie Walker Red Label at $20. If the firm is incorrect in their prior for consumer tastes for this type of spirit, then the firm may have trouble keeping enough product to fill the shelf in the case that their guess of $20 too low, or they may have trouble moving any product at all, in the case that their guess of $20 too high. If either case is true, the grocery store is likely to change their price in the appropriate direction. The reason why learning

[2]According to Seo (2019), as of January 2012 there were 330 state owned/licensed liquor stores, by 2014 there were only 273 total liquor stores in the state.

could be observed in this way is due to the way that spirits are sized. Due to market norms, spirits are almost exclusively sold in containers of the following sizes: shots (44 ml), half pints (200 ml), pints (375 ml), fifths (750 ml), forties (1140 ml) and half gallons (or handles 1750 ml). This allows me to focus only on firms adjusting prices for the products rather than changing the entire price-quantity bundle they offer.

Another way I observe firms learning about the demand for spirits is through changes in the menu of options that they sell. Due to heterogeneity in consumer tastes, a menu of different types and sizes of spirits may be appropriate to maximize profit in one region of the state, and inappropriate in another. To visualize this type of learning, consider the follow simple example. Suppose a grocery store opens a spirits section on June 1, 2012 that includes expensive specialty spirits. If sales do not go over well, the grocery store may decide to price the remaining products at the state minimum and discontinue the section. Additionally, the opposite may occur in that I may observe firms open their sales with a small variety of options, and then expand their menu at a later date.

The main contribution of this paper to the current economics literature is that it identifies firm demand learning in a new final goods consumer market. Of the existing literature on firm learning, Hitsch (2006) identifies learning through entry and exit decisions in the RTE cereal market. The first RTE cereal was introduced to the US market in 1863[3], thus identifying firm learning in the RTE cereal market is systematically different from the brand new market created by Initiative 1183 as generally speaking, the parameters for demand of RTE cereal are known by many firms. [DLP16] examines the electricity market in the UK after deregulation. Since the electricity is being sold through bidding on contracts with the government, learning in [DLP16]'s market setting focuses more on how each firm's competitors compete rather than learning about the true demand parameters.[4] This paper differs from Escobari (2012) in that he focuses on how firms learn about the demand for their product as the expiration date of their good approaches. Spirits have no relevant expiration date, thus his analysis is unlike this one.

The most closely related paper to this one is Huang et. al (2018). Huang et. al (2018) examines the 2012 Washington State liquor law change and estimates a structural model of demand learning by firms focusing entirely on prices. This paper deviates from that approach by also considering the learning effect on the menu of options that firms can offer and by taking a reduced-form, survey-like view of the market as a whole.

The rest of this chapter is organized as follows: section 2.2 discusses the economic literature related to this chapter, section 2.3 provides a description of the market and Initiative 1183, section 2.4 provides a simple framework for firm demand learning, section 2.5 discusses the data I use, section 2.6 considers potential margins for differential learning, section 2.7 provides preliminary evidence of demand learning in this market, section 2.8 estimates reduced-form regressions that are consistent with

[3]Granula by James Caleb Jackson, operator of Our Home on the Hillside

[4]Though their model also identifies firms learning about the governments demand for electricity as well.

demand learning behavior and section 2.9 concludes.

## 2.2 Current Literature on Demand Learning

Little has been written so far on demand learning by firms in the economics literature. This is somewhat puzzling as there has been extensive discussion of consumer demand learning within the literature. The only logical explanation for this could be a lack of firm price/sales data necessary for identification of firm demand learning. This section of the chapter is organized as follows: the first subsection will discuss papers most related to this one and the second subsection will provide a brief overview of the learning literature as a whole.

### 2.2.1 Most Related Economics Literature

This subsection will briefly discuss the papers most related to this one. The papers discussed are from several disciplines as well as both empirical and theoretical.

Huang et. al (2018) considers the Washington spirits law change and through structural estimation techniques, finds that firms have slowed their learning by as early as 2016, only four years after the law change took place.

Hitsch (2006) focuses on the decision of a firm to introduce a new product to an existing market when demand conditions for the product are uncertain. His paper focuses primarily on the decision by firms to remove their new product from the market after observing sales data. His model identifies the optimal demand for information on demand parameters as well as predicts how firms will behave with new products based upon their uncertainty of the demand parameters. This paper will adapt Hitsch (2006)'s model to the Washington state spirits market.

[DLP16] claims to be the first structural paper that identifies firm demand learning. They use the newly deregulated UK electricity market as a case study for how firms learn both about how to compete in a market by bidding on government contracts as well as the government's demand for their product. They find that their model is better for predicting bids than the Bertrand Nash equilibrium predictions during the early stages of competition, suggesting that firms learn about demand in the early stages of deregulation. In later stages of the market, DLP find that the Bertrand Nash equilibrium predictions are better than their model, suggesting that the market eventually reaches an equilibrium.

Another empirical paper from the economics literature that discusses firms learning about the demand for their product is Escobari (2012). Escobari (2012) uses a U.S. airlines panel dataset to observe how firms dynamically price their inventories when demand is uncertain. Escobari uses a reduced-form hedonic pricing model of a ticket as well as a learning model to identify learning behavior by firms. Escobari finds that price increases as the number of seats on the flight decreases as well as price decreases as there is less time to sell, which is consistent with firm demand learning behavior.

Firm demand learning has drawn interest in the operations management literature as well as the industrial organization literature. Sen and Zhang (2008) looks to see when it is most beneficial for firms to learn about demand for their product. Using style goods data[5] as well as computational examples, they find that when initial inventory for these firms is high, learning models are more beneficial than models with no learning in terms of profits from the firm's perspective.

In a related paper from the marketing literature, McGoldrick and Mark (1985) look for evidence of demand learning using grocery market data. They find using different price quantity bundles and looking at unit pricing allows firms to learn about demand for their products.

Though there has been little on firm demand learning in the empirical economics literature, much has been written on the topic within the theoretical literature. This literature dates back to Cournot (1838)'s discussion of the theory of learning in normal-form games. Brown (1951) was the first to discuss belief-based learning (Fictitious Play seen in [DLP16]'s model), where agents note the history of the game and choose actions based upon beliefs formed from past actions of competitors.

More recently, Balvers and Cosimano (1990) discuss the process in which firms adjust prices to learn about demand for their goods. Chatwin (2000) also discusses the optimal dynamic price of perishable goods, something that will be important if looking into a perishable goods market. Dolan and Jeuland (1981) present a methodology for determining the optimal pricing strategy when supply and demand conditions are unstable.

The final paper discussed in this section is Seo (2019). Seo (2019) uses Initiative 1183 to identify the value of firm scope and one-stop shopping. She finds a non-Pareto improving welfare increase for consumers, firms and the government due to this policy. She does so by using stores that previously had liquor only additions before the law change and after the law change do not. Even in the case of only a wall separating the consumers previously, Seo (2019) finds a significant increase in welfare for consumers without the wall. Seo (2019) also uses the same data as this paper as well, the Nielsen Panel and Scanner datasets.

### 2.2.2 Current Literature on Consumer Demand Learning

As stated previously, much of the learning literature has focused on demand learning by consumers. Many of the papers from this literature focus on Bayesian Learning models. In these learning models, consumers are assumed to have some prior belief of their preference for the good, which they update after receiving some sort of signal about the product, either through purchasing and trying the product, through advertising, or both. The standard Bayesian learning model makes several assumptions about how agents learn about their demand parameters and several papers have relaxed these assumptions.

One strand of this literature generally discusses the consumption of some experience good, of which value is unknown to the consumer before purchase [Ackerberg (2003) {Grocery and Advertising}]. Many of the empirical papers of this flavor also consider markets where consumers have some uncertainty of how much they demand

[5] Apparel retailing

the good [Narayanan et al. (2007) {Telephones}, Miravete and Palacios-Huerta (2013) {Cell Phones}]

In a somewhat distantly related literature, health economists have shown evidence for demand learning by doctors for prescription drugs [Ching (2010) {Generic Drugs}, Coscelli and Shum (2003) {New Drug Entry}, Crawford and Shum (2005) {Anti-ulcer drugs}].

Ching, Erdem and Keane (2013) provides a general overview of the history and papers that are a part of this literature.

This literature is somewhat related to this paper in that it shows identification of consumer learning with Bayesian learning models, though the focus on the consumer side of the market somewhat distances these papers from this paper.

### 2.3 Washington State Spirits Market Description

In the period before June 2012, only beer and wine were sold in grocery stores in Washington. During this period, for an individual to purchase spirits, they must go to a state owned or licensed liquor store. Washington Liquor Licensing, Initiative 1183 was approved in November 2011 and enacted in June 2012. This initiative privatized liquor sales within the state, allowing grocery stores to now sell liquor if they obtained the required liquor license from the state for a small fee[6] as well as an additional 17 percent fee[7] on all spirits sales. The initiative required that each store have at least 100,000 square feet of retail and new liquor licenses were provided by the state for purchase.

This market is convenient for this type of analysis for the reasons stated previously, though it does not come without flaws. To estimate consumer demand for spirits, cost estimates are necessary. For this market, it may be difficult to identify marginal cost for each firm. According to the Washington Beer and Wine Distributors Association (WBWDA), liquor sold by grocery stores is still distributed through the state of Washington. Though the prices that consumers see are set by each store, there is a state minimum for each price/quantity bundle. According to the WBWDA, the Washington State Liquor and Cannabis Board (WSLCB) sets the state minimum price equal to the distributor's acquisition cost for the product. I may be able to obtain these state minimum prices for each price/quantity bundle, which will allow me to deal with this issue. It may also be the case that this state minimum is a binding price control, which would limit my ability to observe demand learning through prices, but this would have implications on the way that firms compete in densely populated areas.

There are a couple of other reasons why the Washington state spirits market is an appealing market for a case study of firm learning. The exogenous change in market structure caused by initiative 1183 will provide a clear place to start talking about a shift in market equilibrium. It also seems regulatory changes in alcoholic beverages markets are a hot button topic in the industrial organization literature, as

[6]Fee for liquor licensing is $150 a year.

[7]Which lead to 37.5% effective tax on liquor revenue with previous alcohol taxation.

I have found a couple of job market papers from students attending schools who have a reputation for empirical industrial organization covering this topic using similar models[8]. Additionally, this market shares many of the characteristics that makes the fluid milk market appealing for the price discrimination paper in that the size of the liquor bottles is seemingly exogenous and the market is closely regulated by the WWBDA/WSLCB. The regulation of price by the WWBDA/WSLCB, which creates a price floor at the distributor's cost of obtaining the good, is particularly appealing as this price floor may be a way to proxy for marginal cost for retailers.

## 2.4 A Simple Framework for Firm Demand Learning

The theoretical framework for firm demand learning was adapted from the basic Bayesian Learning model outlined by the Ching, Erdem and Keane (2013) survey of learning models. For this model, I assumed that firms do not know what the exact aggregate demand function of consumers in their geographical market is with certainty. My adaptation of the model Ching et. al (2013)'s basic model deviates in that firms within the Washington spirits market must learn not only what prices to set, but which menu of products is required to maximize profits, whereas Ching et. al (2013) present their framework under the context of consumers learning about some measure of quality for a consumer good. However, I find my adaptation of the basic model to be not only reasonable, but also necessary, as firms consider both dimensions of spirits implied by the aggregate demand function of the consumers they serve for the following reasons:

- Holding inventory is costly for firms. Even though spirits are not a perishable good, it is still in the firm's best interest limit it's back stock inventory. This is particularly true for firms within the grocery/discount store channel, in store inventory space is a commodity and back-stock spirits inventory is more prone to disappearing/being damaged than other types of goods sold by stores within these chains. Thus, if a store manager were to overestimate the consumers purchasing behaviors and order too far too much, or a menu of spirits options that is not ideal for the regional tastes of consumers, it is likely that a large percentage of the back stock will never make it to the store shelf.

- In many states, once the product makes it onto the firm's property and is accounted for, the firm gains ownership of the product. In addition to this, the majority of states do not allow stores to sell back large quantities of back stock to distributors. These states only allow firms to set prices below state minimums after receiving authorization from the state's liquor control agency, which is rarely granted and only under particularly odd circumstances. In combination with the point made before, it is very much in the best interest of these firms to have a reasonable prior belief in regards to the optimal menu of products characterized within the market's aggregate demand function.

[8] Seo (2019), Kim (2016)

- Additionally, recent literature has identified the monetary value of "one-stop" shopping and have found estimates to be non-trivial in magnitude in terms of their economic significance. Thus, it may be the case that a particular firm may not only need to obtain a spirits license and provide some basic menu of options to consumers, but that they need to provide menu that is comparable to their what their competitors offer, or lose customers to their rivals.

Given the points presented in the list above, I find the inclusion of the menu of goods offered by firms to be just as important as which prices to set in terms of what the firm must know to solve their profit maximization problem. As such, my adaption of the model outlined by Ching et. al (2013) constructs a firm's prior and updated beliefs of the relevant aggregate demand function as a function of prior and updated beliefs of menu and price bundles. A complete/more formal list of the assumptions I make in regards to the rules of the demand learning game that firms play are outlined in the list below:

- Firms do not know what the aggregate demand function of consumers in their market is with certainty.
- Firms compete within a market via prices and product menus over periods that are a month in length.[9]
- Beyond seasonality, the aggregate demand function of consumers in a given market does not change over time.[10]
- Firms have some prior belief of the demand conditions[11] in the market they face based upon various demographic characteristics of the firm itself. These characteristics are discussed in further detail in section 2.6 of this chapter.
- Firms learn about demand parameters as follows:
    1. In the first period, $t = 0$, each firm $i$ offers a menu of options $j$, $M_{ijt}$, and at a set of prices, $P_{ijt}$, based upon their prior.
    2. Firms observe the market clear, both within their own store as well as at other stores within the same 3 digit zip code, and adjust their prior. Each firm then may choose to set prices and the menu of options they offer to maximize profits based upon this belief in the current period, or experiment with either the menu of options or prices to learn more about demand in future periods. Given their current set of beliefs in period $t$, firms who choose to experiment are trading current period profits for information about the demand function they face that will lead to higher future period profits.

[9]A month is the approximate amount of time that it takes for a store to receive a new shipment of spirits.

[10]This implies that it is possible for there is a steady state to the game that can be achieved.

[11]Henceforth referred to as their prior.

3. Firms repeat step 2 until they determine there is no incentive to adjust $M_{ijt}$ or $P_{ijt}$ from what they believe is the steady-state equilibrium in the market they serve.

- Due to the potentially infinite number of outcomes that could be profit-maximizing, or close to the profit maximizing outcome, and the noisiness of the signals they receive from observing the market clear, firms must observe several periods of the market clearing in order to have any confidence in their beliefs regarding the demand curve that they face.

- This process ignores potential competitive effects and focuses only on learning about demand, not on how their competitors compete in the market as well.[12]

### 2.5 Data

In this section of this chapter, I discuss the data used to identify learning in the Washington state spirits market, and describe the characteristics of these data.

The WSLCB defines spirits as follows: "any beverage containing alcohol that is obtained by distillation, except flavored malt beverages as defined under RCW 66.04.010(20). These products have labels that indicate "proof" and wines exceeding 24 percent alcohol by volume as indicated on the label."[13] I define "spirits" to be any and all: Bourbon, Canadian Whiskey, Irish Whiskey, Whiskey, Scotch, Gin, Vodka, Rum or Tequila. I omit Cognac and Brandy as their definition is somewhat ambiguous since some Cognac and Brandys are lower than 24 proof and may be considered wine due to their distillation process. As is the case, I omit Cognac and Brandy from these analyses, as I do not observe proof for each UPC with these data.[14]

I use Nielsen Scanner data for this paper. The Nielsen Panel would be more appropriate for identifying the demand parameters in this market, as there may not be sufficient variation in household characteristics when using the Scanner data. Additionally, I use the Nielsen Scanner data to describe how the menu of options that firms offer changes over time.[15]

The Nielsen Panel provides household level consumption of goods over time. Additionally, these data include the characteristics of each household. See section 2.11 and summary statistics of these data.

The Nielsen Scanner data provide the weekly sales of each UPC offered by firms that participate. These data include product characteristic of each UPC sold, as well

[12] i.e. it could be optimal for firms to utilize a mix-strategy in terms of pricing or the menu they offer in more competitive markets. This is outside of my model's framework.

[13] Department of Revenue: Washington State, dor.wa.gov

[14] This is something I can go back to, but I am having trouble finding a reliable source for looking up UPCs at this moment. I have applied for a subscription for a UPC search website, but I am still waiting to hear back from the company (they screen how these descriptions are used, likely due to some agreement with brand producers). With UPC descriptions, I am able to classify each Cognac and Brandy according to WSLCB's definition and add some of these to the sample.

[15] At this time, this is the best way that I can observe the menu of options. I attempt to identify which grocery stores sell spirits and contact them about whether or not they have specialty spirits sections.

as the general location of the store.[16] See section 2.11 summary statistics of these data.

## 2.6 Potential Margins Where Differential Learning May Occur

In the theoretical framework presented in section 2.4 of this chapter, I defined both price and menu prior belief for firm $i$ to be a function of the firm's characteristics. In this section, I provide several margins for which the prior the firm starts with or the rate the firm learns at may vary. For each margin in this section, I first propose the margin which may lead to differential learning. I then discuss how this should matter in the context of learning in terms of the speed in which firms are able to update their prior beliefs to a belief that is far more representative of the true demand conditions within the market or in terms of how the firm constructs their prior.

Below are such margins where learning may vary amongst firms:

- **Firms that are a part of a national chain that operates in states with spirits laws similar to those found in Washington state after the enactment of initiative 1183 should, on average, have an initial prior belief of the demand function they face that is closer to the true demand function, in terms of what products to offer and which prices to set, than firms who do not have chains within states with similar laws. I.e. mom-and-pop/Washington only stores.**

  This statement suggests that national chains, such as Kroger, may be able to use data from markets that are similar to another market they already operate in when determining their prior belief, whereas mom-and-pop stores do not have these data available to them when constructing their prior beliefs. I find a story that Kroger has a menu/prices that it always chooses to start in a new region equally as compelling as this margin, ex-ante.

  To test this empirically, I decompose the variance of prices set by firms over a month-long period and include a variable in the regression that measures how many other locations the parent company owns within Washington state.[17] A negative coefficient estimate would suggest confirm this hypothesis, as it would suggest that as the number of stores a parent company owns increases the variance in the prices they set decreases, which would be consistent with a firm having a better prior belief of demand conditions at the onset of the law change.

[16] 3 digit zip codes, Designated Market Areas, Nielsen Scantrack and County

[17] Due to the size of these data, I only see sales in Washington state and cannot identify the name of the parent company. Thus the number of firms under the same parent company is a proxy for national chain status.

- **Firms located in markets that are more competitive, should learn what prices to set faster than stores located in areas where there is more local monopoly power.**

  Since firms are competing in prices, firms located in geographical markets where there are more competitors present will not be able to extract as much surplus from consumers via large markups vs. stores that serve a market as a local monopolist.[18] The only way that this may have any effect in terms of the menu of products to choose is if firms begin competing via menu they offer[19] rather than prices, though I must make some concession that firms in these areas can walk into their competitor's store to see what menu of spirits they offer. However, I find this statement more compelling in terms of how firms construct their prior beliefs for pricing that firms will be able to update their prior beliefs.

  To test this hypothesis, I include a HHI measure in my variance of prices decomposition. A positive coefficient estimate would confirm this hypothesis, as it would suggest that as markets become more concentrated, the variance in prices increases.

- **The speed of learning may vary based upon the channel type of the firm in question. For example, channel types with that have a lot of foot-traffic on average, such as grocery stores, discount stores, etc., may learn about the demand for spirits quicker than those such as gas stations/convenience store type channels.**

  However, I am less confident that this particular margin matters as it is not clear to me that grocery stores compete in the same market as gas stations/convenience stores. Those purchasing liquor at the gas station/convenience store are far more concerned with convenience whereas those shopping at grocery stores are more interested in prices/variety. As such, the demand functions that stores in channels that are fairly dissimilar may be quite different and it may be potentially unreasonable to apply the same sort of framework due to this.

  However, if it is the case that firms within different channels face the same aggregate demand function,[20] then stores that generally have more foot-traffic due to the other goods that they sell[21], then one may think that one month of sales could provide a less noisy signal to firms operating within channels such as grocery stores rather than those who belong to the gas station/convenience store type channels.

[18] i.e. Seattle and Spokane vs. the Mountainous Region located in the middle of the state.

[19] A specialty store rather vs. a cheaper retail option.

[20] Which happens to be implicitly implied by the assumptions listed in the framework presented in section 2.3.

[21] Seo (2019) estimates the value of one stop shopping and shows that the value of the convenience is not trivial in magnitude.

To test this empirically, I include store channel type fixed effects in my regression.

### 2.7 Preliminary Evidence of Demand Learning

Preliminary market level evidence of learning behavior are presented in figure 2.5, figure 2.6 and figure 2.7, which are each discussed in detail in the following paragraphs.

Figure 2.5 presents the state-wide mean price of pint, fifth, and handle sized containers of spirits during each week following the privatization of spirits sales in December 2016 dollars. From this figure, two things are clear. First, it is clearly the case the spirits prices are seasonal and are sold at higher prices during the winter holiday season. However, by comparing the timing of the price drop following the holiday season over the years shows that firms are learning how and when they are able to charge the higher seasonal prices. This is most clearly illustrated by comparing timing of the price changes following the 2012 and 2015 holiday seasons. Following the 2012 holiday season, prices remained at their holiday premium approximately the beginning of March in the following year, whereas following the 2015 holiday season, prices fell almost immediately in the following January. Looking at the years between these two, it appears that the timing of the price drop has been gradually moving towards the 2015 holiday season timing. From this I conclude that by experimentation with pricing during, firms are learning how to better price discriminate during the holiday season.

Figure 2.6 presents the state-wide standard error of the mean price of pint, fifth and handle sized containers of spirits by week. The standard errors that are presented do not take container size into account, thus if the amount of variation in prices were consistent with the size all else constant, then the SE of pint sized containers should be smaller than fifths which would be smaller than handles.[22] However, this is not the case. Instead, pint sized containers have the largest SE of the three product sizes until about 2014, suggesting firms experiment with the prices of pint sized containers far more than they do with the larger sized products.

Figure 2.7 presents the total number of unique UPCs sold state-wide during each week.[23] This figure shows that firms entered the market with a limited menu of

[22]Since the prices of the large sizes are larger than the smaller sizes due to the difference in the amount of product sold per container.

[23]Note: This may currently be measured with error in two ways. First, measure this does not take into account any changes the total number of retailers in the marketplace during a given week, but according to Seo (2019) the number of retailers serving the market was fairly constant after August 2012.
My other concern is in regards to the spikes in UPCs during the holiday season of every year. A potential explanation for these spikes is that retailers offer more variety in the product menus during the holiday, which would be consistent with (1) in some regards. However, a common practice during the holiday season in the spirits market is to sell fifths in gift package containers which commonly include branded glasses or other collectible memorabilia. These gift sets sell the exact same spirits product, but present it in a more gift friendly way and thus would be treated identically to their counterpart sold on the shelf without the fancy branded memorabilia. I am not certain if the gift sets are sold under a different UPC than the products sold on the shelf – I have submitted this

initially and have expanded this menu to take market specific preferences into account over time. The slowing rate at which this measure has increased in the later years could be consistent with firms slowly approaching a long-run steady state for the optimal menu required in this particular market.

## 2.8 Regression Results

### 2.8.1 Model Specification

To empirically test the margins presented in section 2.6, I estimate three reduced-form regression models. The first of which is a reduced-form regression model that decomposes the variance of the prices set by firms over a month long period. This model aims to explain what factors determine the variance in prices observed in section 2.7. I specify the model as follows:

$$Var(Price)_{ms} = \beta_1 HHI_{ms} + \beta_2 Parent_s + Channel_s \gamma + T_m \alpha + X_{ms} + \epsilon_{ms} \tag{2.1}$$

where Var(Price)$_{ms}$ is the variance of prices during a particular month, $m$, at a particular store, $s$, $Parent_s$ indicates how many other locations store $s$'s parent company owns and operates. $HHI_{ms}$ is the HHI within the 3 digit zip code that store $s$ operates in during month $m$. $Channel_s$ is a vector of channel type dummy variables that indicate whether or not store $s$ is a drug store, grocery store or mass merchandiser; $T_m$ are vectors of month and year fixed effects and $X_{ms}$ is a matrix of UPC specific product characteristics. $\epsilon_{ms}$ represents a classical OLS stochastic error term for store $s$ during month $m$.

The second model I estimate decomposes the variance in the size of the menu of options offered by a firm. I estimate this model to help explain how firms learn which products to offer over time. I specify this model identically to the price variance model with the only exception being that the dependent variable is menu size rather than price variance:

$$MenuSize_{ms} = \beta_1 HHI_{ms} + \beta_2 Parent_s + Channel_s \gamma + T_m \alpha + X_{ms} + \epsilon_{ms} \tag{2.2}$$

Lastly, I estimate a price hedonic model with independent variables identical to those used in the price variance model as a robustness check to ensure that prices generally behave as one would expect. This model is specified below where $Price_{ms}$ is the average price of a UPC set by firm $s$ during month $m$.

$$Price_{ms} = \beta_1 HHI_{ms} + \beta_2 Parent_s + Channel_s \gamma + T_m \alpha + X_{ms} + \epsilon_{ms} \tag{2.3}$$

### 2.8.2 Results: Tables 2.1 & 2.2 Price Variance Regression

Table 2.1 presents the regression results of equation (2.1) with HHI entering the regression linearly whereas table 2.2 presents the estimates of equation (2.1) with HHI binned. In both tables, columns (1) and (2) present the model without brand

question to the Nielsen data forum and have yet to hear back.

fixed effects whereas (3) and (4) include brand fixed effects. In each specification of the model where HHI enters linearly, the coefficient estimate associated with HHI is positive and statistically significant. This suggests that as markets become more concentrated, firms experiment more with prices, a result consistent with this idea that firms in more competitive markets benefit from significant informational spillover effects by observing the prices set by their competitors.

The binned HHI results found in table 2.2 tell a very similar story given that the omitted case is a bin for highly concentrated markets with HHI greater than 2500. The negative parameter estimates associated with the HHI bins of markets less concentrated suggests that more competitive markets learn faster than those with less competition. The positive coefficient estimates associated with the monopolist suggest that the monopolist in my sample sets prices with far more variation. The rest of the parameter estimates associated with table 2.2 agree in sign with those in table 2.1 and have similar magnitudes, thus I will only discuss the results associated with table 2.1 for the rest of this discussion.

The only other statistically significant results I find in this model are that firms appear to be experimenting more with fifth and half-gallon sized containers in comparison to pints, a result that appears at odds with the anecdotal evidence presented in section 2.7. This is likely due to a sample size problem. Firms sell far more fifth/half-gallon sized containers of spirits than they do pints. Since I only observe market sales rather than the actual menu of options offered by firms with the Nielsen data, it could be the case that firms are actually experimenting more with the prices of pint-sized containers, but if this were the case it would be to their own detriment.

It is also interesting that there is no statistically significant difference to the degree of experimentation in prices across channel types. Ex-ante, one may expect drug stores[24] to be more open to experimentation than grocery or mass mechandisers, due to their corporate structure being seemingly less centrally determined. However, my findings do not suggest that this is the case.

In addition, it is interesting that the November and December dummy variables, which are compared to January are not statistically significant. I find this counter intuitive given the anecdotal evidence of price discrimination across the holiday season that is present in section 2.7.

### 2.8.3 Results: Tables 2.3 & 2.4 Determinants of the Total Number of UPCs Sold

Similarly, table 2.3 presents the results of the regression specified by equation (2.2) with HHI entering the regression linearly, whereas table 2.4 presents estimates with of the same model, but with HHI entering the regression nonlinear via bins. These results seek to explain how firms determine/experiment with the menu of options they offer to consumers. In table 2.3, specification (1) presents the model without product characteristics whereas specification (2) includes product characteristics, however the results are almost identical. The most glaring result from this model is that the

[24] The omitted channel type.

coefficient estimate associated with HHI is more than an order of magnitude larger than the rest of the coefficient estimates. However, this is likely due scaling issues as the direct interpretation of this coefficient estimate is "as HHI increases by 1, i.e. going from a perfectly competitive market to a monopoly, the monopolist offers 85 less options to its consumers." Considering that direct interpretation, as well as that the number of options offered during a store-month period in my data ranges from 2 to 220 options, this result is much easier to digest.

Table 2.4's parameter estimates associated with HHI are much more reasonable in terms of their magnitude. The omitted bins for table 2.4 are for highly concentrated markets with HHI greater than 2500 and the monopolist in my sample. Although the magnitudes of these parameter estimates are much more reasonable than those presented previously in table 2.3, the signs tell the same story as before. The statistically significant parameter estimate of 12.2939 for highly competitive markets in specification (2) of table 2.4 is interpreted to say that highly competitive markets offer 12.2939 more unique UPCs to their customers on average than markets that are highly concentrated or served by a monopolist. The magnitude of the parameter estimates associated with "Unconcentrated" markets and the even smaller parameters associated with markets that are "Moderately Concentrated" only further confirm the story told by table 2.3. Similarly to the comparison of tables 2.1 and 2.2, the rest of the parameter that are estimated in each specification are very similar in terms of magnitude and are identical in terms of sign and statistical significance. Thus, the rest of this discussion will only be in regards to table 2.3's results.

Another particularly interesting result is that the coefficient estimate associated with the total number of stores in parent chain is positive and statistically significant. This would suggest that a chain such as Kroger, offers its customers a greater selection of products compared to a mom-and-pop store. This result is consistent with the idea that larger chains must provide more variety, as they serve a much more diverse consumer-base than non-chain stores.

I also find it interesting that the coefficient estimate associated with the Food/Grocery channel is negative and statistically significant, but the Mass Merchandiser is not,[25] suggesting that Food/Grocery stores offer less selection than the other two chains.

Lastly, I find the coefficient estimates associated with November and December to both be statistically significant and negative. This is consistent with the idea that the gift boxes offered during the holiday seasons, which usually include a glass to drink the spirit from or something of the like at no additional monetary fee, come at the fee of having fewer options to choose from during the holiday season.

### 2.8.4 Results: Tables 2.5 & 2.6 Price Hedonic Models

Following the same pattern as the other two sections of results in this chapter, table 2.5 presents the regression results associated with the price hedonic model outlined by equation (2.3) with HHI entering the regression linearly, whereas table 2.6 presents equation (2.3) with HHI nonlinearly binned. The first two columns of this table

[25]Both of which are being compared to Drug Stores.

presents the model without brand fixed effects, whereas specifications (3) and (4) present the model with brand fixed effects. I find the brand fixed effects specifications more compelling in this setting, thus I will continue my discussion on this table only considering those results.

Surprisingly, the regression results presented for the linear specification in table 2.5 suggest that there is no statistically significant difference in prices that can be attributed to market concentration. This result could be due to either of the following: the WSLCB imposes a price floor in this market and this result is consistent with that price floor being binding in all markets in Washington State regardless of HHI; it could be an observation issue in that I only observe sales to consumers.

However, table 2.6's parameter estimates tell a different story. When HHI enters the model nonlinear via the bins associated with the FTC's guidelines, my results are statistically significant in many of the different specifications. However, since the omitted bin in this specification represents highly concentrated markets with HHI greater than 2500, the sign of these parameter estimates is somewhat puzzling. The direct interpretation of the parameter estimate of 0.7423 that is associated with highly competitive markets and is estimated with controls and brand fixed effects is that on average, liquor is about 74 cents more expensive in highly competitive markets than in highly concentrated markets, which is the opposite of what economic theory would suggest.[26] These almost certainly endogenous results associated with highlight the discussion presented in chapter 1 of this dissertation where I motivate the need for structural modeling when estimating demand models in the fluid milk market. Since the simultaneity of the demand system causes prices to be endogenous these parameters are difficult to identify without an instrumental variables approach when doing so in a reduced-form setting.

However, one encouraging feature of these parameter estimates is that controlling for product characteristics and including brand fixed effects in particular, seems to do some good in dealing with the endogeneity of these parameters. Although the sign of the parameter estimates associated with the HHI bins in specifications (3) and (4), which include brand fixed effects, are still opposite of what economic would suggest, they are much smaller in terms of magnitude than the parameter estimates in specifications (1) - (2) where brand fixed effects are absent. This suggests that some, but unfortunately not all, of the endogeneity associated with these estimates is dealt with via the inclusion of brand fixed effects.

The remaining estimated parameters between tables 2.5 and 2.6 are identical in terms of sign and statistical significance, and are very similar in magnitude, thus for the rest of this section I will discuss the results presented in table 2.5.

The coefficient estimates associated with the size indicator variables are consistent with prices increasing somewhat linearly across the menu of sizes offered on average.

The November and December coefficient estimates are consistent with the holiday season price discrimination story outlined in section 2.1.

[26] At least this is the case if we ignore some story about how more competitive markets present firms with an opportunity to coordinate and collude that is not available in those that are more concentrated and less competitive.

## 2.9 Conclusion

In summary, this paper identifies firm demand learning via reduced-form methods in the Washington state spirits market ex-post initiative 1183. Through graphical/anecdotal evidence, I show prices and menu offerings leveling off towards a long-run steady state equilibrium over time and through the reduced-form regressions estimated, I am able to confirm several, but not all, of the margins considered in section 2.6.

This paper contributes to the existing literature on demand learning by being one of the first to examine firm learning in a completely new market. Other papers have shown the value of information and firm demand learning, but to my knowledge, no paper has shown this type of behavior in a brand new market without utilizing structural estimation techniques. Additionally, it may be interesting to estimate the priors that firms may have on demand conditions for an area that are a function of another region of a similar consumer distribution. Firms who are looking to expand their territory should be interested in this analysis as if I find that these priors are fairly close to the true demand parameters, they will be able to assume that they can use the menu and pricing strategies that they are currently using in other markets.

## 2.10 Tables and Figures

Figure 2.1: Number of Stores by 3DZ: 2012

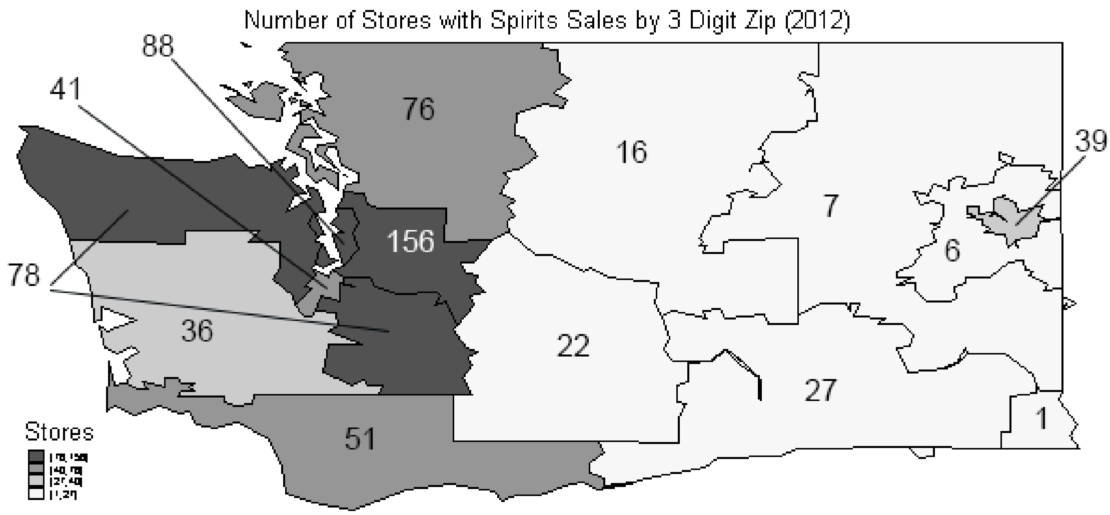



Figure 2.2: Number of Stores by 3DZ: 2013

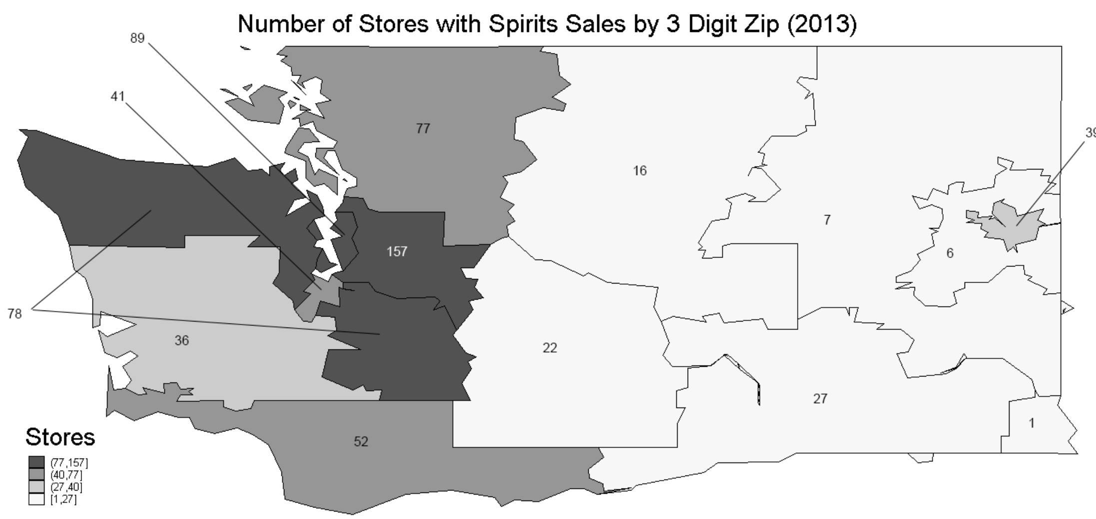


Figure 2.3: Number of Stores by 3DZ: 2014

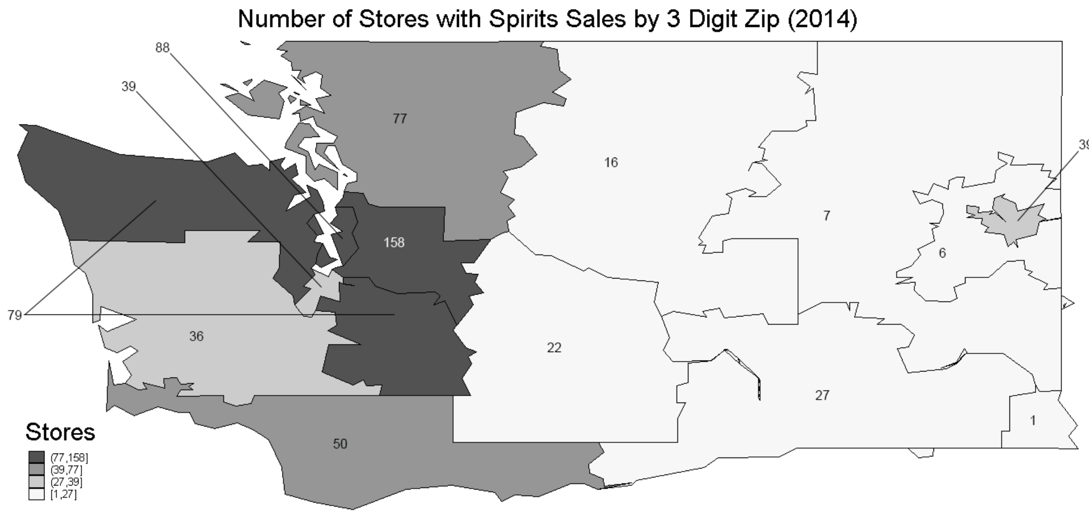



Figure 2.4: Number of Stores by 3DZ: All Years

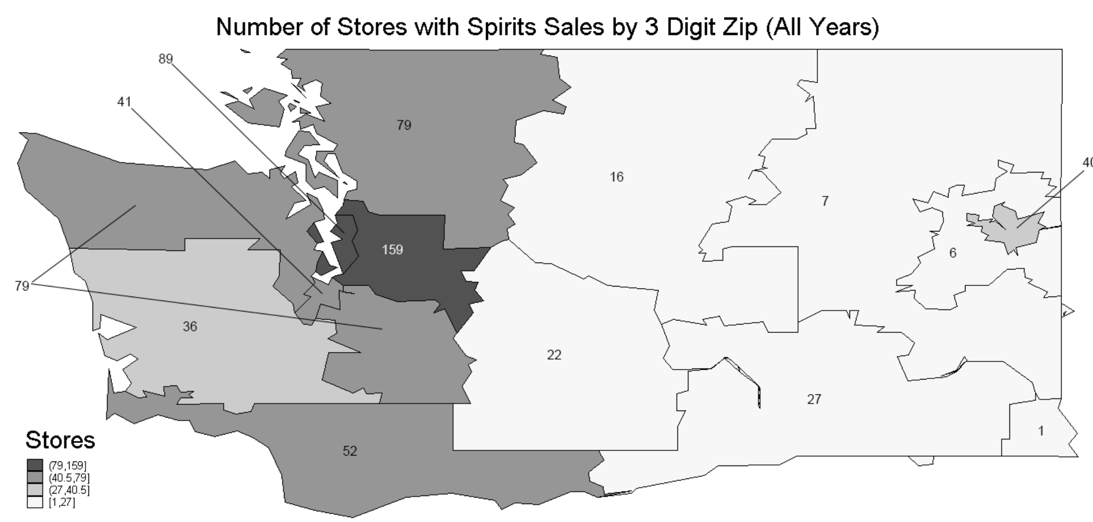


Figure 2.5: Spirits Summary Statistics: Mean Price by Size and Week

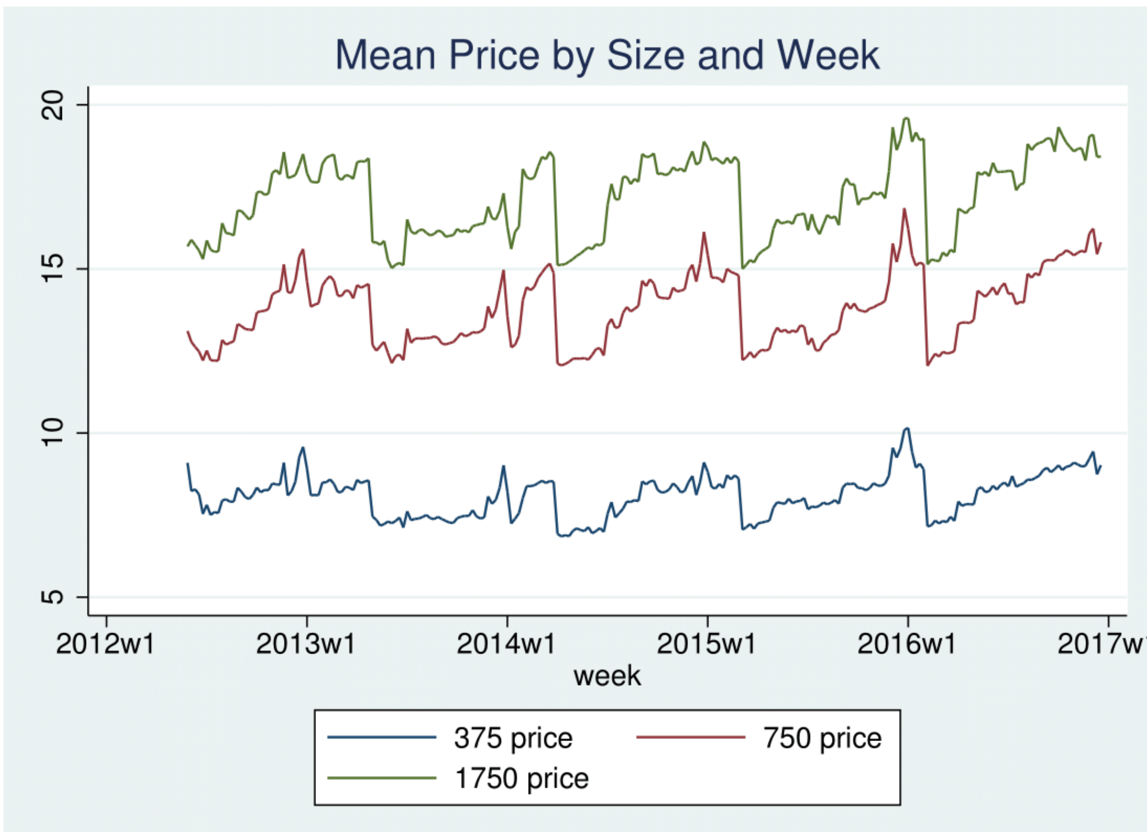


Figure 2.6: Spirits Summary Statistics: Standard Error by Size and Week

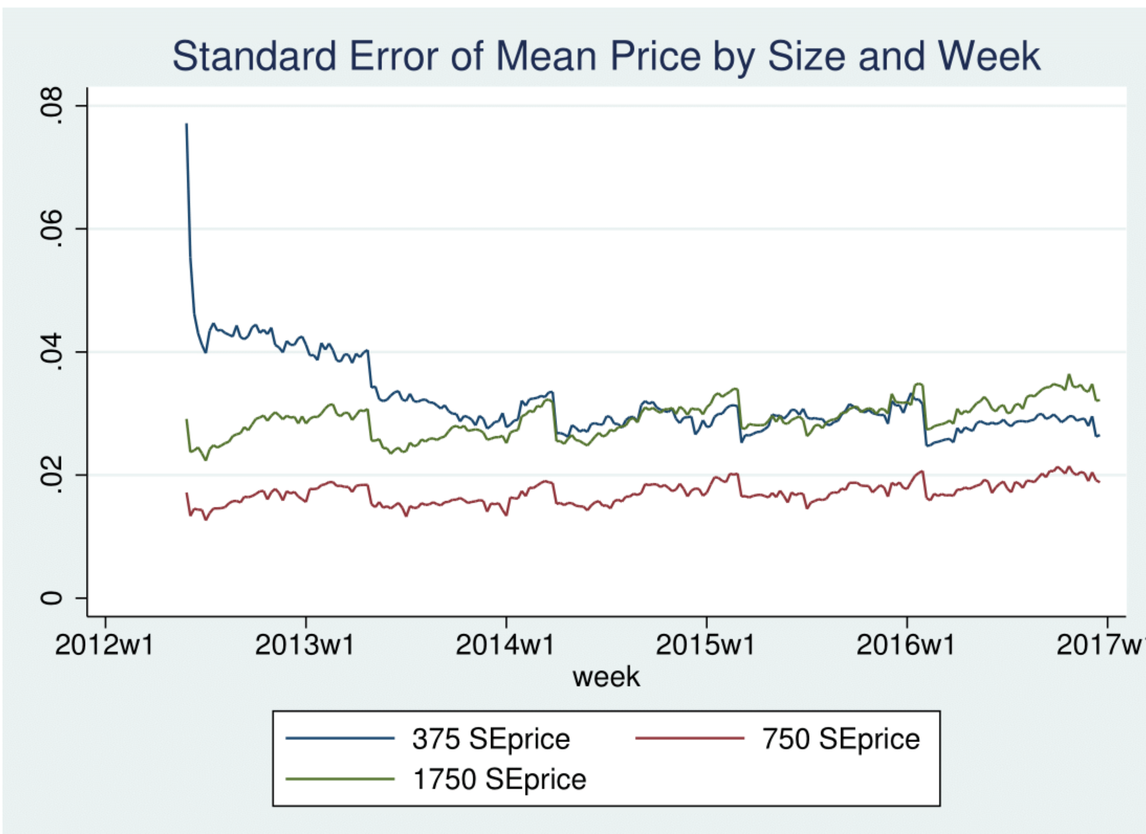


Figure 2.7: Spirits Summary Statistics: Total Number of Unique UPCs Sold

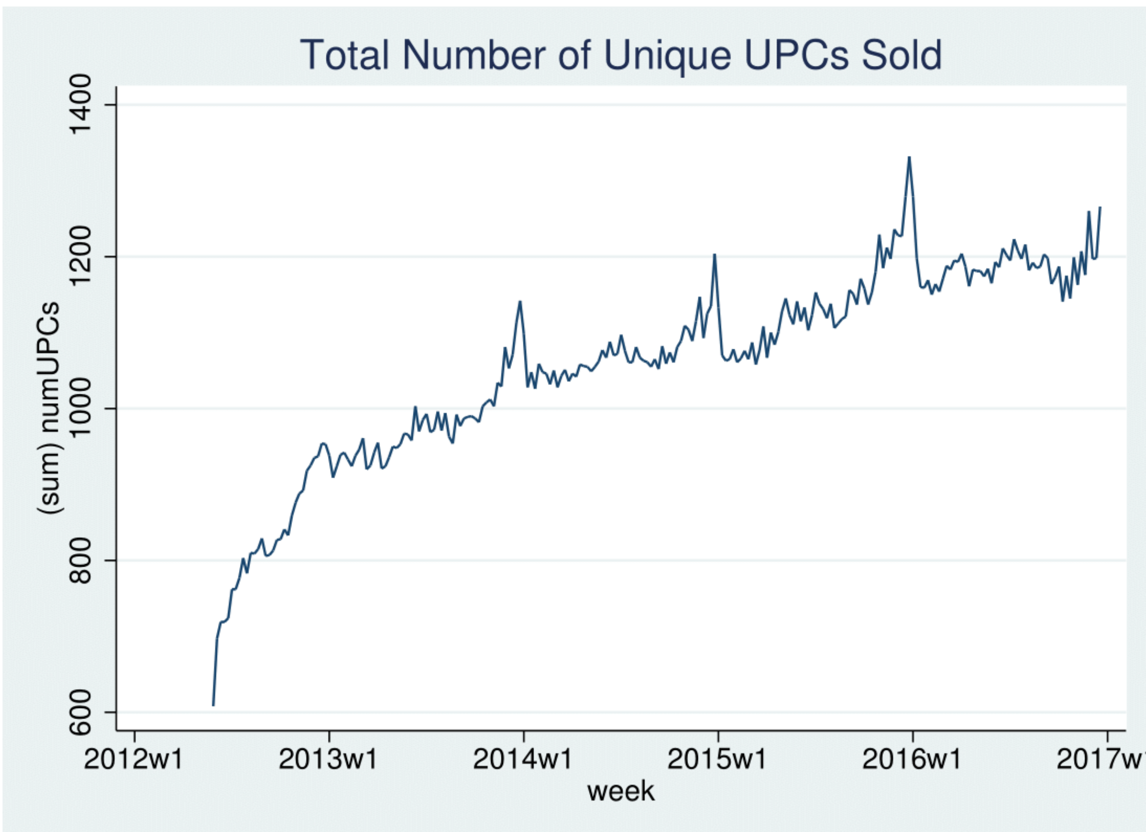


Table 2.1: Decomposition of the Variance of Prices

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| HHI in 3-Digit Zipcode | 0.5753*** | 0.5791*** | 0.5242*** | 0.5303*** |
| | (0.1258) | (0.1263) | (0.1189) | (0.1228) |
| Total Number of Stores in Parent Chain | −0.0005 | −0.0005 | −0.0007 | −0.0007 |
| | (0.0007) | (0.0007) | (0.0006) | (0.0006) |
| Food Channel Dummy | −0.0295 | −0.0368 | 0.0236 | 0.0212 |
| | (0.0888) | (0.0886) | (0.0733) | (0.0740) |
| Mass Merchandiser Dummy | −0.0255 | −0.0307 | 0.0214 | 0.0162 |
| | (0.0880) | (0.0874) | (0.0703) | (0.0701) |
| Fifth Size Container Dummy | 0.1994*** | 0.2061*** | 0.3931*** | 0.3875*** |
| | (0.0502) | (0.0516) | (0.0846) | (0.0846) |
| Half Gallon Size Container Dummy | 0.0812** | 0.0778** | 0.3180*** | 0.3144*** |
| | (0.0295) | (0.0302) | (0.0707) | (0.0708) |
| November Dummy | −0.0047 | −0.0044 | −0.0033 | −0.0029 |
| | (0.0044) | (0.0043) | (0.0038) | (0.0036) |
| December Dummy | 0.0027 | 0.0027 | 0.0040 | 0.0042 |
| | (0.0025) | (0.0021) | (0.0026) | (0.0023) |
| Flavor Dummy | | −0.0571** | | 0.0272 |
| | | (0.0228) | | (0.0717) |
| Special Blend Dummy | | 0.0496 | | 0.0997 |
| | | (0.0425) | | (0.2150) |
| Controls | N | Y | N | Y |
| Brand Fixed Effects | N | N | Y | Y |
| Obs. | 1,365,820 | 1,365,820 | 1,365,820 | 1,365,820 |
| Brands | . | . | 466 | 466 |

Note: There are 1,365,820 sales of spirits. January is the omitted month dummy variable. Controls include month, year, channel, spirit type, feature, display, flavor and special blend dummy variables. Standard errors are clustered at the store chain level for all regressions. Standard errors are shown in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 2.2: Decomposition of the Variance of Prices: Binned HHI

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| HHI3DZ: Highly Comp (HHI < 150) | −0.2022** | −0.2025** | −0.1904** | −0.1919** |
| | (0.0698) | (0.0677) | (0.0544) | (0.0530) |
| HHI3DZ: Uncon (150 < HHI < 1500) | −0.1808** | −0.1835** | −0.1603** | −0.1627** |
| | (0.0717) | (0.0698) | (0.0573) | (0.0558) |
| HHI3DZ: Mod Con (1500 < HHI < 2500) | −0.1374 | −0.1357 | −0.1040* | −0.1036* |
| | (0.0723) | (0.0706) | (0.0519) | (0.0502) |
| HHI3DZ: Monopolist | 0.2204** | 0.2230** | 0.1677* | 0.1712* |
| | (0.0876) | (0.0893) | (0.0770) | (0.0782) |
| Total Number of Stores in Parent Chain | −0.0005 | −0.0005 | −0.0007 | −0.0007 |
| | (0.0008) | (0.0008) | (0.0007) | (0.0007) |
| Food Channel Dummy | −0.0287 | −0.0356 | 0.0249 | 0.0227 |
| | (0.0892) | (0.0890) | (0.0745) | (0.0750) |
| Mass Merchandizer Dummy | −0.0255 | −0.0306 | 0.0211 | 0.0161 |
| | (0.0883) | (0.0876) | (0.0706) | (0.0701) |
| Fifth Size Container Dummy | 0.1999*** | 0.2066*** | 0.3931*** | 0.3876*** |
| | (0.0500) | (0.0514) | (0.0842) | (0.0842) |
| Half Gallon Size Container Dummy | 0.0821** | 0.0787** | 0.3186*** | 0.3151*** |
| | (0.0295) | (0.0302) | (0.0708) | (0.0708) |
| November Dummy | −0.0047 | −0.0045 | −0.0032 | −0.0029 |
| | (0.0044) | (0.0043) | (0.0038) | (0.0036) |
| December Dummy | 0.0026 | 0.0027 | 0.0039 | 0.0042 |
| | (0.0025) | (0.0021) | (0.0026) | (0.0023) |
| Flavor Dummy | | −0.0572** | | 0.0260 |
| | | (0.0230) | | (0.0702) |
| Special Blend Dummy | | 0.0494 | | 0.1013 |
| | | (0.0426) | | (0.2154) |
| Controls | N | Y | N | Y |
| Brand Fixed Effects | N | N | Y | Y |
| Obs. | 1,365,820 | 1,365,820 | 1,365,820 | 1,365,820 |
| Brands | . | . | 466 | 466 |

Note: There are 1,365,820 sales of spirits. January is the omitted month dummy variable. Controls include month, year, channel, spirit type, feature, display, flavor and special blend dummy variables. Standard errors are clustered at the store chain level for all regressions. Standard errors are shown in parentheses *** p<0.01, ** p<0.05, * p<0.1. HHI bins constructed based upon HHI classifications found in the FTC's Merger Guidelines. The omitted bin for HHI is for highly concentrated markets (HHI > 2500).

Table 2.3: Determinants of the Total Number of UPCs Sold

| | (1) | (2) |
|---|---|---|
| HHI in 3-Digit Zipcode | $-84.8904^{***}$ | $-85.2861^{***}$ |
| | (12.7817) | (12.7591) |
| Total Number of Stores in Parent Chain | $0.0447^{***}$ | $0.0445^{***}$ |
| | (0.0035) | (0.0036) |
| Food Channel Dummy | $-1.6474^{***}$ | $-1.7306^{***}$ |
| | (0.3721) | (0.3588) |
| Mass Merchandiser Dummy | 0.8159 | 0.8268 |
| | (0.4903) | (0.4796) |
| Fifth Size Container Dummy | 0.5877 | 0.8417 |
| | (0.4960) | (0.5140) |
| Half Gallon Size Container Dummy | 0.5361 | 0.5116 |
| | (0.5408) | (0.5557) |
| November Dummy | $-0.0762^{**}$ | $-0.0798^{**}$ |
| | (0.0227) | (0.0215) |
| December Dummy | $-0.0393^{*}$ | $-0.0455^{**}$ |
| | (0.0165) | (0.0160) |
| Flavor Dummy | | $-1.3502^{***}$ |
| | | (0.3187) |
| Special Blend Dummy | | 0.0344 |
| | | (0.5885) |
| Controls | N | Y |
| Obs. | 1,715,381 | 1,715,381 |

Note: There are 1,715,381 sales of spirits. January is the omitted month dummy variable. Controls include month, year, channel, spirit type, feature, display, flavor and special blend dummy variables. Standard errors are clustered at the store chain level for all regressions. Standard errors are shown in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 2.4: Determinants of the Total Number of UPCs Sold: Binned HHI

| | (1) | (2) |
|---|---|---|
| HHI3DZ: Highly Comp (HHI < 150) | 12.2190*** | 12.2939*** |
| | (1.4039) | (1.4087) |
| HHI3DZ: Uncon (150 < HHI < 1500) | 6.0730*** | 6.1255*** |
| | (0.2189) | (0.2223) |
| HHI3DZ: Mod Con (1500 < HHI < 2500) | 1.3328*** | 1.3578*** |
| | (0.2700) | (0.2694) |
| Total Number of Stores in Parent Chain | 0.0396*** | 0.0395*** |
| | (0.0030) | (0.0030) |
| Food Channel Dummy | −1.1408** | −1.2346** |
| | (0.3738) | (0.3621) |
| Mass Merchandizer Dummy | 1.0536* | 1.0475* |
| | (0.4633) | (0.4466) |
| Fifth Size Container Dummy | 0.6104 | 0.8576 |
| | (0.4896) | (0.5185) |
| Half Gallon Size Container Dummy | 0.5946 | 0.5591 |
| | (0.5496) | (0.5620) |
| November Dummy | −0.0772** | −0.0800** |
| | (0.0234) | (0.0225) |
| December Dummy | −0.0419** | −0.0476*** |
| | (0.0134) | (0.0128) |
| Flavor Dummy | | −1.3849*** |
| | | (0.3336) |
| Special Blend Dummy | | 0.0700 |
| | | (0.5935) |
| Controls | N | Y |
| Obs. | 1,715,381 | 1,715,381 |

Note: There are 1,715,381 sales of spirits. January is the omitted month dummy variable. Controls include month, year, channel, spirit type, feature, display, flavor and special blend dummy variables. Standard errors are clustered at the store chain level for all regressions. Standard errors are shown in parentheses *** $p<0.01$, ** $p<0.05$, * $p<0.1$. HHI bins constructed based upon HHI classifications found in the FTC's Merger Guidelines. The omitted bins for HHI are for highly concentrated markets (HHI > 2500) and the monopolist in my sample.

Table 2.5: Price Hedonic Models

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| HHI in 3-Digit Zipcode | −9.7111 | −9.6121 | −1.1691 | −1.3179 |
| | (5.0411) | (4.9834) | (1.1968) | (1.2375) |
| Total Number of Stores in Parent Chain | −0.0026 | −0.0011 | −0.0022 | −0.0025 |
| | (0.0050) | (0.0054) | (0.0033) | (0.0035) |
| Food Channel Dummy | −2.4533*** | −2.6008*** | −0.0562 | −0.0026 |
| | (0.4999) | (0.5474) | (0.4672) | (0.4911) |
| Mass Merchandiser Dummy | −1.2302** | −1.1333* | −0.2012 | −0.0964 |
| | (0.4892) | (0.5770) | (0.4227) | (0.4575) |
| Fifth Size Container Dummy | 5.4377*** | 5.6731*** | 7.0999*** | 7.1938*** |
| | (0.4201) | (0.4434) | (1.1107) | (1.1371) |
| Half Gallon Size Container Dummy | 10.4680*** | 10.6474*** | 17.2310*** | 17.3052*** |
| | (0.4452) | (0.5542) | (1.2864) | (1.3097) |
| November Dummy | −0.2516*** | −0.2561*** | −0.1713*** | −0.1748*** |
| | (0.0198) | (0.0180) | (0.0176) | (0.0179) |
| December Dummy | 0.3777*** | 0.3730*** | 0.4446*** | 0.4417*** |
| | (0.0266) | (0.0259) | (0.0510) | (0.0501) |
| Flavor Dummy | | −0.1694 | | −0.0210 |
| | | (0.3917) | | (0.2324) |
| Special Blend Dummy | | 5.3456** | | −1.3948** |
| | | (1.8823) | | (0.3875) |
| Controls | N | Y | N | Y |
| Brand Fixed Effects | N | N | Y | Y |
| Obs. | 1,365,820 | 1,365,820 | 1,365,820 | 1,365,820 |
| Brands | . | . | 466 | 466 |

Note: There are 1,365,820 sales of spirits. January is the omitted month dummy variable. Controls include month, year, channel, spirit type, feature, display, flavor and special blend dummy variables. Standard errors are clustered at the store chain level for all regressions. Standard errors are shown in parentheses *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6: Price Hedonic Models: Binned HHI

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| HHI3DZ: Highly Comp (HHI < 150) | 3.6910*** | 3.6783*** | 0.7078** | 0.7423** |
| | (0.2732) | (0.3030) | (0.1973) | (0.2238) |
| HHI3DZ: Uncon (150 < HHI < 1500) | 2.1074*** | 2.1037*** | 0.4018** | 0.4554** |
| | (0.2833) | (0.3078) | (0.1252) | (0.1591) |
| HHI3DZ: Mod Con (1500 < HHI < 2500) | 0.4185* | 0.4463** | 0.0620 | 0.0576 |
| | (0.1789) | (0.1765) | (0.1084) | (0.1385) |
| HHI3DZ: Monopolist | −0.2847 | −0.2495 | 0.7394** | 0.6505* |
| | (0.4222) | (0.4604) | (0.2875) | (0.3045) |
| Total Number of Stores in Parent Chain | −0.0015 | −0.0000 | −0.0019 | −0.0021 |
| | (0.0042) | (0.0046) | (0.0031) | (0.0033) |
| Food Channel Dummy | −2.4905*** | −2.6395*** | −0.0679 | −0.0179 |
| | (0.4683) | (0.5152) | (0.4627) | (0.4856) |
| Mass Merchandizer Dummy | −1.2282** | −1.1350* | −0.1951 | −0.0930 |
| | (0.4664) | (0.5551) | (0.4276) | (0.4620) |
| Fifth Size Container Dummy | 5.4425*** | 5.6757*** | 7.1083*** | 7.2002*** |
| | (0.4379) | (0.4594) | (1.1110) | (1.1376) |
| Half Gallon Size Container Dummy | 10.4768*** | 10.6533*** | 17.2321*** | 17.3049*** |
| | (0.4554) | (0.5589) | (1.2850) | (1.3083) |
| November Dummy | −0.2542*** | −0.2587*** | −0.1721*** | −0.1755*** |
| | (0.0188) | (0.0170) | (0.0175) | (0.0178) |
| December Dummy | 0.3765*** | 0.3718*** | 0.4442*** | 0.4414*** |
| | (0.0265) | (0.0259) | (0.0508) | (0.0499) |
| Flavor Dummy | | −0.1740 | | −0.0144 |
| | | (0.3921) | | (0.2350) |
| Special Blend Dummy | | 5.3539** | | −1.3977** |
| | | (1.8660) | | (0.3833) |
| Controls | N | Y | N | Y |
| Brand Fixed Effects | N | N | Y | Y |
| Obs. | 1,365,820 | 1,365,820 | 1,365,820 | 1,365,820 |
| Brands | . | . | 466 | 466 |

Note: There are 1,365,820 sales of spirits. January is the omitted month dummy variable. Controls include month, year, channel, spirit type, feature, display, flavor and special blend dummy variables. Standard errors are clustered at the store chain level for all regressions. Standard errors are shown in parentheses *** p<0.01, ** p<0.05, * p<0.1. HHI bins constructed based upon HHI classifications found in the FTC's Merger Guidelines. The omitted bin for HHI is for highly concentrated markets (HHI > 2500).

**References for Chapter 1**

Aryal, Gaurab. "An empirical analysis of competitive nonlinear pricing." Available at SSRN 2262664 (2014).

Berry, Steven T. "Estimating discrete-choice models of product differentiation." The RAND Journal of Economics (1994): 242-262.

Buchanen, J. M. "The Theory of Monopolistic Quantity Discounts." Review of Economic Studies 20.3 (1953): 199-208.

Chao, Yong. "Strategic Effects Of Three Part Tariffs Under Oligopoly." International Economic Review 54.3 (2013): 977-1015.

Hee-Jung Choi, Michael K. Wohlgenant, Xiaoyong Zheng; Household-Level Welfare Effects of Organic Milk Introduction, American Journal of Agricultural Economics, Volume 95, Issue 4, 1 July 2013, Pages 10091028, https://doi.org/10.1093/ajae/aat021

Cohen, Andrew. "Package size and price discrimination in the paper towel market." International Journal of Industrial Organization 26.2 (2008): 502-516.

DeSalvo, Joseph S., and Mobinul Huq. "Introducing nonlinear pricing into consumer choice theory." The Journal of Economic Education 33.2 (2002): 166-179.

Dolan, Robert J. "Quantity discounts: Managerial issues and research opportunities." Marketing Science 6.1 (1987): 1-22.

Dupuit, Jules. "On tolls and transport charges." Annales des ponts et chausses. Vol. 11. No. 1962. 1849.

Escobari, Diego, and Paan Jindapon. "Price discrimination through refund contracts in airlines." International Journal of Industrial Organization 34 (2014): 1-8.

Friebel, Kerstin, Oleksandr Perekhozhuk, and Thomas Glauben. "Price Discrimination in Russian Wheat Exports: Evidence from Firm-level Data." Journal of Agricultural Economics (2015).

Hamilton, Jonathan H., and Jacques-Franois Thisse. "Nonlinear pricing in spatial oligopoly." Economic design 2.1 (1996): 379-397.

Hee-Jung Choi, Michael K. Wohlgenant, Xiaoyong Zheng, Household-Level Welfare Effects of Organic Milk Introduction, American Journal of Agricultural Economics, Volume 95, Issue 4, July 2013, Pages 10091028, https://doi.org/10.1093/ajae/aat021

Leslie, Phillip. "Price discrimination in Broadway theater." RAND Journal of Economics (2004): 520-541.

Lewis, Tracy R., and David EM Sappington. "Supplying information to facilitate price discrimination." International Economic Review (1994): 309-327.

Liu, Yizao and Shu Shen. "Price Discrimination with Asymmetric Firms: The Case of the U.S. Carbonated Soft Drinks Market." Working Paper (2012)

Manchester, Alden C., and Donald P. Blayney. Milk pricing in the United States. No. 33612. United States Department of Agriculture, Economic Research Service, 2001.

Maskin, Eric, and John Riley. "Monopoly with incomplete information." The RAND Journal of Economics 15.2 (1984): 171-196.

McManus, Brian. "Nonlinear pricing in an oligopoly market: The case of specialty coffee." The RAND Journal of Economics 38.2 (2007): 512-532.

Miller, Nathan H., and Matthew Osborne. "Spatial differentiation and price discrimination in the cement industry: evidence from a structural model." The RAND Journal of Economics 45.2 (2014): 221-247.

Miravete, Eugenio J., and Lars-Hendrik Roller. "Competitive non-linear pricing in duopoly equilibrium: the early US cellular telephone industry." (2003).

Miravete, E. "Are all those Calling Plans Really Necessary." The Limited Gains From Complex Tariffs (2004).

Miravete, Eugenio J. "Screening consumers through alternative pricing mechanisms." Journal of Regulatory Economics 9.2 (1996): 111-132.

Moffitt, Robert. "The econometrics of kinked budget constraints." The Journal of Economic Perspectives 4.2 (1990): 119-139.

Mussa, Michael, and Sherwin Rosen. "Monopoly and product quality." Journal of Economic theory 18.2 (1978): 301-317.

Nevo, Aviv. "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand." Journal of Economics & Management Strategy 9.4 (2000): 513-548.

Pigou, Arthur. "C, 1920." The economics of welfare (1920).

Reiss, Peter C., and Matthew W. White. Evaluating welfare with nonlinear prices. No. w12370. National Bureau of Economic Research, 2006.

Robinson, Joan. "The economics of imperfect competition." (1933): 519-521.

Sharkey, William W., and David S. Sibley. "Optimal non-linear pricing with regulatory preference over customer type." Journal of Public Economics 50.2 (1993): 197-229.

Kawaguchi, Tsunemasa, Nobuhiro Suzuki, and Harry M. Kaiser. "A spatial equilibrium model for imperfectly competitive milk markets." American Journal of Agricultural Economics 79.3 (1997): 851-859.

United States Department of Agriculture, National Agricultural Statistics Service. http://www.nass.usda.gov/

**References for Chapter 2**

Ackerberg, Daniel A. "Advertising, learning, and consumer choice in experience good markets: an empirical examination*." International Economic Review 44.3 (2003): 1007-1040.

Balvers, Ronald & Cosimano, Thomas. (1990). Actively Learning About Demand and the Dynamics of Price Adjustment. Economic Journal. 100. 882-98. 10.2307/ 2233664.

Brown, G.W., "Iterative Solutions of Games by Fictitious Play," in T.C. Koopmans, ed., Activity Analysis of Production and Allocation, Wiley, 1951.

Chatwin, Richard E., 2000. "Optimal dynamic pricing of perishable products with stochastic demand and a finite set of prices," European Journal of Operational Research, Elsevier, vol. 125(1), pages 149-174, August.

Ching, Andrew T. "Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration." International Journal of Industrial Organization 28.6 (2010): 619-638.

Ching, Andrew T., Tlin Erdem, and Michael P. Keane. "Invited paper-learning models: An assessment of progress, challenges, and new developments." Marketing Science 32.6 (2013): 913-938.

Cournot, Antoine Augustine, "Recherches sur les principes mathmatiquies de la thoric des richesses," Paris: Hachette, 1838.

Coscelli, Andrea, and Matthew Shum. "An empirical model of learning and patient spillovers in new drug entry." Journal of Econometrics 122.2 (2004): 213-246.

Crawford, Gregory S., and Matthew Shum. "Uncertainty and learning in pharmaceutical demand." Econometrica 73.4 (2005): 1137-1173.

Dolan, Robert J., and Abel P. Jeuland. "Experience curves and dynamic demand models: Implications for optimal pricing strategies." The Journal of Marketing (1981): 52-62.

Doraszelski, Ulrich, Gregory Lewis, and Ariel Pakes. "Just starting out: Learning and equilibrium in a new market." NBER w21996, (2016).

Erdem, T., M. Keane (1996) Decision-making under uncertainty: capturing dynamic brand choice processes in turbulent consumer goods markets. Marketing Science, 15(1): 1-20.

Escobari, Diego. "Dynamic Pricing, Advance Sales and Aggregate Demand Learning in Airlines." The Journal of Industrial Economics 60.4 (2012): 697-724.

Huang, Yufeng and Ellickson, Paul B. and Lovett, Mitch, Learning to Set Prices in the Washington State Liquor Market (October 8, 2018). Available at SSRN: https://ssrn.com/abstract=3267701 or http://dx.doi.org/10.2139/ssrn.3267701

Hitsch, Gunter J. "An empirical model of optimal dynamic product launch and exit under demand uncertainty." Marketing Science 25.1 (2006): 25-50.

McGoldrick, Peter J., and Helen J. Marks. "Price-size relationships and customer reactions to a limited unit-pricing programme." European Journal of Marketing 19.1 (1985): 47-64.

Miravete, Eugenio J., and Ignacio Palacios-Huerta. "Consumer inertia, choice dependence, and learning from experience in a repeated decision problem." Review of Economics and Statistics 96.3 (2014): 524-537.

Narayanan, Sridhar, Pradeep K. Chintagunta, and Eugenio J. Miravete. "The role of self selection, usage uncertainty and learning in the demand for local telephone service." Quantitative Marketing and Economics 5.1 (2007): 1-34.

Sen, Alper, and Alex X. Zhang. "Style goods pricing with demand learning." European Journal of Operational Research 196.3 (2009): 1058-1075.

Seo, Boyoung, Firm Scope and the Value of One-Stop Shopping in Washington State's Deregulated Liquor Market (April 21, 2019). Kelley School of Business Research Paper No. 16-70.

Washington Beer and Wine Distributors Association, www.wbwwa.com.

**Vita**

# Benjamin E. Wallace

---

## Education

| | |
|---|---|
| **M.S. Economics**, University of Kentucky, Lexington, KY | 2014 |
| **B.A. Economics**, University of Michigan-Flint, Flint, MI | 2012 |
| **B.S. Actuarial Mathematics**, University of Michigan-Flint, Flint, MI | 2012 |

## Professional Experience

**Visiting Assistant Professor**

| | |
|---|---|
| Roanoke College, Economics Department | 2018 |

**Research Assistant**

| | |
|---|---|
| University of Kentucky, Economics Department | Summers of 2015 & 2016 |
| Robert Wood Johnson Foundation | 2015 |

## Teaching Experience

**Primary Instructor**

| | | |
|---|---|---|
| **UK** | ECO 401, Intermediate Microeconomic Theory | (1 Section) |
| **UK** | ECO 391, Economic and Business Statistics | (8 Sections) |
| **UK** | ECO 201, Principles of Microeconomics | (3 Sections) |
| **UK** | ECO 101, Contemporary Eco Issues | (2 Sections) |
| **UK** | Actuary Exam 1/P Preparation Seminar | (1 Section) |
| **RC** | ECON 122, Principles of Macroeconomics | (1 Section) |
| **RC** | ECON 121, Principles of Microeconomics | (2 Sections) |

**Teaching Assistant**

ECO 610, Managerial Economics
ECO 461, Market Structure and Anti-Trust Policy
ECO 391, Economic and Business Statistics
ECO 327, Strategic Decision Marking: An Introduction to Game Theory

## Awards and Certificates

| | |
|---|---|
| Gatton Fellowship, University of Kentucky, | Fall 2014, Fall 2015, Fall 2016 |
| Max Steckler Fellowship, University of Kentucky, | Fall 2013 |
| Teaching Assistantship, | 2013-2014, 2014-2015, 2015-2016, 2016-2017 |

## Presentations

"Nonlinear Pricing: Evidence of Price Discrimination in the Fluid Milk Market"
– International Industrial Organization Conference, Indianapolis, Indiana, 2018
– Centre College, Economics Department, Danville, Kentucky, 2017
– Kentucky Economics Association Conference, University of Kentucky, Lexington, KY, 2016

University of Kentucky Graduate Student Workshop, Lexington, Kentucky, 2017
– Overview of Nielsen Datasets
– Introduction to Matlab