Title
Essays on Consumer Purchase Behavior and Competitive Firm Strategies

Author
Chen, Andy Wei-Rong

Publication Date
2016

Peer reviewed|Thesis/dissertation

Essays on Consumer Purchase Behavior and Competitive Firm Strategies

By

Andy Wei-Rong Chen

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Przemyslaw Jeziorski, Chair

Professor Ganesh Iyer

Professor Ben Handel

Spring 2016

Essays on Consumer Purchase Behavior and Competitive Firm Strategies

Abstract

Essays on Consumer Purchase Behavior and Competitive Firm Strategies

by

Andy Wei-Rong Chen

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Przemyslaw Jeziorski, Chair

The two essays explore two topics in marketing - consumer purchase behavior and competitive firm strategies. The first essay examines consumer stockpiling behavior in the retail gasoline market and aims to shed light on what factors affect consumer stockpiling. Past research on consumer stockpiling behavior such as Hendel and Nevo (2006) finds evidence of inter-temporal substitution by consumers that implies stockpiling behavior. However, they do not observe actual inventory or consumption and have to rely on simplifying assumptions about these quantities. I collect a novel data set of gasoline purchase history of consumers with actual inventory and consumption to test several hypotheses that relate consumer stockpiling to price, duration between purchases, and consumption. First, I find that consumers holding more inventory are more price sensitive. Higher inventory increases the impact of price on purchase decisions and those with higher inventory can afford to do more price search before making a purchase. I also find that all else equal, consumers will reduce consumption following a purchase made during high prices; that consumers with lower inventory have a higher probability of purchasing; and that duration from previous purchase is shorter for purchases made during low prices and longer during high prices.

The second essay examines competition in a dynamic setting between Wal-Mart and Target in the context of location choices and expansion strategies. Studying this topic sheds light on how an industry evolves and how the market structure is shaped by decisions such as when to enter and exit and where to locate new stores as well as the driving force behind different expansion strategies. One of the early papers by Bresnahan and Reiss (1991) study entry of retail and professional services into isolated markets in a one-shot game. In their model, firms are allowed to enter once and open one store. Collard-Wexler (2014) estimates a model of investment and entry in the ready-mix concrete industry. Again, each firm is assumed to own a single plant in the model. In my paper, I build a structural model in which firms can open multiple stores over a period of time. This setting is closer to reality as competition between firms lasts over many periods with firms making regular entry decisions and often opening multiple stores in

the same market. I estimate how entry decisions are affected by competitor's presence and market characteristics and learn about the evolution of market structure and expansion strategies of the firms. The firms are forward-looking and engage in Markov perfect equilibrium (MPE) strategies. The results show that firm profits are affected asymmetrically by the competitor's presence. First, Wal-Mart is the dominant firm with higher profits regardless of Target's presence. On the other hand, Target's profits depend a lot on Wal-Mart's presence. It must have more stores than Wal-Mart (called store advantage) to profit. I also find asymmetry in the expansion strategies of the two firms. Wal-Mart tends to explore new and smaller markets by being the first and often the only firm to enter, while Target tends to focus on major markets with high population and GDP and strives to maintain store advantage over Wal-Mart by matching or outdoing Wal-Mart's decision to open new stores.

Contents

List of Figures

List of Tables

Acknowledgements

The completion of the dissertation is by no means the end of the academic journey. It is simply the end of one stage shortly followed by the beginning of another. Nonetheless, it is a milestone worth mentioning and thanking the people who have helped me along way.

First, I would like to express sincere gratitude to my dissertation chair Professor Przemek Jeziorski for his guidance, encouragement, and motivation throughout the dissertation and Ph.D. program at UC Berkeley. I would also like to thank other members of the dissertation committee, Professor Ganesh Iyer and Professor Ben Handel, for their advice, patience, and support throughout the Ph.D. program. I also appreciate the comments and support from other faculty members in the Marketing Group at UC Berkeley – Professors Miguel Villas-Boas, Zsolt Katona, Ming Hsu, Yuichiro Kamada, Clayton Critcher, Leif Nelson, and Ellen Evers. Also, many thanks to the Ph.D. program office (Melissa Hacker, Bradley Jong, and Kim Guilfoyle) for their help and kind advice over the years.

Finally, I would like to express my deepest gratitude to my parents for providing advice and unreserved support throughout my years of study. This accomplishment would not have been possible without them. Thank you.

Chapter 1

Consumer Stockpiling Behavior in the Retail Gasoline Market

1.1 Introduction

It is a common observation that consumers tend to purchase more storable goods when the prices drop temporarily. However, it is often not clear whether the increase during low prices is due to immediate consumption or stockpiling behavior (inter-temporal substitution). Consumers, on one hand, may experience consumption shocks and therefore need to increase the amount of purchase immediately. On the other hand, they may anticipate higher prices in the future and decide to take advantage of current low prices. It has been shown in past research that significant relationships exist between inventory of goods held by consumers, consumption, purchase amount and probability in markets such as coffee, yogurt, ketchup, ice cream, soft drinks, and laundry detergents (Chiang (1991), Ailawadi and Neslin (1998), Bell and Boztug (2004), Erdem, Imai, and Keane (2003), and Hendel and Nevo (2006)). In the retail gasoline literature, despite that many topics have been explored (such as gasoline price cycles, price gouging, price dispersions, and gasoline market structures), consumer stockpiling behavior has not been explored in depth.

In summary, several key differences distinguish this paper and existing literature on consumer stockpiling behavior. First, among existing literature in consumer stockpiling behavior in the retail gasoline market, prior work is limited and has not explored the direct relationship between purchase and price without using proxies. Prior work has used proxies such as consumer search for gasoline prices to investigate how it relates to price changes. This could be due to lack of data on actual inventory consumers hold at the time of purchase. In contrast, my paper directly tests the relationship between purchase, inventory, and price in the gasoline market. Second, prior work on the retail gasoline market has not explored how past purchases influence the current one. I investigate the relationship between past purchases on the current purchase and whether purchase history causes consumer purchase behavior to change over time. Third, existing literature on consumer stockpiling behavior in general has not used actual data on consumer inventory. My paper is one of the first to use such data. I will explore questions such as whether consumption and purchase probability change during low prices; how significant consumer stockpiling behavior is when consumers make purchase decisions; and how different kinds of consumers respond to price changes.

In this paper I test the following hypotheses: All else equal, consumers with lower current inventory have a higher probability of purchase and buy more gasoline per purchase; all else equal, duration from the previous purchase is shorter for purchases during low prices, and longer during high prices; and all else equal, consumers will reduce consumption following a purchase during high prices and increase consumption following purchases during low prices.

The results confirm the hypotheses. First, consumers who have lower inventory of gasoline in the gas tanks have a higher probability of purchasing and buy more gasoline each purchase. Also, purchases made during low prices are followed by shorter periods of time until the next purchase compared to purchases made during high prices. Moreover, consumption tends to decrease following purchases made during high prices. In addition, consumers with more inventory are more price sensitive than those with lower inventory.

This paper is organized as follows. Section 1.2 provides a literature review of static and dynamic models in consumer stockpiling literature and an overview of the literature in the retail gasoline industry in general. Section 1.3 discusses the theories and hypotheses tested in this paper. Section 1.4 describes the data used in this paper to test the hypotheses. Section 1.5 describes the theoretical and empirical models used, and Section 1.6 provides the conclusion and directions for future research.

1.2 Literature Review

The early papers studying sales of storable goods focus on static demand estimation which excludes dynamic effects due to inter-temporal substitution. One example is the paper by Bresnahan (1987) who estimates the demand in the car market. Another is the classic paper by Berry, Levinsohn, and Pakes (1995) who use a static demand estimation model with random coefficients to capture idiosyncratic consumer tastes for the different products. These models have become popular for static demand estimation and form the foundation of many papers in the sales and demand estimation literature. However, they are not sufficient in the presence of dynamic effects when consumers are forward-looking and may purchase when there is a price reduction or when they anticipate a price surge in the future. Some notable papers discussing static demand models include the ones by Gupta (1988), Chiang (1991), Ailawadi and Neslin (1998), and Bell and Boztug (2004).

Gupta (1988) presents a static model exploring the effectiveness of sales promotions by decomposing the sales "bump" during the promotion period into sales increase due to brand switching, purchase time acceleration, and stockpiling. He finds that more than 84% of the sales increase are due to brand switching, 14% are due to purchase acceleration, and less than 2% are due to stockpiling. Chiang (1991) presents a static model that simultaneously examines consumer purchase decisions of whether to buy, what to buy, and how much to buy in the context of the coffee market. The idea is to deduce an unobservable threshold price about which a consumer switches from non-purchase to purchase by solving a consumer utility maximization problem. The paper's main drawback is that it cannot handle situations which have an inherently dynamic nature. Such issues may arise from consumer stockpiling behavior due to sudden price changes, promotion, and income expectation.

Ailawadi and Neslin (1998) use a static model to empirically demonstrate the existence of flexible consumption rates in packaged-goods products (yogurt and ketchup), how this phenomenon can be modeled, and its importance in assessing the effectiveness of sales promotions. They specify an incidence, choice, and quantity model in which category consumption varies with the level of household inventory. Like other papers mentioned above, the authors do not observe inventory or consumption rate, so they have to make assumptions about inventory level and consumption. One way is to let the starting inventory for each household equal to the average weekly consumption level of the household. The total volume of product purchased by the household over the duration of the period is then divided by the number of days in the period. Bell and Boztug (2004) propose a novel static model to study the effects of inventory on the probability of purchase. They estimate inventory using standard inventory calculation formula plus an "inventory estimation bias." The estimated thresholds are plausible across multiple food categories, with hot dogs, ice cream and soft drinks showing the largest effects. They show that their model, calibrated on ten product categories, fits better than the standard nested logit model. However, the static results should be viewed with caution as with other papers above because of potential bias shown in the Hendel and Nevo (2006) paper discussed below.

In the 2000's, dynamic models are developed to include dynamic effects due to consumer stockpiling behavior. Papers by Erdem, Imai, and Keane (2003) and Hendel and Nevo (2006) are good examples. Erdem, Imai, Keane develop a model of household demand for frequently purchased consumer goods that are branded, storable, and subject to stochastic price fluctuations. They account for how inventories and expectations of future prices affect current period purchase decisions. The model is estimated using scanner data for the ketchup category. One main result is that price expectations and the nature of the price process have important effects on demand elasticities. Specifically, long-run cross price elasticities of demand are more than twice as great as short-run cross price elasticities. The model can be used to predict how consumer purchase decisions respond to changes in the entire retail pricing process such as a shift from high/low pricing to everyday low pricing. Like Hendel and Nevo, the authors do not observe inventory or consumption. The authors assume that in each period, barring a stock out, the usage rate does not depend on the inventory level. They further assume that in each period, households use each brand in their inventory proportionately to meet their usage needs. Hendel and Nevo (2006) find evidence of inter-temporal substitution by consumers, implying that consumers exhibit stockpiling behavior by buying more when prices are low and substituting away from purchases when prices are high. They also show that static estimates of long-run price elasticities are biased.

In a more recent paper, Ching, Erdem, and Keane (2014) present an estimation method allowing them to estimate dynamic models without solving a dynamic programming problem. Their paper also relaxes strong assumptions about how consumers form expectations about the future and consequently reduce the computational burden. The reduced computational requirement allows them to include learning in their model, thus combining the two mainstream of consumer dynamic choice models in learning and stockpiling (inventory) behavior. The authors find that learning, strategic trial, inventories and category consideration are all important factors in consumer purchase decisions. The market they study is the diaper market.

The existing literature studying consumer stockpiling in the retail gasoline market has been limited. One occasional example is a recent paper by Byrne and Roos (2015) who explore cross-sectional and inter-temporal consumer search behavior in the gasoline market and test for stockpiling behavior by measuring the relationship between consumer search and the timing before and after price jumps. The authors obtain data on the number of clicks on a website maintained by the Australian government that tracks and displays countrywide gasoline prices in Australia. Their hypothesis is that consumers anticipate future sales, and therefore engage in inter-temporal search for storable goods, where stockpiling is possible. They find evidence of both cross-sectional and inter-temporal consumer search behavior as well as stockpiling behavior. The results show that when price-dispersion is high at a specific point in time consumer search would increase. They also show that consumer search increases during days leading up to price jumps, and decreases during days following price jumps. This active inter-temporal search behavior presents evidence for stockpiling behavior.

Nonetheless, prior work has explored a wide range of other issues in the retail gasoline market. For example, it has been found that retail gasoline prices display a phenomenon known as the "rock and feather" pricing, where prices rise fast when costs are high, but drop slowly when costs are low (Borenstein, Cameron, and Gilbert (1997), Bachmeier and Grin (2003)). It has also been found that gas stations tend to undercut one another by small amounts frequently and consistently until retail prices fall close to cost. Then one firm raises its price substantially, other firms follow within a few days or even a few hours, and another round of undercutting begins (Zimmerman, Yun, and Taylor (2013), Lewis and Noel (2011)). Many papers have explored the factors that affect price dispersion and price elasticities. Lewis (2009) finds that the extent of price dispersion is related to the density of local competition, and the relationship varies depending on the type of sellers in the same region. Other researchers have found that greater station density leads to lower prices (Shepard (1993), Sen (2005), Cooper and Jones (2007), Houde (2012)). Prior work also explores the retail gasoline market structure (Zimmerman (2012)) and vertical relationships between manufacturers and retailers (Shepard (1993), Slade (1998)). There have also been a handful of papers exploring the laws that govern competition in the retail gasoline market. For example, it has been found that laws prohibiting retail gasoline stations from pricing either below cost, or a positive minimum-margin price tend to increase prices (Anderson and Johnson (1999), Brannon (2003)). Despite the amount of existing literature on a variety of issues in the retail gasoline market, the study of consumer stockpiling and the relationship between stockpiling, price, and purchase behavior in the retail gasoline market has been limited. My paper attempts to provide some insights to this area of literature.

The papers in consumer stockpiling literature mentioned above study a variety of markets. For example, Erdem, Imai, and Keane's paper studies the ketchup market and Hendel and Nevo study the detergent market. They use data readily available from the consulting firms such as A.C. Nielsen. The data is detailed and convenient to obtain. However, it does not include data on consumer inventory. Prior work has had to infer consumer inventory from assumptions such as a constant consumption rate. In contrast, I use a unique data set that includes price paid, quantity sold, and consumer inventory in the retail gasoline market. The data set was compiled from several resources on the Internet. Another benefit of studying the retail gasoline market is that it potentially allows for a more comprehensive coverage of the market price as retail gasoline prices have been made publicly available by consumers and retailers with precision and frequency.

As mentioned in the introduction, in relation to past research, this paper tests the relationship between purchase, inventory, and price in the gasoline market without using proxies such as consumer search for gasoline prices. Also, I explore the relationship between past and current purchases in the gasoline market with actual data on consumer inventory. The next section discusses the hypotheses I will test in the paper.

1.3 Theory and Hypotheses

1.3.1 Probability of Purchase and Inventory

Prior research has shown that in theory, in a dynamic market with forward-looking consumers, quantity purchased and probability of purchase decline in the current inventory (Hendel and Nevo, 2006). The intuition is that consumers have a target inventory they want to reach every period. This target inventory is independent of the current inventory. Therefore, the higher the current inventory, the smaller purchase is required to reach the target inventory level. However, Hendel and Nevo cannot directly test this hypothesis without making assumptions because they, like most researchers studying inventory and stockpiling, do not observe the exact inventories consumers hold. With the data I have, I can test this hypothesis directly:

H1: All else equal, consumers with lower current inventory have a higher probability of purchase and buy more gasoline per purchase.

1.3.2 Purchase Frequency and Inventory

A hypothesis that follows is concerned with the frequency of purchase, or duration between purchases. The theory proposed by Hendel and Nevo (2006) assumes that the target inventory level S and threshold inventory level s that triggers purchase both decline in price. Therefore, during low prices, the target inventory (or end-of-period inventory) level will be higher. Consumers purchase more during low prices and this should allow them to last longer until the next purchase, all else equal. In addition, during low prices, the inventory that triggers purchase should also be higher. Consumers purchase while holding a higher inventory during low prices, meaning that a shorter time has passed since the last purchase. Thus I will test the following hypothesis:

H2: All else equal, duration from the previous purchase is shorter for purchases during low prices, and longer during high prices.

1.3.3 Stockpiling, Price, and Consumption

The next test is to explore whether consumption of gasoline changes based on previous purchases. In other words, do consumers adjust consumption after getting a good or bad deal? Prior work in the consumer behavior literature has found that consumers often set budgets and track expenses against their budget (Heath and Soll, 1996). They show that people may under-consume or over-consume because of unexpected consumption opportunities or constraints, most likely impacted by changes in prices. This phenomenon is called mental budgeting by Heath and Soll. Much existing work in the finance literature has also explored similar issues in the context of how tax policies affect consumer spending. Poterba (1988) finds that temporary tax policies affect consumption rates. For example, their results suggest that a $1 increase in transitory income resulting from the 1975 tax rebate increases spending by about 20 cents. Browning and Collado (2001) and Parker (1999) also find similar results. In the context of the gasoline purchases, I will test the following hypothesis:

H3: All else equal, consumers will reduce consumption following purchase during high prices and increase consumption following purchases during low prices.

1.4 Data

I compile a unique data set using purchase data from a website that collects gasoline purchases of drivers in the world as well as car manufacturer websites that offer specifications of vehicles and government websites that make public retail gasoline prices. First, the data on gasoline purchases is obtained from Fuelly, a website that allows users from the world to track gasoline purchases. Users enter data such as the date of the purchase, number of gallons purchased, price paid, odometer reading or miles driven since last purchase, location (state and/or city), car brand, car model, year the car was made, and special notes. Miles per gallon (MPG) over time is generated automatically and displayed publicly for users and others interested in comparing MPG within specific regions or specific type of cars or across users. The data is self-reported, so extensive checking was done to remove data points that are obviously not reasonable such as entries with extremely high prices, gallons bought, and miles driven.

Figure 1.1: A User's Fuelly Profile

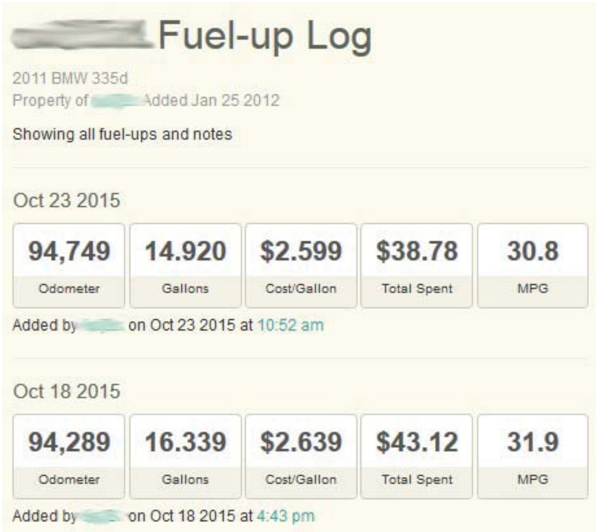


Figure 1.2: A User's Log of Gasoline Purchases

Variables that are mandatory for the users to enter are date of the purchase, number of gallons purchased, price paid, odometer reading or miles driven since last purchase, car brand, car model, and car year. A summary of the purchase variables is shown on each user's profile page. The website also provides a log of each user's gasoline purchases. Location of the user is not mandatory, and as a result around 40% of the users entered the state they reside in. The users without state information are not included in the sample. To check for under-reporting due to purchases that are not reported, I calculate the MPG between purchases using the miles driven and gallons bought and compare them with the MPG specified by the manufacturer. Any user with at least one MPG between purchases greater or less than 1.5 times the MPG specified by the manufacturer was removed. Moreover, the website was launched in 2008, so any user with any entry dated before 2008 are not in the sample as they show a tendency to misreport. I try my best to include in the sample only users with the most extensive information provided and show no tendency to misreport.

Variable	Description
Days between Purchases	Number of days passed since the last gasoline purchase
Price Paid ($/Gallon)	The actual cost per gallon of gasoline paid by the user
State Weekly Average Price ($/Gallon)	The average cost per gallon of gasoline in the state the user is registered in
Inventory at Purchase (% Full)	The percent of gasoline still remaining in the tank at the time of purchase
Consumption (Gallons)	Number of gallons consumed on a particular day
Gallons Purchased	Number of gallons purchased
Miles Driven between Purchases	Number of miles driven since the last gasoline purchase
Tank Capacity (Gallons)	Number of gallons the car's tank can hold
Car Age (Years)	Number of years since the car was made
MSRP ($)	The manufacturer's suggested retail price of the car
State	The state the user is registered in
Car Type	The type of car the user owns
Car Brand	The manufacturer of the car
Purchase	Dummy for purchase (Equals 1 if a purchase is made on a particular day and 0 otherwise)

Table 1.1: Description of Variables

Extra information about the cars in the sample are obtained from various online sources. For example, most of the MPG specified by the manufacturer are obtained from the U.S. Department of Energy. The same website also provides information on car categories defined by passenger and cargo volume known as the EPA size class. There are ten categories of cars in the sample: compact car, large car, mid-size car, minivan, pickup truck, sports utility vehicle (SUV), station wagon, subcompact car, two-seater, and van. Information on tank capacity and the type of fuel (regular, mid-grade, premium) that should be used as specified by the manufacturer is also obtained from the same source. Tank capacity is the key data that allows me to calculate the inventory level when a purchase is made. This feature is what essentially distinguishes this paper from previous work which usually simulates or estimates inventory and consumption.

State	Frequency	Percent
California	189	34.18
Colorado	46	8.32
Florida	68	12.30
Massachusetts	19	3.44
Minnesota	18	3.25
New York	38	6.87
Ohio	47	8.50
Texas	85	15.37
Washington	43	7.78

Table 1.2: Frequency of Each State

Car Type	Frequency	Percent
Compact Car	157	28.39
Large Car	12	2.17
Midsize Car	100	18.08
Minivan	11	1.99
Pickup Truck	79	14.29
SUV	110	19.89
Station Wagon	28	5.06
Subcompact Car	35	6.33
Two Seater	17	3.07
Van	4	0.72

Table 1.3: Frequency of Each Car Type

Finally, data on retail gasoline prices is obtained from U.S. Energy Information Administration. The data contains average weekly retail gasoline prices by type (regular, mid-grade, premium) at the country, state, and city level from 2000 to present. One restriction is that state-level data is only available for nine states (California, Colorado, Florida, Massachusetts, Minnesota, New York, Ohio, Texas, and Washington) and a number of cities. As a result, only users from these states are included in the sample. There is definitely data on retail gasoline prices in all the states and great number of cities, but the only source I know for now charges a hefty price. The data used for this paper is currently only for cars in the above nine states. Finally data on MSRP (manufacturer's suggested retail price) is obtained from various sources of the Internet such ad www.edmunds.com and www.cars.com. The MSRP's are adjusted to the 2015 value using a car price index.

Variable	Observation	Mean	SD	Min	Max
Number of Purchases Per Car	553	8.40	6.67	1	44
Days between Purchases	4095	8.89	10.00	0	132
Price Paid ($/Gallon)	4647	3.29	0.55	0.1	4.8
State Weekly Average Price ($/Gallon)	4647	3.32	0.48	1.86	4.25
Inventory at Purchase (% Full)	4647	0.29	0.17	0	0.99
Gallons Purchased	4647	12.09	4.41	0.01	34.16
Miles Driven between Purchases	4094	338.66	205.68	10	2755

Table 1.4: Summary of Gasoline Purchase Related Variables

The final sample contains 553 cars in nine states. In total there are 4694 purchases and 42195 observations. Each observation is a decision on whether to purchase made by a car owner on a particular day. As described before, there are ten car types, which are broken down into 37 car brands and 197 car models manufactured in 30 different years. Table 1.2 shows the frequency of the state and Table 1.3 shows the frequency of car type respectively. Figure 1.3 shows the distribution of the tank capacity; Figure 1.4 shows the distribution of the car age; and Figure 1.5 shows the distribution of the MSRP. Table 1.4 shows the summary variables related to gasoline purchases.

	Observation	Mean	SD	Min	Max
Tank Capacity	553	17.30	5.25	8.5	42
Car Age	553	5.01	6.06	0	35
MSRP	532	23458.34	9923.32	9340	79000

Table 1.5: Summary of Car Attributes

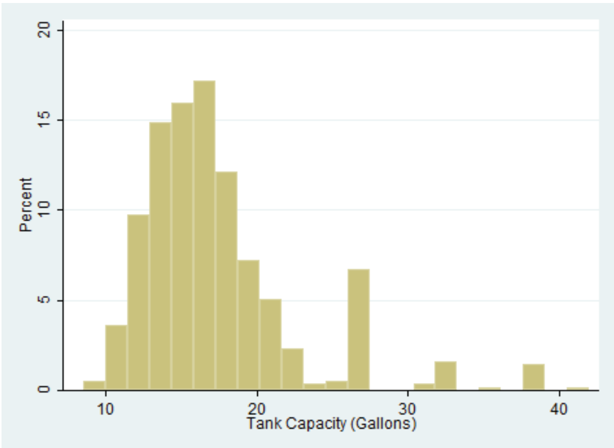


Figure 1.3: Distribution of Tank Capacity

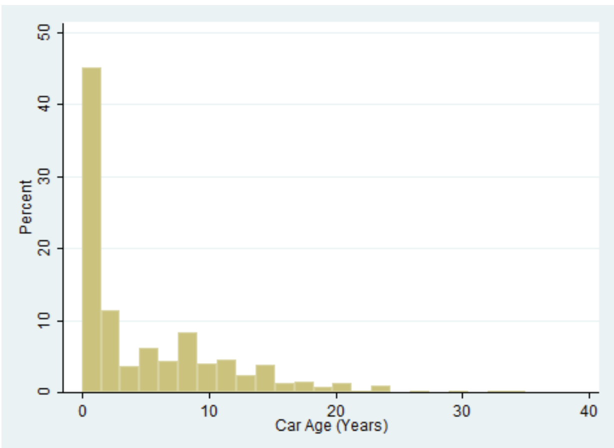


Figure 1.4: Distribution of Car Age

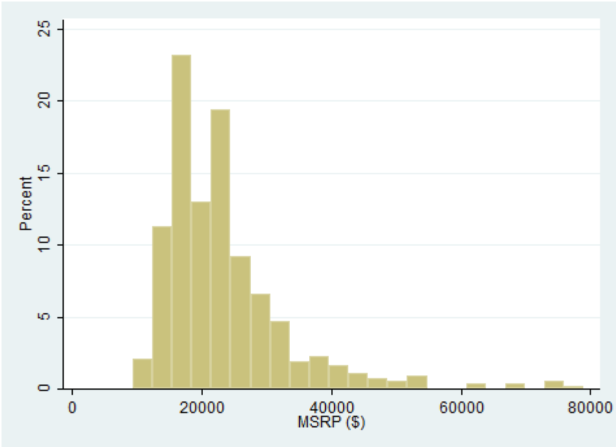


Figure 1.5: Distribution of MSRP

1.5 Model

1.5.1 Theoretical Model

First I present a theoretical model of consumer utility. Suppose at the beginning of period t, consumer i's utility is

$$U_{it} = \gamma \ln c_{it} - \alpha p_t x_{it} - \delta \ln(\bar{I}_i - i_t)$$

subject to constraints

$$i_{t+1} = i_t + x_{it} - c_{it}$$

where c_{it} is the number of gallons consumed by consumer i in period t and has a log functional form, p_t is the price paid by the consumer in period t, $\bar{I}_i$ is the target inventory consumer i wants to reach by the end of period t, and x_{it} is the number of gallons purchased in period t. Each period, consumers want to maximize the utility by choosing how much to consume and how much to buy, which will determine the end-of-period inventory. The Lagrangian of the above maximization problem is

$$L = \gamma \ln c_{it} - \alpha p_t x_{it} - \delta \ln(\bar{I}_i - i_t) + \lambda(i_t + x_{it} - c_{it} - i_{t+1})$$

Differentiating with respect to the decision variables c_{it} and x_{it} yields

$$
\begin{aligned}
\frac{\partial L}{\partial c_{it}} &= \frac{\gamma}{c_{it}} - \frac{\delta}{\bar{I}_i - i_t}\left(\frac{\partial\left(\bar{I}_i - i_t\right)}{\partial c_{it}}\right) - \lambda \\
&= \frac{\gamma}{c_{it}} - \frac{\delta}{\bar{I}_i - i_t} - \lambda \\
&= 0 \\
\frac{\partial L}{\partial x_{it}} &= -\alpha p_t + \frac{\delta}{\bar{I}_i - i_t}\left(\frac{\partial\left(\bar{I}_i - i_t\right)}{\partial c_{it}}\right) + \lambda \\
&= -\alpha p_t + \frac{\delta}{\bar{I}_i - i_t} + \lambda \\
&= 0
\end{aligned}
$$

Substituting $\lambda = \alpha p_t - \frac{\delta}{\bar{I}_i - i_t}$ into the first equation yields optimal consumption $c_{it}^* = \frac{\gamma}{\alpha p_t}$. Basic economic theory suggests that γ and α are positive because more consumption provides more utility while higher price decreases utility. It follows that $\frac{\partial c_{it}^*}{\partial p_t} = -\frac{\gamma}{\alpha p_{it}^2} < 0$; consumption c_{it} decreases with price p_t. The higher the price, the less consumption in that period is for consumers. It follows that inventory i_t will be higher and that purchase amount x_{it} will be smaller, especially when there is a capacity constraint on the car tank.

1.5.2 Empirical Model - Probability of Purchase and Inventory

To test Hypothesis 1 concerning the relationship between purchase probability, purchase quantity, and inventory, I estimate the following logit model using the panel data I compiled:

$$
\text{Dependent variable } y_{it} = \begin{cases} 1 & \text{if a purchase is made by user } i \text{at time } t \\ 0 & \text{otherwise} \end{cases}
$$

Let p_{it} be the probability that $y_{it} = 1$, then

$$
\begin{aligned}
\ln\left(\frac{p_{it}}{1-p_{it}}\right) \quad = \quad & \beta_0 + \beta_1 Price_{it} + \beta_2 Inventory_{it} + \beta_3 Inventory_{it} \times Price_{it} \\
& + \beta_4 Consumption_{it} + \beta_5 Consumption_{it} \times Price_{it} \\
& + \beta_6 MPG_{it} y_i + \beta_7 MSRP_i + \beta_8 TankCapacity \\
& + \sum_{r=1}^{R-1} \beta_{9,r} State_i + \sum_{s=1}^{S-1} \beta_{10,s} CarMake_i \\
& + \sum_{t=1}^{T-1} \beta_{11,t} CarType_i + \sum_{u=1}^{U-1} \beta_{12,u} CarAge_{it} + \epsilon_{it}
\end{aligned}
$$

where an observation is a decision made each day by the consumer on whether to purchase gasoline or not. The independent variables include the actual price paid by the driver for each gasoline purchase, the inventory the user holds when the purchase is made, and consumption or the amount of gasoline used on a particular day. Other controls included in the model include the state the user is registered in, the car's efficiency since the last purchase measured in miles per gallon, plus time-invariant car attributes such as MSRP, tank capacity, make or manufacturer, type, and age. R, S, T, and U are the number of states, car makes, car types, and car ages in the data respectively.

Before estimating the model, some exploratory analysis of the relationships between price paid, average state price, and the difference between these two values against inventory level at purchase are plotted in Figures 1.6 and 1.7. Figure 1.6 shows that the percentage difference between the average state price and actual price paid by consumers decrease in inventory level. Consumers with higher inventory buy at prices much lower than the average price relative to consumers holding low inventory. The reason could be that those who hold low inventory tend to buy gasoline closer to the average price because their immediate need for gasoline does not allow them to wait longer or price search. Figure 1.7 plots the average price paid and average state price against the inventory level. Figure 1.9 and Table 1.6 show that there is a significant portion of consumers who fill up when they still hold relatively high inventory. Half the consumers purchase gasoline when their gasoline tank is more than 25% full. Also, 25% of the consumers purchase gas when their tank is more than 36% full. It is likely that these consumers with high inventory levels engage in price searching because they are more price sensitive and want to take advantage of price cuts. Figure 1.8 shows an approximately negative linear relationship between the number of gallons purchased as inventory increases. The exploratory analysis shows that hypothesis 1 is plausible.

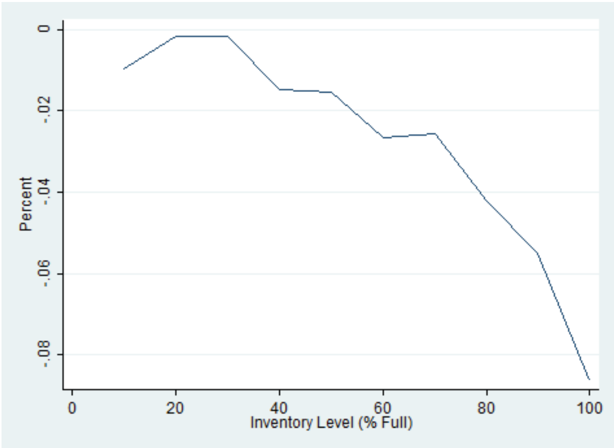


Figure 1.6: Percentage Difference Between Average Price Paid and State Price vs. Inventory

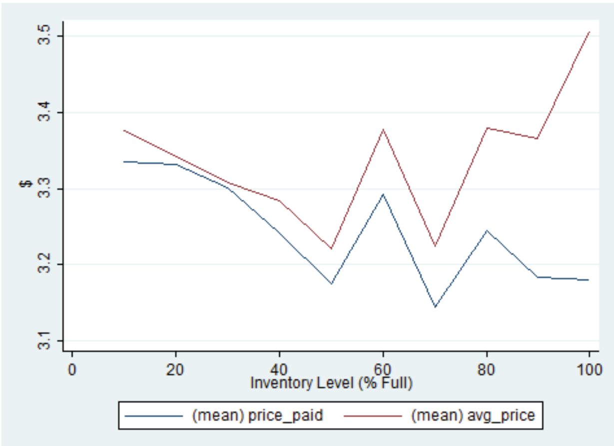


Figure 1.7: Average Price Paid and Average State Price vs. Inventory

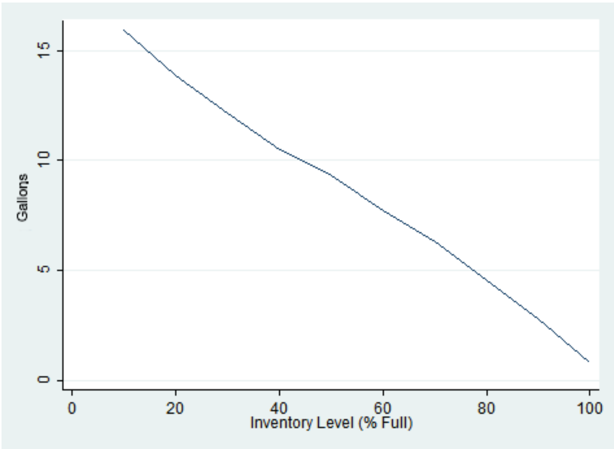


Figure 1.8: Gallons Purchased vs. Inventory Level

The estimates of the logit model are presented in Table 1.7. The parameters of interest for testing Hypotheses 1 are β_1, β_2, and β_3, all of which are highly statistically significant with p-values at least as small as 0.01. The results show the expected negative sign for price which indicates that higher price reduces the probability of buying. Inventory also has the expected negative sign; people with higher inventory have lower probability of buying. The interaction term between inventory and price is also negative, indicating that people with higher inventory are more price sensitive. In other words, higher inventory increases the impact of price on the probability of purchase. Regarding the control variables, most of the state dummies are statistically significant and have positive signs, indicating that people in states such as Colorado, Florida, Massachusetts, Minnesota, and Ohio are more likely to make a purchase than people in California (the base category), all else equal. Most of the estimates for car make and car type are not statistically significant. For each model, the clustered standard errors by state are also shown to account for the possibility that the standard errors may be correlated within each state. The clustered standard errors are greater and increase the p-values of the estimates. However, most estimates are still significant or marginally significant at the 0.1 or smaller significance levels.

Percentile	Inventory Level
10%	0.066
25%	0.161
50%	0.252
75%	0.367
90%	0.538
95%	0.641

Table 1.6: Percentiles of Inventory Level at Purchase

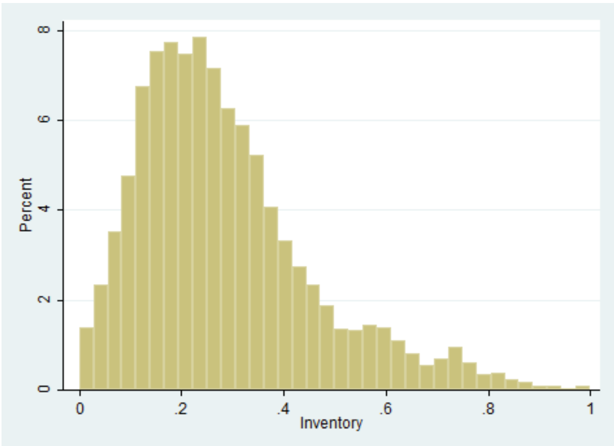


Figure 1.9: Distribution of Inventory Level at Purchase

Table 1.7: Logit Model of Purchase Probability

	(1) Purchase		(2) Purchase	
Purchase				
Price Paid	-0.421***	(0.0853)	-0.421*	(0.251)
Inventory	-2.470***	(0.807)	-2.470+	(1.535)
Inventory x Price	-1.470***	(0.249)	-1.470***	(0.510)
Consumption	12.55***	(2.539)	12.55**	(5.396)
Consumption x Price	0.0192	(0.774)	0.0192	(1.801)
Tank Capacity	0.0234**	(0.0105)	0.0234	(0.0242)
MSRP	-0.0000190***	(0.00000473)	-0.0000190	(0.0000125)

MPG	-0.00654***	(0.00124)	-0.00654***	(0.00224)
State - Colorado	0.290***	(0.0875)	0.290***	(0.0857)
State - Florida	0.355***	(0.0821)	0.355***	(0.115)
State - Massachusetts	0.553***	(0.129)	0.553***	(0.0370)
State - Minnesota	0.645***	(0.147)	0.645***	(0.0732)
State - New York	-0.000591	(0.0976)	-0.000591	(0.0901)
State - Ohio	0.270***	(0.0897)	0.270***	(0.0883)
State - Texas	0.0332	(0.0808)	0.0332	(0.103)
State - Washington	0.0501	(0.109)	0.0501	(0.166)
Car Make - Audi	0.0299	(0.194)	0.0299	(0.332)
Car Make - BMW	-1.230***	(0.289)	-1.230***	(0.401)
Car Make - Buick	0.150	(0.212)	0.150	(0.636)
Car Make - Cadillac	-0.485+	(0.296)	-0.485	(0.690)
Car Make - Chevrolet	-0.503***	(0.175)	-0.503	(0.401)
Car Make - Chrysler	-0.837**	(0.364)	-0.837	(0.582)
Car Make - Dodge	-1.184***	(0.222)	-1.184***	(0.361)
Car Make - Fiat	-0.570	(0.697)	-0.570	(0.389)
Car Make - Ford	0.305*	(0.159)	0.305	(0.475)
Car Make - GMC	0.266	(0.372)	0.266	(0.889)
Car Make - Honda	-0.0108	(0.166)	-0.0108	(0.445)
Car Make - Hyundai	0.231	(0.199)	0.231	(0.553)
Car Make - Infiniti	-0.319	(0.256)	-0.319	(0.366)
Car Make - Isuzu	0.160	(0.548)	0.160	(0.481)
Car Make - Jeep	-0.504**	(0.210)	-0.504	(0.840)
Car Make - Kia	0.161	(0.203)	0.161	(0.422)
Car Make - Land	1.656***	(0.634)	1.656***	(0.439)
Car Make - Lexus	0.607**	(0.305)	0.607+	(0.378)
Car Make - Lincoln	0.0656	(0.335)	0.0656	(0.641)
Car Make - Lotus	0.569	(0.452)	0.569	(0.513)
Car Make - Mazda	0.135	(0.184)	0.135	(0.379)
Car Make - Mercedes-Benz	-0.681***	(0.253)	-0.681**	(0.308)
Car Make - Mercury	0	(.)	0	(.)
Car Make - Mini	0.580*	(0.336)	0.580	(0.407)
Car Make - Mitsubishi	0.694**	(0.301)	0.694**	(0.290)
Car Make - Nissan	0.483*	(0.258)	0.483	(0.593)
Car Make - Pontiac	-1.335***	(0.385)	-1.335	(1.436)
Car Make - Porsche	0.603*	(0.308)	0.603**	(0.282)
Car Make - Ram	-0.463**	(0.215)	-0.463	(0.505)
Car Make - Saab	0.458	(0.322)	0.458	(0.510)
Car Make - Saturn	-0.477	(0.378)	-0.477	(0.606)
Car Make - Scion	0.501**	(0.200)	0.501	(0.656)
Car Make - Subaru	-0.522***	(0.170)	-0.522+	(0.323)
Car Make - Toyota	0.256	(0.162)	0.256	(0.329)
Car Make - Volkswagen	-0.978***	(0.166)	-0.978**	(0.468)
Car Make - Volvo	-0.179	(0.220)	-0.179	(0.509)
Car Type - Large Car	-0.0656	(0.192)	-0.0656	(0.269)
Car Type - Midsize Car	-0.314***	(0.0794)	-0.314**	(0.153)

Car Type - Minivan	0.435**	(0.217)	0.435	(0.490)
Car Type - Pickup Truck	-0.211	(0.154)	-0.211	(0.357)
Car Type - SUV	-0.0283	(0.0860)	-0.0283	(0.193)
Car Type - Station Wagon	0.354***	(0.124)	0.354	(0.229)
Car Type - Subcompact Car	-0.163	(0.124)	-0.163	(0.277)
Car Type - Two Seater	-0.291*	(0.155)	-0.291**	(0.123)
Car Type - Van	-0.119	(0.594)	-0.119	(0.672)
Car Age - 1	0.281***	(0.0739)	0.281**	(0.114)
Car Age - 2	0.369***	(0.0995)	0.369*	(0.209)
Car Age - 3	0.143	(0.116)	0.143	(0.402)
Car Age - 4	0.279**	(0.125)	0.279	(0.313)
Car Age - 5	1.067***	(0.177)	1.067***	(0.401)
Car Age - 6	0.233*	(0.125)	0.233	(0.199)
Car Age - 7	0.344***	(0.129)	0.344**	(0.161)
Car Age - 8	-1.078***	(0.150)	-1.078***	(0.185)
Car Age - 9	0.457***	(0.101)	0.457	(0.304)
Car Age - 10	0.294**	(0.125)	0.294	(0.317)
Car Age - 11	0.0384	(0.211)	0.0384	(0.295)
Car Age - 12	-0.0672	(0.147)	-0.0672	(0.285)
Car Age - 13	-0.325*	(0.166)	-0.325*	(0.172)
Car Age - 14	-0.564***	(0.148)	-0.564***	(0.215)
Car Age - 15	-0.107	(0.249)	-0.107	(0.318)
Car Age - 16	-0.629**	(0.296)	-0.629+	(0.388)
Car Age - 17	0.209	(0.253)	0.209	(0.843)
Car Age - 18	-0.769	(0.534)	-0.769	(0.900)
Car Age - 19	0.176	(0.513)	0.176	(0.334)
Car Age - 20	-0.693	(0.451)	-0.693*	(0.403)
Car Age - 21	-0.599**	(0.254)	-0.599*	(0.347)
Car Age - 22	0.857	(0.895)	0.857***	(0.318)
Car Age - 23	0.361	(0.448)	0.361	(0.322)
Car Age - 24	3.871***	(0.718)	3.871***	(0.627)
Car Age - 27	0	(.)	0	(.)
Car Age - 29	0	(.)	0	(.)
Car Age - 32	4.176***	(0.559)	4.176***	(0.595)
Car Age - 35	0	(.)	0	(.)
Constant	-1.826***	(0.361)	-1.826*	(1.057)
Observations	34466		34466	

Standard errors in parentheses
Base category is California for state, Acura for car make, compact car for car type and 0 for car age.
Column 2 shows the clustered standard errors.
$^{+}$ $p < 0.11$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Next, I add individual fixed effects to the model to account for unobserved heterogeneity among the consumers. This can avoid potential endogeneity issue in the form of omitted variable bias and reduce potential self-selection problem. Two logit models are estimated with individual fixed effects. The first has the probability of purchase regressed on the basic variables price, inventory, and consumption. Price and inventory are significant at the 0.01 significance level and both have the expected negative sign. The second is a logit model that regresses the probability of purchase on price, inventory, consumption, and consumer attributes state, car type, car model, and car year. Again, price and inventory are significant at the 0.01 significance level and both are negative. In addition, car model, car type, and car year are also significant.

Table 1.8: Logit Model of Purchase Probability with Individual FE

	(1) Purchase		(2) Purchase	
Purchase				
Price Paid	-2.022***	(0.326)	-2.400**	(0.962)
Inventory x Price	-3.642***	(0.867)	-3.273***	(0.910)
Consumption x Price	-0.00593	(1.424)	0.890	(1.472)
State			-0.0235	(0.0490)
Car Type			-0.203***	(0.0635)
Car Model			-0.0180***	(0.00297)
Car Year			0.113***	(0.0298)
Observations	58142		58142	

Standard errors in parentheses
Base category is California for state, Acura for car make, compact car for car type and 0 for car age.
Column 2 shows the clustered standard errors.
$^{+}$ $p < 0.11$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.5.3 Empirical Model - Purchase Frequency and Inventory

To test Hypothesis 2 concerning the duration between purchases during low prices, I estimate the following OLS model:

$$
\begin{aligned}
Duration_{it} \quad = \quad & \beta_0 + \beta_1 PriceDev_{it} + \beta_2 Inventory_{it} \\
& + \beta_3 Inventory_{it} \times PriceDev_{it} + \beta_4 Consumption_{it} \\
& + \beta_5 Consumption_{it} \times PriceDev_{it} + \beta_6 MPG_{it} \\
& + \beta_7 MSRP_i + \beta_8 TankCapacity_i + \sum_{r=1}^{R-1} \beta_{9,r} State_i \\
& + \sum_{s=1}^{S-1} \beta_{10,s} CarMake_i + \sum_{t=1}^{T-1} \beta_{11,t} CarType_i \\
& + \sum_{u=1}^{U-1} \beta_{12,u} CarAge_{it} + \epsilon_{it}
\end{aligned}
$$

where the dependent variable $Duration_{it}$ is the number of periods passed since the last purchase made by user i at time $t-1$. The independent variables include percent difference between actual price paid by the user and the state average retail gasoline price for that week ($PriceDev_{it}$), the inventory, consumption on a particular day, and interactions between $PriceDev_{it}$ and inventory and consumption. The same controls for car attributes and states are included as in the logit model in the last section. An observation in this model is a day in which an actual purchase is made.

The results in Table 1.9 show the expected positive sign for $PriceDev_{it}$, implying that during high prices (i.e. Price paid > Average Price), the greater the difference between state average and the actual price paid, the longer the duration since the previous purchase. The intuition is that for a user to pay a high price, he or she must have waited long enough for his inventory to run out and is left with no choice but to buy at the high price. Similarly, if the actual price paid by the user is lower than the state average price, it must mean that the user is eager to take advantage of the lower price despite that he or she may have just made a purchase not long ago. This supports Hypothesis 2. The sign of inventory is negative and expected. The higher the inventory, the shorter the duration from last purchase must be. Here the interaction between inventory and $PriceDev_{it}$ is not statistically significant, despite that the sign is as expected - the higher inventory, the greater the impact of $PriceDev_{it}$. The estimate for consumption is also negative, implying that the greater the consumption is, the faster gasoline is used up by the user and hence the shorter the duration from the last purchase is. The estimate for tank capacity is positive. The intuition is that cars with larger tanks can hold more fuel, so when users with bigger cars make a purchase, a longer time period has passed since the last one. Similar to the last section, most estimates even with clustered standard errors are still significant or marginally significant at the 0.1 or smaller significance levels.

Table 1.9: OLS Model of Duration Since Last Purchase

	(1)		(2)	
	Periods Since Last Purchase		Periods Since Last Purchase	
Price Deviation	2.513	(2.908)	2.513	(5.082)
Inventory	-11.75***	(0.807)	-11.75***	(1.086)
Inventory x Price Deviation	3.065	(8.395)	3.065	(9.882)
Consumption	-25.04***	(0.723)	-25.04***	(1.613)
Consumption x Price Deviation	-12.25*	(7.436)	-12.25*	(5.434)
Tank Capacity	0.271***	(0.0634)	0.271**	(0.0985)
MSRP	0.0000792***	(0.0000292)	0.0000792	(0.0000653)
MPG	0.179***	(0.00695)	0.179***	(0.0131)
State - Colorado	-0.623	(0.488)	-0.623	(0.375)
State - Florida	-1.044**	(0.441)	-1.044***	(0.215)
State - Massachusetts	-0.385	(0.713)	-0.385	(0.247)
State - Minnesota	-3.902***	(0.815)	-3.902***	(0.481)
State - New York	0.989*	(0.589)	0.989	(0.582)
State - Ohio	-1.939***	(0.497)	-1.939***	(0.427)
State - Texas	0.684^{+}	(0.421)	0.684*	(0.326)
State - Washington	-0.354	(0.609)	-0.354	(0.457)
Car Make - Audi	0.774	(1.148)	0.774	(1.654)
Car Make - BMW	7.655***	(1.927)	7.655*	(4.080)
Car Make - Buick	3.115**	(1.281)	3.115	(2.174)
Car Make - Cadillac	-0.324	(1.718)	-0.324	(1.877)
Car Make - Chevrolet	2.861***	(1.064)	2.861	(1.996)
Car Make - Chrysler	4.303*	(2.246)	4.303**	(1.817)
Car Make - Dodge	5.934***	(1.462)	5.934	(3.447)
Car Make - Fiat	4.865	(4.441)	4.865**	(1.835)
Car Make - Ford	2.360**	(0.961)	2.360	(2.369)
Car Make - GMC	2.725	(2.356)	2.725	(1.925)
Car Make - Honda	2.837***	(0.991)	2.837	(1.750)
Car Make - Hyundai	3.483***	(1.185)	3.483	(2.058)
Car Make - Infiniti	7.623***	(1.521)	7.623	(4.931)
Car Make - Isuzu	1.749	(3.312)	1.749	(5.944)
Car Make - Jeep	2.290*	(1.233)	2.290	(1.622)
Car Make - Kia	2.813**	(1.199)	2.813	(2.129)
Car Make - Land	-5.708*	(3.077)	-5.708**	(2.085)
Car Make - Lexus	2.765	(1.986)	2.765	(1.670)
Car Make - Lincoln	-1.420	(1.913)	-1.420	(1.899)

Car Make - Lotus	5.029**	(2.546)	5.029	(2.955)
Car Make - Mazda	2.978***	(1.105)	2.978	(2.047)
Car Make - Mercedes-Benz	10.62***	(1.721)	10.62*	(5.456)
Car Make - Mercury	0	(.)	0	(.)
Car Make - Mini	9.411***	(1.833)	9.411*	(4.973)
Car Make - Mitsubishi	0.543	(1.687)	0.543	(5.253)
Car Make - Nissan	0.610	(1.467)	0.610	(1.808)
Car Make - Pontiac	8.157***	(2.080)	8.157+	(4.403)
Car Make - Porsche	1.503	(2.001)	1.503	(3.539)
Car Make - Ram	-0.859	(1.295)	-0.859	(1.796)
Car Make - Saab	-2.569	(1.878)	-2.569	(2.101)
Car Make - Saturn	1.815	(2.383)	1.815	(1.584)
Car Make - Scion	1.852*	(1.099)	1.852	(1.689)
Car Make - Subaru	3.942***	(1.036)	3.942*	(1.815)
Car Make - Toyota	1.377	(0.980)	1.377	(1.808)
Car Make - Volkswagen	2.958***	(1.039)	2.958	(2.053)
Car Make - Volvo	2.528**	(1.232)	2.528**	(1.041)
Car Type - Large Car	-1.557	(1.113)	-1.557	(1.617)
Car Type - Midsize Car	1.346***	(0.477)	1.346*	(0.717)
Car Type - Minivan	-0.722	(1.152)	-0.722	(1.706)
Car Type - Pickup Truck	1.671*	(0.922)	1.671	(0.939)
Car Type - SUV	0.218	(0.494)	0.218	(0.802)
Car Type - Station Wagon	0.829	(0.687)	0.829	(1.502)
Car Type - Subcompact Car	1.253*	(0.713)	1.253	(1.013)
Car Type - Two Seater	0.834	(0.934)	0.834	(1.695)
Car Type - Van	-0.697	(3.512)	-0.697	(1.046)
Car Age - 1	-0.323	(0.426)	-0.323	(0.581)
Car Age - 2	-0.448	(0.558)	-0.448	(0.719)
Car Age - 3	-0.0426	(0.701)	-0.0426	(0.846)
Car Age - 4	-0.217	(0.678)	-0.217	(0.855)
Car Age - 5	-0.966	(0.917)	-0.966	(0.955)
Car Age - 6	-0.564	(0.668)	-0.564	(0.547)
Car Age - 7	-1.893***	(0.702)	-1.893*	(0.822)
Car Age - 8	1.346	(0.868)	1.346	(1.134)
Car Age - 9	-0.373	(0.569)	-0.373	(0.805)
Car Age - 10	1.379**	(0.661)	1.379	(1.215)
Car Age - 11	5.743***	(1.131)	5.743*	(2.846)
Car Age - 12	3.668***	(0.900)	3.668	(2.322)
Car Age - 13	3.218***	(1.014)	3.218	(1.832)
Car Age - 14	2.339**	(0.940)	2.339	(1.330)
Car Age - 15	-1.595	(1.624)	-1.595	(2.789)
Car Age - 16	9.927***	(1.706)	9.927*	(4.402)
Car Age - 17	3.246**	(1.616)	3.246	(3.340)
Car Age - 18	8.747**	(3.465)	8.747**	(3.084)
Car Age - 19	-0.625	(2.587)	-0.625	(4.584)
Car Age - 20	0.530	(2.628)	0.530	(0.758)
Car Age - 21	9.171***	(1.439)	9.171**	(3.287)

Car Age - 22	1.927	(5.374)	1.927	(1.533)
Car Age - 23	-7.656***	(2.452)	-7.656*	(3.729)
Car Age - 24	4.232	(3.740)	4.232**	(1.277)
Car Age - 27	0	(.)	0	(.)
Car Age - 29	0	(.)	0	(.)
Car Age - 32	1.328	(3.571)	1.328	(3.336)
Car Age - 35	0	(.)	0	(.)
Constant	1.876	(1.350)	1.876	(2.393)
Observations	3887		3887	

Standard errors in parentheses
Base category is California for state, Acura for car make, compact car for car type and 0 for car age.
Column 2 shows the clustered standard errors.
$^{+}$ $p < 0.11$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1.5.4 Empirical Model - Stockpiling, Price, and Consumption

To test Hypothesis 3 concerning the change in consumption after a purchase, I estimate the following OLS model:

$$\begin{aligned} ConsumptionChange_{it} &= \beta_0 + \beta_1 PriceDev_{it} + \beta_2 Consumption_{it} \\ &\quad + \beta_3 Consumption_{it} \times PriceDev_{it} + \beta_4 MPG_{it} \\ &\quad + \beta_5 MSRP_i + \beta_6 TankCapacity_i + \sum_{r=1}^{R-1} \beta_{7,r} State_i \\ &\quad + \sum_{s=1}^{S-1} \beta_{8,s} CarMake_i + \sum_{t=1}^{T-1} \beta_{9,t} CarType_i \\ &\quad + \sum_{u=1}^{U-1} \beta_{10,u} CarAge_{it} + \epsilon_{it} \end{aligned}$$

where $ConsumptionChange_{it}$ is the change in consumer i's consumption before and after a purchase at time t. The independent variables are consumption, consumption interacted with price, MPG, MSRP, tank capacity and car characteristics as in the previous model. The parameter β_1 has the expected negative sign and is statistically significant at the 0.1 significance level. This implies that the lower the actual price paid is relative to the state average price, the greater positive change in consumption there will be after a purchase. In other words, consumers who purchase during low prices will increase consumption in the next period, while consumers who purchase during high prices will decrease consumption in the next period. The estimate for $Consumption_{it}$ is highly significant with a negative sign, suggesting that high consumption in the current period t will lead to a smaller consumption next period and the difference increases with consumption in the current period.

Table 1.10: OLS Model of Change in Consumption

	(1) Change in Consumption		(2) Change in Consumption	
Price Deviation	-0.0397*	(0.0222)	-0.0397	(0.0249)
Consumption	-0.966***	(0.00897)	-0.966***	(0.0110)
Consumption x Price Deviation	0.146	(0.0923)	0.146	(0.101)
Tank Capacity	-0.00310***	(0.000779)	-0.00310*	(0.00159)
MSRP	-0.000000346	(0.000000363)	-0.000000346	(0.000000651)
MPG	0.000181**	(0.0000859)	0.000181*	(0.0000946)
State - Colorado	-0.00120	(0.00604)	-0.00120	(0.00373)
State - Florida	0.0298***	(0.00547)	0.0298***	(0.00304)
State - Massachusetts	0.0332***	(0.00884)	0.0332***	(0.00385)
State - Minnesota	0.0540***	(0.0101)	0.0540***	(0.0142)
State - New York	-0.0123*	(0.00729)	-0.0123	(0.00732)
State - Ohio	0.0107*	(0.00617)	0.0107^{+}	(0.00592)
State - Texas	-0.00426	(0.00522)	-0.00426	(0.00406)
State - Washington	-0.00530	(0.00756)	-0.00530	(0.00504)
Car Make - Audi	-0.0254*	(0.0143)	-0.0254	(0.0368)
Car Make - BMW	-0.0565**	(0.0239)	-0.0565	(0.0366)
Car Make - Buick	0.0102	(0.0159)	0.0102	(0.0475)
Car Make - Cadillac	-0.0196	(0.0213)	-0.0196	(0.0338)
Car Make - Chevrolet	-0.0439***	(0.0132)	-0.0439	(0.0331)
Car Make - Chrysler	-0.0343	(0.0279)	-0.0343	(0.0342)
Car Make - Dodge	-0.0198	(0.0181)	-0.0198	(0.0351)
Car Make - Fiat	-0.0811	(0.0551)	-0.0811**	(0.0311)
Car Make - Ford	-0.0203*	(0.0119)	-0.0203	(0.0439)
Car Make - GMC	-0.0428	(0.0291)	-0.0428	(0.0468)
Car Make - Honda	-0.0275**	(0.0123)	-0.0275	(0.0301)
Car Make - Hyundai	-0.0405***	(0.0147)	-0.0405	(0.0400)
Car Make - Infiniti	-0.00466	(0.0189)	-0.00466	(0.0517)
Car Make - Isuzu	-0.0415	(0.0410)	-0.0415	(0.0511)
Car Make - Jeep	-0.0350**	(0.0153)	-0.0350	(0.0370)
Car Make - Kia	-0.0136	(0.0149)	-0.0136	(0.0277)

Car Make - Land	0.104***	(0.0382)	0.104**	(0.0381)
Car Make - Lexus	-0.0605**	(0.0246)	-0.0605	(0.0367)
Car Make - Lincoln	-0.0264	(0.0237)	-0.0264	(0.0423)
Car Make - Lotus	-0.0683**	(0.0316)	-0.0683**	(0.0272)
Car Make - Mazda	-0.0264*	(0.0137)	-0.0264	(0.0261)
Car Make - Mercedes-Benz	-0.0526**	(0.0214)	-0.0526	(0.0312)
Car Make - Mercury	0	(.)	0	(.)
Car Make - Mini	-0.0852***	(0.0227)	-0.0852*	(0.0453)
Car Make - Mitsubishi	-0.0649***	(0.0209)	-0.0649	(0.0624)
Car Make - Nissan	-0.0219	(0.0182)	-0.0219	(0.0197)
Car Make - Pontiac	-0.0204	(0.0258)	-0.0204	(0.0381)
Car Make - Porsche	-0.0458*	(0.0248)	-0.0458	(0.0368)
Car Make - Ram	-0.0198	(0.0160)	-0.0198	(0.0423)
Car Make - Saab	0.0224	(0.0233)	0.0224	(0.0459)
Car Make - Saturn	-0.0639**	(0.0296)	-0.0639*	(0.0303)
Car Make - Scion	-0.0179	(0.0136)	-0.0179	(0.0458)
Car Make - Subaru	-0.00891	(0.0129)	-0.00891	(0.0380)
Car Make - Toyota	-0.0377***	(0.0122)	-0.0377	(0.0345)
Car Make - Volkswagen	-0.0255**	(0.0129)	-0.0255	(0.0453)
Car Make - Volvo	-0.0207	(0.0153)	-0.0207	(0.0309)
Car Type - Large Car	-0.00553	(0.0138)	-0.00553	(0.0156)
Car Type - Midsize Car	-0.00249	(0.00592)	-0.00249	(0.00870)
Car Type - Minivan	0.0154	(0.0143)	0.0154	(0.0260)
Car Type - Pickup Truck	0.0278**	(0.0114)	0.0278	(0.0278)
Car Type - SUV	0.00976	(0.00614)	0.00976	(0.00887)
Car Type - Station Wagon	-0.0215**	(0.00851)	-0.0215	(0.0200)
Car Type - Subcompact Car	-0.0134	(0.00885)	-0.0134	(0.0104)
Car Type - Two Seater	0.00206	(0.0116)	0.00206	(0.0116)
Car Type - Van	0.00331	(0.0436)	0.00331	(0.0295)
Car Age - 1	0.0134**	(0.00529)	0.0134	(0.00869)
Car Age - 2	0.0297***	(0.00693)	0.0297**	(0.0102)
Car Age - 3	0.0212**	(0.00870)	0.0212	(0.0158)
Car Age - 4	0.0322***	(0.00840)	0.0322	(0.0183)
Car Age - 5	-0.0111	(0.0113)	-0.0111	(0.0110)
Car Age - 6	0.0241***	(0.00829)	0.0241*	(0.0127)
Car Age - 7	0.0418***	(0.00871)	0.0418**	(0.0151)
Car Age - 8	0.0283***	(0.0107)	0.0283*	(0.0140)
Car Age - 9	0.0574***	(0.00703)	0.0574**	(0.0173)
Car Age - 10	-0.00386	(0.00820)	-0.00386	(0.0135)
Car Age - 11	-0.0231*	(0.0140)	-0.0231	(0.0187)
Car Age - 12	0.0160	(0.0112)	0.0160	(0.0196)
Car Age - 13	-0.0318**	(0.0126)	-0.0318***	(0.00842)
Car Age - 14	-0.00260	(0.0117)	-0.00260	(0.0202)
Car Age - 15	0.00456	(0.0202)	0.00456	(0.00966)
Car Age - 16	0.0348+	(0.0212)	0.0348	(0.0272)
Car Age - 17	-0.0178	(0.0201)	-0.0178	(0.0194)
Car Age - 18	-0.0940**	(0.0430)	-0.0940*	(0.0428)

Car Age - 19	0.135***	(0.0321)	0.135**	(0.0478)
Car Age - 20	0.0193	(0.0326)	0.0193*	(0.00893)
Car Age - 21	-0.0156	(0.0179)	-0.0156	(0.0157)
Car Age - 22	-0.0775	(0.0667)	-0.0775***	(0.0148)
Car Age - 23	0.0674**	(0.0304)	0.0674***	(0.0187)
Car Age - 24	-0.0445	(0.0463)	-0.0445	(0.0293)
Car Age - 27	0	(.)	0	(.)
Car Age - 29	0	(.)	0	(.)
Car Age - 32	-0.0238	(0.0439)	-0.0238	(0.0220)
Car Age - 35	0	(.)	0	(.)
Constant	0.153***	(0.0167)	0.153***	(0.0360)
Observations	3887		3887	

Standard errors in parentheses
Base category is California for state, Acura for car make, compact car for car type and 0 for car age.
Column 2 shows the clustered standard errors.
$^{+}$ $p < 0.11$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

To explore the association between consumption and gallons purchased per transaction, I estimate the following OLS models:

$$\begin{aligned} Gallons_{it} = \ & \beta_0 + \beta_1 PriceDev_{it} + \beta_2 Consumption_{it-1} \\ & + \beta_3 MPG_{it} \\ & + \beta_4 MSRP_i + \beta_5 TankCapacity_i + \sum_{r=1}^{R-1} \beta_{6,r} State_i \\ & + \sum_{s=1}^{S-1} \beta_{7,s} CarMake_i + \sum_{t=1}^{T-1} \beta_{8,t} CarType_i \\ & + \sum_{u=1}^{U-1} \beta_{9,u} CarAge_{it} + \epsilon_{it} \end{aligned}$$

Table 1.11: OLS Model of Consumption on Amount of Gas Purchased

	(1) Gallons	(2) Gallons

Consumption (t-1)	0.302	(0.313)		
Price Deviation	2.599***	(0.615)	-9.527***	(0.680)
Tank Capacity	0.423***	(0.0261)	0.0339***	(0.0104)
MSRP	0.0000267**	(0.0000122)	-0.00000778*	(0.00000450)
MPG	-0.0163***	(0.00277)	-0.00931***	(0.000684)
State - Colorado	0.430**	(0.195)	0.288***	(0.0863)
State - Florida	0.178	(0.177)	0.271***	(0.0817)
State - Massachusetts	-0.461+	(0.285)	0.223*	(0.130)
State - Minnesota	0.828**	(0.327)	1.094***	(0.152)
State - New York	-0.328	(0.238)	-0.178*	(0.0955)
State - Ohio	0.349*	(0.200)	0.203**	(0.0891)
State - Texas	-0.0709	(0.171)	0.149**	(0.0744)
State - Washington	0.138	(0.249)	-0.0647	(0.105)
Car Make - Audi	-0.191	(0.466)	0.206	(0.193)
Car Make - BMW	1.531*	(0.842)	-0.436*	(0.246)
Car Make - Buick	0.155	(0.529)	0.309	(0.215)
Car Make - Cadillac	3.056***	(0.680)	0.768**	(0.312)
Car Make - Chevrolet	-0.500	(0.437)	-0.414**	(0.171)
Car Make - Chrysler	1.297	(0.902)	-0.530	(0.335)
Car Make - Dodge	3.576***	(0.618)	-0.650***	(0.205)
Car Make - Fiat	-1.191	(2.013)	-1.061+	(0.648)
Car Make - Ford	-0.849**	(0.401)	0.0103	(0.160)
Car Make - GMC	-6.843***	(1.015)	-1.017***	(0.382)
Car Make - Honda	-0.665+	(0.410)	-0.0939	(0.166)
Car Make - Hyundai	-1.325***	(0.485)	-0.200	(0.199)
Car Make - Infiniti	1.003	(0.637)	-0.613***	(0.218)
Car Make - Isuzu	-2.594*	(1.350)	-0.650	(0.466)
Car Make - Jeep	0.0556	(0.505)	-0.107	(0.203)
Car Make - Kia	-1.162**	(0.490)	-0.281	(0.210)
Car Make - Land	0.681	(1.234)	2.287***	(0.614)
Car Make - Lexus	-1.850**	(0.812)	-0.255	(0.294)
Car Make - Lincoln	0.764	(0.770)	0.259	(0.328)
Car Make - Lotus	-2.366**	(1.027)	-0.204	(0.410)
Car Make - Mazda	-0.675	(0.457)	-0.166	(0.183)
Car Make - Mercedes-Benz	0.294	(0.759)	-0.854***	(0.223)
Car Make - Mercury	0	(.)	0	(.)
Car Make - Mini	-0.503	(0.744)	0.485	(0.324)
Car Make - Mitsubishi	-2.189***	(0.713)	-0.257	(0.299)
Car Make - Nissan	-1.014*	(0.597)	0.347	(0.264)
Car Make - Pontiac	0.594	(0.878)	-0.529	(0.337)
Car Make - Porsche	-2.680***	(0.825)	-0.144	(0.299)
Car Make - Ram	0.952*	(0.531)	0.0837	(0.216)
Car Make - Saab	0.867	(0.779)	0.687**	(0.328)
Car Make - Saturn	-0.741	(1.005)	-0.438	(0.373)
Car Make - Scion	-0.859*	(0.449)	0.129	(0.211)
Car Make - Subaru	0.00115	(0.430)	-0.306*	(0.168)
Car Make - Toyota	-0.933**	(0.405)	-0.0108	(0.162)

Car Make - Volkswagen	0.588	(0.433)	-0.328**	(0.164)
Car Make - Volvo	0.513	(0.501)	-0.0782	(0.210)
Car Type - Large Car	0.00873	(0.449)	-0.0881	(0.182)
Car Type - Midsize Car	0.141	(0.194)	-0.230***	(0.0794)
Car Type - Minivan	0.927**	(0.463)	0.411*	(0.224)
Car Type - Pickup Truck	1.714***	(0.380)	0.147	(0.157)
Car Type - SUV	0.537***	(0.200)	0.0567	(0.0870)
Car Type - Station Wagon	-1.298***	(0.275)	-0.265**	(0.130)
Car Type - Subcompact Car	-0.663**	(0.295)	-0.320**	(0.125)
Car Type - Two Seater	-0.288	(0.376)	-0.320**	(0.151)
Car Type - Van	-0.800	(2.032)	0.592	(0.656)
Car Age - 1	-0.134	(0.175)	0.452***	(0.0767)
Car Age - 2	0.108	(0.227)	0.298***	(0.105)
Car Age - 3	0.498*	(0.291)	0.358***	(0.122)
Car Age - 4	1.094***	(0.271)	0.794***	(0.130)
Car Age - 5	-1.542***	(0.369)	0.149	(0.171)
Car Age - 6	0.248	(0.263)	0.377***	(0.131)
Car Age - 7	0.401	(0.280)	0.595***	(0.142)
Car Age - 8	2.327***	(0.361)	-0.0564	(0.146)
Car Age - 9	-1.081***	(0.223)	0.0978	(0.104)
Car Age - 10	-0.491*	(0.266)	0.425***	(0.131)
Car Age - 11	-1.150**	(0.468)	0.0418	(0.189)
Car Age - 12	1.040***	(0.379)	-0.0566	(0.142)
Car Age - 13	-0.241	(0.403)	-0.207	(0.154)
Car Age - 14	1.058***	(0.382)	-0.125	(0.136)
Car Age - 15	0.482	(0.743)	0.498**	(0.241)
Car Age - 16	1.229*	(0.738)	-0.293	(0.252)
Car Age - 17	-1.137+	(0.698)	-0.445*	(0.252)
Car Age - 18	0.549	(1.656)	-0.622	(0.476)
Car Age - 19	3.806***	(1.093)	1.812***	(0.590)
Car Age - 20	2.105**	(1.040)	0.213	(0.442)
Car Age - 21	-1.070*	(0.570)	-0.781***	(0.226)
Car Age - 22	0	(.)	1.730	(1.504)
Car Age - 23	1.508	(1.027)	1.289***	(0.433)
Car Age - 24	-1.410	(2.142)	1.487*	(0.804)
Car Age - 27	0	(.)	0	(.)
Car Age - 29	0	(.)	0	(.)
Car Age - 32	-7.401***	(1.445)	-0.769	(0.719)
Car Age - 35	0	(.)	0	(.)
Constant	4.688***	(0.550)	1.210***	(0.216)
Observations	3377		34711	

Standard errors in parentheses

Base category is California for state, Acura for car make, compact car for car type and 0 for car age.

$^{+}$ $p < 0.11$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The coefficient for consumption in the previous period is not significant in the first regression, so changes in consumption in the period before a purchase does not significantly affect the purchase amount. When consumers purchase more or less than usual, it is not likely due to consumption shocks in the past. In other words, any unusual over- or under-use of gasoline will not change consumer's purchase amount. The coefficient for price deviation from the state average price is significant with or without consumption in the previous period as a predictor variable. The sign is negative as expected in the second regression; higher price compared to the state average results in less gallons purchased. The sign for price deviation is positive in the first regression. However, it should be taken with a caveat because consumption is included and not significant.

I also estimate the following OLS model to explore the association between gallons purchased and consumption in the period following the purchase.

$$
\begin{aligned}
Consumption_{it} \quad = \quad & \beta_0 + \beta_1 PriceDev_{it} + \beta_2 Gallons_{it-1} \\
& + \beta_3 MPG_{it} \\
& + \beta_4 MSRP_i + \beta_5 TankCapacity_i + \sum_{r=1}^{R-1} \beta_{6,r} State_i \\
& + \sum_{s=1}^{S-1} \beta_{7,s} CarMake_i + \sum_{t=1}^{T-1} \beta_{8,t} CarType_i \\
& + \sum_{u=1}^{U-1} \beta_{9,u} CarAge_{it} + \epsilon_{it}
\end{aligned}
$$

Table 1.12: OLS Model of Gallons Purchased on Consumption

	(1) Consumption (t)		(2) Consumption (t)	
Price Deviation	0.0219	(0.0345)		
Gallons	-0.00740***	(0.000975)	-0.00736***	(0.000973)
Tank Capacity	0.00621***	(0.00152)	0.00626***	(0.00152)
MSRP	0.00000114*	(0.000000683)	0.00000114*	(0.000000683)
MPG	0.000428***	(0.000155)	0.000437***	(0.000155)
State - Colorado	0.0487***	(0.0109)	0.0478***	(0.0108)
State - Florida	0.0127	(0.00990)	0.0130	(0.00989)

State - Massachusetts	0.0000706	(0.0160)	-0.000134	(0.0160)
State - Minnesota	0.0126	(0.0183)	0.0132	(0.0183)
State - New York	0.00813	(0.0133)	0.00734	(0.0132)
State - Ohio	-0.00365	(0.0112)	-0.00426	(0.0112)
State - Texas	0.0446***	(0.00953)	0.0443***	(0.00952)
State - Washington	0.0208	(0.0139)	0.0205	(0.0139)
Car Make - Audi	-0.0274	(0.0261)	-0.0268	(0.0261)
Car Make - BMW	-0.0746	(0.0472)	-0.0719	(0.0470)
Car Make - Buick	0.0276	(0.0296)	0.0294	(0.0295)
Car Make - Cadillac	0.0337	(0.0382)	0.0340	(0.0382)
Car Make - Chevrolet	-0.0183	(0.0245)	-0.0167	(0.0244)
Car Make - Chrysler	0.0407	(0.0505)	0.0418	(0.0505)
Car Make - Dodge	0.0149	(0.0348)	0.0167	(0.0347)
Car Make - Fiat	-0.0563	(0.113)	-0.0531	(0.113)
Car Make - Ford	0.00124	(0.0225)	0.00283	(0.0223)
Car Make - GMC	-0.0376	(0.0572)	-0.0367	(0.0572)
Car Make - Honda	0.0513**	(0.0230)	0.0529**	(0.0228)
Car Make - Hyundai	0.00825	(0.0272)	0.00948	(0.0271)
Car Make - Infiniti	-0.00803	(0.0357)	-0.00824	(0.0357)
Car Make - Isuzu	0.0234	(0.0756)	0.0215	(0.0756)
Car Make - Jeep	-0.00141	(0.0283)	-0.000429	(0.0283)
Car Make - Kia	-0.0176	(0.0275)	-0.0168	(0.0274)
Car Make - Land	-0.00923	(0.0691)	-0.00683	(0.0690)
Car Make - Lexus	-0.0664	(0.0455)	-0.0651	(0.0454)
Car Make - Lincoln	-0.0941**	(0.0431)	-0.0929**	(0.0430)
Car Make - Lotus	0.154***	(0.0575)	0.159***	(0.0571)
Car Make - Mazda	0.000702	(0.0256)	0.00241	(0.0255)
Car Make - Mercedes-Benz	0.0371	(0.0425)	0.0373	(0.0425)
Car Make - Mercury	0	(.)	0	(.)
Car Make - Mini	0.0126	(0.0417)	0.0125	(0.0417)
Car Make - Mitsubishi	0.00856	(0.0400)	0.00885	(0.0400)
Car Make - Nissan	-0.0270	(0.0335)	-0.0252	(0.0333)
Car Make - Pontiac	0.00793	(0.0492)	0.00932	(0.0491)
Car Make - Porsche	-0.0410	(0.0463)	-0.0405	(0.0463)
Car Make - Ram	0.00320	(0.0298)	0.00498	(0.0296)
Car Make - Saab	-0.0631	(0.0436)	-0.0610	(0.0435)
Car Make - Saturn	-0.0476	(0.0563)	-0.0461	(0.0562)
Car Make - Scion	0.00752	(0.0251)	0.00868	(0.0251)
Car Make - Subaru	-0.00418	(0.0241)	-0.00196	(0.0238)
Car Make - Toyota	0.0203	(0.0227)	0.0217	(0.0226)
Car Make - Volkswagen	-0.0199	(0.0243)	-0.0181	(0.0241)
Car Make - Volvo	0.0229	(0.0281)	0.0235	(0.0281)
Car Type - Large Car	-0.0409+	(0.0251)	-0.0417*	(0.0251)
Car Type - Midsize Car	-0.00370	(0.0109)	-0.00418	(0.0108)
Car Type - Minivan	-0.0534**	(0.0259)	-0.0544**	(0.0259)
Car Type - Pickup Truck	-0.0134	(0.0213)	-0.0135	(0.0213)
Car Type - SUV	-0.0161	(0.0112)	-0.0168	(0.0111)

Car Type - Station Wagon	-0.0177	(0.0155)	-0.0180	(0.0154)
Car Type - Subcompact Car	-0.0268^{+}	(0.0165)	-0.0262	(0.0165)
Car Type - Two Seater	-0.0353*	(0.0211)	-0.0356*	(0.0211)
Car Type - Van	-0.112	(0.114)	-0.115	(0.114)
Car Age - 1	0.00412	(0.00979)	0.00391	(0.00978)
Car Age - 2	-0.00123	(0.0127)	-0.00157	(0.0127)
Car Age - 3	0.00138	(0.0163)	0.000364	(0.0162)
Car Age - 4	0.0526***	(0.0152)	0.0523***	(0.0152)
Car Age - 5	-0.0257	(0.0207)	-0.0260	(0.0207)
Car Age - 6	0.0140	(0.0147)	0.0132	(0.0147)
Car Age - 7	-0.00730	(0.0157)	-0.00752	(0.0157)
Car Age - 8	0.0326^{+}	(0.0203)	0.0326^{+}	(0.0203)
Car Age - 9	-0.0189	(0.0125)	-0.0188	(0.0125)
Car Age - 10	0.0785***	(0.0149)	0.0788***	(0.0148)
Car Age - 11	0.0692***	(0.0262)	0.0692***	(0.0262)
Car Age - 12	0.0529**	(0.0212)	0.0525**	(0.0212)
Car Age - 13	-0.0164	(0.0226)	-0.0161	(0.0225)
Car Age - 14	-0.0202	(0.0214)	-0.0203	(0.0214)
Car Age - 15	-0.0226	(0.0416)	-0.0233	(0.0416)
Car Age - 16	0.0324	(0.0413)	0.0337	(0.0413)
Car Age - 17	-0.0266	(0.0391)	-0.0274	(0.0390)
Car Age - 18	-0.0180	(0.0927)	-0.0206	(0.0926)
Car Age - 19	-0.00391	(0.0613)	-0.00394	(0.0613)
Car Age - 20	0.0289	(0.0582)	0.0285	(0.0582)
Car Age - 21	0.0959***	(0.0319)	0.0968***	(0.0319)
Car Age - 22	0	(.)	0	(.)
Car Age - 23	0.0836	(0.0575)	0.0813	(0.0574)
Car Age - 24	0.282**	(0.120)	0.284**	(0.120)
Car Age - 27	0	(.)	0	(.)
Car Age - 29	0	(.)	0	(.)
Car Age - 32	-0.0828	(0.0812)	-0.0832	(0.0812)
Car Age - 35	0	(.)	0	(.)
Constant	-0.0118	(0.0311)	-0.0145	(0.0309)
Observations	3377		3377	

Standard errors in parentheses
Base category is California for state, Acura for car make, compact car for car type and 0 for car age.
Column 2 shows the clustered standard errors.
$^{+}$ $p < 0.11$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The first regression shows that price deviation from the state average price is not significant, so it is not included in the second regression. Only gallons purchased in included in the second regression and it is found to be significant. The sign is negative. This seems counter-intuitive that higher gallons purchased is associated with less consumption in the following period. However, the magnitude is very small, so this should not be a big concern. Overall, consumption seems to be an important predictor for gallons purchased in the period immediately after. This makes sense as the consumers tend to fill up their car tanks after positive consumption shocks. If they use less in a period, their purchase in the next period will be significant smaller. On the other hand, the amount of gasoline purchased does not seem to be a strong predictor for consumption next period. This means that when consumers purchase, they are probably not buying for immediate consumption.

1.6 Conclusion and Future Research

This paper empirically estimates models that relate stockpiling behavior to various factors involved in purchasing gasoline such as price, inventory, duration between purchases, and consumption before and after purchases. Using a unique data set compiled from several sources, I find evidence that stockpiling behavior exists among consumers of gasoline. Consumers low on inventory are more likely to purchase and purchase more gasoline on a given day. On the other hand, consumers high on inventory can afford to delay purchase to wait for a better deal and they are more price sensitive. Duration can also be used to infer stockpiling behavior. When a purchase is made during low prices, the duration from the previous purchase tends to be shorter than when a purchase is made during high prices. Furthermore, the consumption of gasoline tends to decrease following purchase during high prices and increase following purchases during low prices. These findings support the hypotheses and provide insights to how consumers behave when deciding whether to purchase gasoline.

There are several directions for future research. It would be interesting to investigate consumer stockpiling behavior at a higher frequency because retail gasoline prices are known to fluctuate very frequently. An extension in this direction would require data tracking consumer decisions at a finer frequency such as daily or hourly. There are certain regions that require gasoline stations to post prices hourly or daily by law, so obtaining high frequency retail prices for these regions is possible. Another extension would be to explore the decision making process made by consumers. This would be possible if there exists data tracking the prices consumers see before making a purchase. I do not know any resource offering this kind of data, so it will likely require a time-consuming data collection process tracking consumers over time. One way would be to track the movement of consumer vehicles over time and identify the gas stations the vehicles passed by and the time of passing. Then map each passing to the retail price of that gas station at that particular time. In addition, a structural model could also offer insights to consumer decision making process and stockpiling behavior. It will likely require heavy computational power to estimate such a structural model with high frequency data points, but with the improving computational tools it can likely be overcome.

Chapter 2

Competitive Expansion Strategies Between Retailers

2.1 Introduction

In 1962, Wal-Mart and Target opened their first stores in Arkansas and Minnesota respectively. Today, both firms have grown into giant retail chains generating billions of revenues annually with stores across North America and the rest of the world. By 2013, Wal-Mart has over 7000 stores and Target has over 1700 stores in the U.S. Competition between the firms have constantly been a popular topic in the media and among academics.

It has been widely acknowledged that retailers make strategic decisions such as market entry and exit. In order to study a firm's expansion strategy over time, we need a dynamic model that will encompass the firm's forward-looking behavior rather than a static model that only provides a snapshot of the firm. A handful of research has been done on the retail industry focusing on main players such as Wal-Mart and Target using static models (Jia 2008 and Zhu et al 2009). One advantage of static models is that they are more tractable than dynamic models so these models tend to include more firms and make it easier to study competitive effects between firms. However, static models study the market at only one point in time and neglect the forward-looking behavior of firms. On the other hand, dynamic models allow researchers to study firms over a period of time taking into account their forward-looking competitive behavior. However, due the complexity of such models, dynamic models usually include only one firm (for example, Holmes 2011) and neglect the effect of competitors. With new development of numerical methods, it is possible to estimate more complex dynamic models in a tractable way.

To address the gap in literature on competitive dynamics of multiple firm decisions in retail market entry and exit, this paper examines how Wal-Mart and Target's expansion strategies differ and how their variable profits and fixed costs, hence total profits, are affected by the entry decisions of themselves and competitors as well as market characteristics. The firms are assumed to be forward-looking and compete by playing Markov perfect equilibrium (MPE) strategies in a dynamic entry game. Entry is represented by the opening of a new store in a market, which is defined to be a city. This paper assumes that once a firm enters the market, it does not exit (i.e. no store closure). This assumption is supported by the data and is used by Holmes in his paper on the diffusion pattern of Wal-Mart (2011). The estimation method used is nested pseudo-likelihood (NPL) proposed by Aguirregabiria and Mira (2007). It is a two-step numerical method that speeds up the traditional one-step nested fixed point method significantly. One main result is that Wal-Mart dominates the competition in almost all market structures. Wal-Mart's profits are unaffected by Target's presence in small and medium markets. In large markets, Wal-Mart's profits decrease initially with store disadvantage if it is operating less than four stores in the market. With more than four stores, Wal-Mart's profits increase regardless of the number of stores Target has. On the other hand, Target's profits increase when it has two stores in the market and store advantage (more stores than Wal-Mart), and decrease beyond two stores. The initial increase is only high enough for Target to profit in a large markets with sufficient population. In small and medium markets, Target does not profit regardless of the number of stores owned by Target and Wal-Mart. This is probably due to Target's high fixed costs outweighing the variable profits which depend on the market size. Target would need a large market size in order to cover the fixed costs which increase more than linearly with the number of stores it has in the market.

2.2 Literature Review

An example of static model is the paper by Jia (2008) who examines the impact of chain stores such as Wal-Mart on competing discount stores such as Kmart in the same market. She finds that Wal-Mart's presence on Kmart was much stronger in 1997 than in 1988, and that scale economies were important for Wal-Mart, but less for Kmart. A limitation of Jia's paper is that the model is static. It assumes that all Wal-Mart and Kmart stores are opened in one period; it does not allow firms to be forward-looking or delay store openings. Zhu et al. (2009) examine the competition among Wal-Mart, Target, and Kmart by estimating a discrete game model in which each firm's decision on entry and store format depend on their rival's decisions and presence and market characteristics. They find that Wal-Mart, as a dominant firm, often expands into new markets where it has no competitors despite that the new market may have less attractive market characteristics. Moreover, they find that Wal-Mart can withstand competition to a greater extent than its rivals. The limitation of their research is that it is also a static model, providing only a snapshot at a point of time. Perhaps this feature enables them to include a third competitor, Kmart, due to the tractability of the model, but a static model ignores the competitive dynamics inherent in strategic decisions of market entry and exit. An example of a dynamic model is the paper by Holmes (2011) who estimates the benefit of the economies of density for Wal-Mart from its dynamic expansion process and finds that Wal-Mart benefits significantly from high store density and is willing to suffer some initial losses to achieve economies of density. Although Holmes' model is dynamic, it only focuses on one firm. This reduces the computational burden dramatically, and allows him to include rich geographic variables on store format and location of stores and distribution centers. However, it does not account for the location and decision of the competitors. This paper addresses this gap in literature by studying how two retailers compete in using expansion strategies in a dynamic setting.

2.3 Model

As in Aguirregabiria, I start with a simple model of two firms selling a homogeneous product in an oligopoly. Suppose the demand function at time t is $Q_t = S_t(b_0 - b_1 P_t)$, where b_0 and b_1 are parameters, $Q_t = Q_{1t} + Q_{2t}$ is the aggregate demand, S_t is the exogenous market size defined by the number of consumers in the market, and P_t is the product price charged by both firms. There are two stages in each period. First, the firms compete in a Cournot game by setting quantities simultaneously. Then they decide whether to invest in extra capacity by setting up new stores. This stage is dynamic because the firms are assumed to be forward-looking. Production costs are assumed to be $C_{it} = MC_{it}Q_{it}$, which is linear in quantity produced. MC_{it} is the marginal cost and is assumed to depend on firm i's presence by equation $MC_{it} = c_i + d(X_{it} + Y_{it})$, where c_i and d are parameters, X_{it} is the current number of stores in a market for firm i , and $Y_{it} \in \{0, 1\}$, $Y_{it} = 1$ being firm i's decision to set up a new store at time t and $Y_{it} = 0$ if firm does not open a new store. A state is defined as $\{X_{it}, X_{jt}\}$, which is the number of stores owned by firm i and j respectively. The entry decision is assumed to be independent across markets.

To derive the value function, I start with the demand function $Q_t = S_t(b_0 + b_1 P_t)$ with market size S_t,

$$P_t = \frac{1}{b_1}\left(b_0 - \frac{Q_t}{S_t}\right) = \frac{1}{b_1}\left(b_0 - \frac{Q_1 + Q_2}{S_t}\right)$$

Variable profit functions are

$$\begin{aligned}
\pi_{1t} &= Q_{1t}\left(\frac{1}{b_1}\left(b_0 - \frac{Q_{1t} + Q_{2t}}{S_t}\right) - MC_{1t}\right) \\
\pi_{2t} &= Q_{2t}\left(\frac{1}{b_1}\left(b_0 - \frac{Q_{1t} + Q_{2t}}{S_t}\right) - MC_{2t}\right)
\end{aligned}$$

Taking FOC,

$$\begin{aligned}
\frac{\partial \pi_{1t}}{\partial Q_{1t}} &= \frac{1}{b_1}\left(b_0 - \frac{Q_{1t} + Q_{2t}}{S_t}\right) - MC_{1t} - Q_{1t}\left(\frac{1}{b_1 S_t}\right) \\
&= 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \pi_{2t}}{\partial Q_{2t}} &= \frac{1}{b_1}\left(b_0 - \frac{Q_{1t} + Q_{2t}}{S_t}\right) - MC_{2t} - Q_{2t}\left(\frac{1}{b_1 S_t}\right) \\
&= 0
\end{aligned}$$

We have

$$\begin{aligned}
Q_{1t} &= \frac{1}{2}\left(b_0 S_t - b_1 S_t MC_{1t} - Q_{2t}\right) \\
Q_{2t} &= \frac{1}{2}\left(b_0 S_t - b_1 S_t MC_{2t} - Q_{1t}\right)
\end{aligned}$$

Solving for Q_{it} yields

$$
\begin{aligned}
Q_{1t} &= \frac{1}{3}\left(b_0 S_t - b_1 S_t(2MC_{1t} - MC_{2t})\right) \\
Q_{2t} &= \frac{1}{3}\left(b_0 S_t - b_1 S_t(2MC_{2t} - MC_{1t})\right)
\end{aligned}
$$

$$
\begin{aligned}
Q_t &= Q_{1t} + Q_{2t} \\
&= \frac{S_t}{3}\left(2b_0 - b_1(MC_{1t} + MC_{2t})\right)
\end{aligned}
$$

We have variable profits for each firm

$$
\begin{aligned}
VP_{it} &= P_t Q_{it} \\
&= \left(\frac{1}{b_1}\left(b_0 - \frac{Q_{it} + Q_{jt}}{S_t}\right) - MC_{it}\right) Q_{it} \\
&= \left[\frac{1}{b_1}\left(b_0 - \frac{(2b_0 - b_1(MC_{it} + MC_{jt}))}{3}\right) - MC_{it}\right] \cdot \\
&\quad \frac{1}{3}\left(b_0 S_t - b_1 S_t(2MC_{it} - MC_{jt})\right) \\
&= \left(\frac{b_0}{3b_1} - \frac{2}{3}MC_{it} + \frac{1}{3}MC_{jt}\right)\frac{S_t}{3}\left(b_0 - b_1(2MC_{it} - MC_{jt})\right) \\
&= \frac{S_t}{9b_1}\left(b_0 - 2b_1 MC_{it} + b_1 MC_{jt}\right)\left(b_0 - 2b_1 MC_{it} + b_1 MC_{jt}\right) \\
&= \frac{S_t}{b_1}\left(\frac{b_0 - 2b_1 MC_{it} + b_1 MC_{jt}}{3}\right)^2
\end{aligned}
$$

Using $MC_{it} = c_i + d_i(X_{it} + Y_{it})$, we can expand VP_{it} into three parts as follows.

$$
\begin{aligned}
VP_{it} &= \frac{(b_0 + b_1(c_j - 2c_i))^2}{9b_1} S_t 1\{X_{it} + Y_{it} > 0\} \\
&\quad + \frac{2b_1 d\,(b_0 + b_1(c_j - 2c_i))}{9b_1} S_t\,(2(X_{it} + Y_{it}) - (X_{jt} + Y_{jt})) \\
&\quad + \frac{(b_1 d)^2}{9b_1} S_t\,(2(X_{it} + Y_{it}) - (X_{jt} + Y_{jt}))^2 \\
&= \theta_{0i}^{VP} S_t 1\{X_{it} + Y_{it} > 0\} + \theta_{1i}^{VP} S_t\,(2(X_{it} + Y_{it}) - (X_{jt} + Y_{jt})) \\
&\quad + \theta_{2i}^{VP} S_t\,(2(X_{it} + Y_{it}) - (X_{jt} + Y_{jt}))^2
\end{aligned}
$$

where parameters θ_{0i}^{VP}, θ_{1i}^{VP}, θ_{2i}^{VP} are to be estimated and are functions of b_0, b_1, c_i, and d_i. Specifically, $\theta_{0i}^{VP} \equiv (b_0 + c_j - c_i)^2$, $\theta_{1i}^{VP} \equiv 2d_i(b_0 + c_j - c_i)$, and $\theta_{2i}^{VP} \equiv d_i^2$. This is different from Aguirregabiria's result, where

$$
\begin{aligned}
VP_{it} &= \theta_{0i}^{VP} S_t 1\{X_{it} + Y_{it} > 0\} + \theta_{1i}^{VP} S_t\,((X_{it} + Y_{it}) - (X_{jt} + Y_{jt})) \\
&\quad + \theta_{2i}^{VP} S_t\,((X_{it} + Y_{it}) - (X_{jt} + Y_{jt}))^2
\end{aligned}
$$

which essentially means that firm i's store advantage $(X_{it} + Y_{it}) - (X_{jt} + Y_{jt})$ has the same effect on variable profits regardless of the actual values of X_{it}, Y_{it}, X_{jt}, and Y_{jt}. For example, suppose at the end of time period t, Wal-Mart has five stores and Target has three stores, this store advantage for Wal-Mart would have the same effect on variable profits when Wal-Mart has six stores and Target has four stores, or when Wal-Mart has seven stores and Target has five stores.

The firm's total profits (variable profits minus fixed costs plus market characteristic effects) are

$$\begin{aligned}
\Pi_{it} &= VP_{it} \\
&\quad -\theta_{0i}^{FC} 1\{X_{it}+Y_{it}>0\} - \theta_{1i}^{FC}(X_{it}+Y_{it}) - \theta_{2i}^{FC}(X_{it}+Y_{it})^2 \\
&\quad +\theta_{Denstiy} DENSITY(X_{it}+Y_{it}) + \theta_{GDP} GDP(X_{it}+Y_{it}) \\
&\quad +\theta_{Rent} RENT(X_{it}+Y_{it}) + \theta_{Wages} WAGES(X_{it}+Y_{it}) \\
&\quad -Y_{it}\epsilon_{it} \\
&= Z_{it}(Y_{it,}Y_{jt})\boldsymbol{\theta_i} - Y_{it}\epsilon_{it}
\end{aligned}$$

where ϵ_{it} is firm i's private information shock in its investment cost. It is assumed to be normally distributed and i.i.d. Also, define

$$\begin{aligned}
Z_{it}(Y_{it,}Y_{jt}) \equiv\ & (S_t 1\{X_{it}+Y_{it}>0\},\ S_t(2(X_{it}+Y_{it}) - (X_{it}+Y_{it})), \\
& S_t\left(2(X_{it}+Y_{it}) - (X_{jt}+Y_{jt})\right)^2, \\
& -1\{X_{it}+Y_{it}>0\},\ -(X_{it}+Y_{it}),\ -(X_{it}+Y_{it})^2, \\
& DENSITY,\ GDP,\ RENT,\ WAGES)
\end{aligned}$$

and

$$\boldsymbol{\theta_i} \equiv \left(\theta_{0i}^{VP},\ \theta_{1i}^{VP},\ \theta_{2i}^{VP},\ \theta_{0i}^{FC},\ \theta_{1i}^{FC},\ \theta_{2i}^{FC}, \theta_{Denstiy},\ \theta_{GDP},\ \theta_{Rent},\ \theta_{Wages}\right)$$

where θ_{0i}^{FC}, θ_{1i}^{FC}, θ_{2i}^{FC} are parameters in the fixed cost function. θ_{0i}^{FC} can be considered a lump-sum cost associated with having any positive number of stores in the market and interpreted as an entry cost. The part

$$-\theta_{1i}^{FC}(X_{it}+Y_{it}) - \theta_{2i}^{FC}(X_{it}+Y_{it})^2$$

assumes that fixed costs increase with the number of stores quadratically. Similarly,

$$\theta_{1i}^{VP} S_t\left(2(X_{it}+Y_{it}) - (X_{jt}+Y_{jt})\right) + \theta_{2i}^{VP} S_t\left(2(X_{it}+Y_{it}) - (X_{jt}+Y_{jt})\right)^2$$

assumes that variable profits increase with the number of stores quadratically.

The firms are assumed to play strategies that result in a Markov perfect equilibrium (MPE). The rest of the paper follows the estimation method described in Aguirregabiria, with the modification in the multiplier of firm i's store presence and decision $X_{it} + Y_{it}$. In an MPE, the firm strategies depend only on the current state, which only depends on the variables S_t, X_{1t}, X_{2t}, and ϵ_{it} . Let χ be the set of all possible states $(X_{1t},\ X_{2t})$ and $\boldsymbol{X_t}$ be the common knowledge state variables $(S_t,\ X_{1t}, X_{2t})$. Let $\sigma_i(\boldsymbol{X_t}, \epsilon_{it})$ be the strategy function of firm i. If (σ_1, σ_2) constitutes a MPE then for every firm i, the strategy σ_i maximizes the expected value of firm i at every state given the opponent's strategy. The firm's strategy can be represented in terms of conditional choice probabilities (CCP). Let $P_i(Y_{it} = 1|\boldsymbol{X_t})$ be the probability of firm i opening a store at state $\boldsymbol{X_t}$. This probability can be calculated by integrating the strategy function σ_i over the distribution of the private information ϵ_{it}.

$$P_i(Y_{it} \;\; = \;\; 1|\boldsymbol{X_t}) = \int 1\{\sigma_i(\boldsymbol{X_t}, \epsilon_{it}) = 1\} dG_i(\epsilon_{it})$$

Then the MPE can be represented as a pair of probability functions

$$\{P_1(\boldsymbol{X_t}),\ P_2(\boldsymbol{X_t})\}$$

such that strategy P_i maximizes firm i's payoffs at every state given the opponent's strategy P_j. $P_i(\boldsymbol{X_t})$ is the shorthanded form of $P_i(Y_{it} = 1|\boldsymbol{X_t})$.

As shown before, firm i's total profits are $\Pi_{it} = Z_{it}(Y_{it,} Y_{jt})\boldsymbol{\theta_i} - Y_{it}\epsilon_{it}$. Its expected profits at time t are therefore

$$\begin{aligned} \Pi_{it} &= (1 - P_j(\boldsymbol{X_t}))Z_{it}(Y_{it,}0)\boldsymbol{\theta_i} + P(\boldsymbol{X_t})Z_{it}(Y_{it,}1)\boldsymbol{\theta_i} - Y_{it}\epsilon_{it} \\ &= z_{it}^P(Y_{it})\boldsymbol{\theta_i} - Y_{it}\epsilon_{it} \end{aligned}$$

where

$$z_{it}^P(Y_{it}) \;\; \equiv \;\; (1 - P_j(\boldsymbol{X_t}))Z_{it}(Y_{it,}0) + P(\boldsymbol{X_t})Z_{it}(Y_{it,}1)\boldsymbol{\theta_i}$$

Assuming consumers are forward-looking, with a discount factor of β_i,

$$\tilde{z}_{it}^{P}(Y_{it}) \equiv z_{it}^{P}(Y_{it}) + \left(\sum_{s=1}^{\infty} \beta^{s} z_{it+s}^{P} Y_{it+s} | \boldsymbol{X_t},\ Y_{it}\right)$$

is the sum of expected values of current and future z_{it+s} for $s = 0,\ 1,\ 2, ...$ Similarly,

$$\tilde{e}_{it}^{P}(Y_{it}) \equiv \left(\sum_{s=0}^{\infty} \beta^{s} \epsilon_{it+s} Y_{it+s} | \boldsymbol{X_t},\ Y_{it}\right)$$

is the sum of expected values of current and future private information shocks $\epsilon_{it+s} Y_{it+s}$ for $s = 0,\ 1,\ 2, ...$

Given the opponent's strategy, firm i will open a new store if and only if the profits of doing so is at least as good as not opening one. i.e.

$$\{Y_{it} = 1\} \longleftrightarrow \{z_{it}^{P}(1)\theta_i - \epsilon_{it} \geq z_{it}^{P}(0)\theta_i\}$$

This condition drops Π_i and allows us to estimate parameters θ_i without data on the actual revenues; we only need data on the variables contained in z_{it}^{P}. In terms of CCP, firm i's best response is

$$P_i(Y_{it} = 1 | \boldsymbol{X_{it}}) = G_i\left([z_{it}^{P}(1) - z_{it}^{P}(0)]\theta_i\right)$$

When consumers are forward-looking, the MPE as a vector of CCPs is

$$P_i(Y_{it} = 1 | \boldsymbol{X_{it}}) = G_i\left([\tilde{z}_{it}^{P}(1) - \tilde{z}_{it}^{P}(0)]\theta_i - [\tilde{e}_{it}^{P}(1) - \tilde{e}_{it}^{P}(0)]\right)$$

where $\tilde{z}_{it}^{P}(Y_{it})$ is the expected plus discounted sum of all z_{it}^{P}'s from the current time t. Similarly, $\tilde{e}_{it}^{P}(Y_{it})$ is the expected plus discounted sum of all e_{it}^{P}'s from the current time t. Note that this is a probit model assuming that ϵ_{it} is normally distributed. The MPE as a function of CCPs consists of a system of $N \times K \times m_s$ equations, where N is the number of players, K is the number of actions a player can take at time t, and m_s is the number of states. Pesendorfer and Schmidt-Denglers (2008) explain that in any MPE, the probability vector P satisfies $P_i(Y_{it} = 1|\boldsymbol{X_{it}}) = G_i\left([\tilde{z}_{it}^{P}(1) - \tilde{z}_{it}^{P}(0)]\theta_i - [\tilde{e}_{it}^{P}(1) - \tilde{e}_{it}^{P}(0)]\right)$. Conversely, any P that satisfies this equation can be extended to represent a MPE. They further show that because this equation is a continuous mapping of $[0,1]^{N\times K\times m_s}$ onto itself, by Brouwer's theorem, it has a fixed point P. This P is a vector of CCPs that satisfies this equation and it can be extended to construct a decision rule that constitutes a Markov perfect equilibrium. The caveat is that there may be multiple equilibrium. This issue is addressed in the next section on estimation.

2.4 Estimation

The estimation method used is the nested pseudo-likelihood (NPL) method proposed by Aguirregabiria and Mira (2007). This is a two-step estimator. It is chosen because traditional one-step numerical estimation such as nested-fixed point maximum likelihood is computationally demanding. The two key assumptions for using NPL are "no unobserved common knowledge variables" and "single equilibrium" in the data. In the NPL, for any vector of the CCPs, the pseudo log-likelihood function is:

$$\begin{aligned}&Q(\theta,\mathbf{P})\\&=\sum_{m=1}^{M}\sum_{i=1}^{2}\sum_{t=1}^{T}Y_{imt}lnG_i\left([\tilde{z}_{imt}^{P}(1)-\tilde{z}_{imt}^{P}(0)]\theta_i-[\tilde{e}_{imt}^{P}(1)-\tilde{e}_{imt}^{P}(0)]\right)\\&+(1-Y_{imt})\ln\left(1-G_i\left([\tilde{z}_{imt}^{P}(1)-\tilde{z}_{imt}^{P}(0)]\theta_i-[\tilde{e}_{imt}^{P}(1)-\tilde{e}_{imt}^{P}(0)]\right)\right)\end{aligned}$$

There are CCPs for each player at every state in each market. Each player chooses a best response choice probability based on his belief about the other player's choice probability. A NPL fixed point estimator is defined as a pair $(\hat{\theta},\hat{\mathbf{P}})$ that satisfies two conditions:

$$\hat{\theta} = \mathrm{argmax}_{\theta} Q(\theta,\hat{\mathbf{P}})$$

and

$$\begin{aligned}\hat{P}(\mathbf{X_t}) &= \Phi(\tilde{z}^P_{imt}, \hat{\theta}_i) \\ &= G_i\left([\tilde{z}^P_{imt}(1) - \tilde{z}^P_{imt}(0)]\theta_i - [\tilde{e}^P_{imt}(1) - \tilde{e}^P_{imt}(0)]\right)\end{aligned}$$

The estimation algorithm is as follows: Start with a nonparametric estimator of $\hat{\mathbf{P}}^{\mathbf{0}}$ as the initial guess for the CCPs of each player. This initial guess can be a vector of random values, or the frequency of entry calculated from the data as using

$$\hat{\mathbf{P}}^{\mathbf{0}}_{\mathbf{i}}(\mathbf{X_t}) = \frac{\sum_{m=1}^{M}\sum_{t=1}^{T} Y_{imt} 1\{\mathbf{X}_{mt} = \mathbf{X}\}}{\sum_{m=1}^{M}\sum_{t=1}^{T} 1\{\mathbf{X}_{mt} = \mathbf{X}\}}$$

This initial guess is used to calculate $\hat{\theta}^0 = \text{argmax}_\theta Q(\theta, \hat{\mathbf{P}})$, which is a probit model. $\hat{\theta}^0$ is then used to update the CCP vector by computing

$$\hat{P}^1(\mathbf{X_t}) = G_i\left([\tilde{z}^P_{imt}(1) - \tilde{z}^P_{imt}(0)]\theta_i - [\tilde{e}^P_{imt}(1) - \tilde{e}^P_{imt}(0)]\right)$$

This $\hat{P}^1(\mathbf{X_t})$ is subsequently used in the next iteration to find $\hat{\theta}^{\mathbf{1}}$. This process is repeated until $\hat{P}(\mathbf{X_t})$ converges. i.e. When $|\hat{\theta}^K - \hat{\theta}^{K-1}| < \epsilon$, where ϵ is the criterion for convergence. To account for the possibility of multiple equilibria, several initial values of $\hat{\mathbf{P}}^{\mathbf{0}}$ are used, including a vector of zeros, vector of ones, and the CCPs calculated from the data. All these initial guesses generated the same estimates $\hat{\theta}$ upon convergence.

2.5 Data

The two retailers of interest are Wal-Mart and Target. Each market is defined to be a city, and the dataset is panel of 614 markets in Georgia. The data used to estimate the model is from 2002 to 2006. The maximum number of stores for either firm in any market is five, so the state space is $\chi = \{0, 1, 2, 3, 4, 5\} \times \{0, 1, 2, 3, 4, 5\}$. Two main elements of data are used. The first element is store-level data, which includes the number of stores for both retailers existing by the end of each year and whether any retailer entered the market by opening new stores during that year. The data set includes data on all the store locations and opening dates of Wal-Mart and Target since they opened their first stores in 1962. This model assumes that firms open new store and do not close any for tractability. This assumption is supported by the data. For Wal-Mart, this simplifying assumption is also used in Holmes' paper (2011), which states that Wal-Mart rarely closes stores and that Wal-Mart's annual reports disclose store closings that are on the order of two per year in the U.S. For Target, a similar behavior is observed from its annual reports. Data from the annual reports show that they close an average of three stores without replacement per year countrywide. An average of 17 stores were closed across the country and relocated in the same trade area during 2002 and 2006, but those are not considered different in this model.

The second element of the data is market characteristics, including population density, GDP per capita, average rent, and average wages. Population densities at the city level are obtained by dividing population from 2002 to 2006 by the area of the city. GDP per capita is obtained for 2013 for each county and used for interpolating using state GDP growth index to find the GDP per capita from 2002 to 2006. For most markets, city-level GDP per capita is not available, so county GDP per capita is used. City-level GDP per capita is only available for major cities such as Pittsburgh and Philadelphia. Average rent is obtained from various sources, mainly LoopNet.com, a commercial real estate website that provides retail property rent data at the city, county, metropolitan, and state level. Average rent (per sq. mile per month) is obtained for 2013 for each city and then interpolated using a quarterly index for U.S. northeast retail property value, provided by CoStar, a leading provider of commercial real estate information, analytics and marketing services. Average weekly wages data is obtained from U.S. Department of Labor, which publishes historical quarterly wages data by county.

Table 2.1 presents the total number of stores for Wal-Mart and Target in Georgia from 2002 to 2006. During this period, both firms seemed that they wanted to match their rival's expansion and consequently there were a similar number of new stores opened in Georgia each year. Table 2.2 presents the summary of market characteristics for the markets or cities in Georgia from 2002 to 2006.

Table 2.1: Total Number of Stores for Wal-Mart and Target in Georgia

Year	Wal-Mart	Target
2002	98	30
2003	103	34
2004	103	36
2005	107	38
2006	114	42

Table 2.2: Summary of Market Characteristics

Variable	Mean	Max	Min	SD
Market Size (Number of people)	6549.21	392647.89	20.98	21530.35
Distance to Distribution Center for Wal-Mart (mile)	49.2	118.68	1.86	25.97
Distance to Distribution Center for Target (mile)	141.32	275.27	1.24	59.86
Population Density (1000's people/sq. mile)	0.71	6.46	0.01	0.77
GDP Per Capita ($)	16401.45	84943.31	2325.31	7395.53
Monthly Rent ($/sq. mile)	13.99	27.01	5.27	9.37

Table 2.3 presents the number of markets under different market structures. Wal-Mart operated under monopoly in over twenty times as many as market as Target did in 2002, and the advantage remained consistent throughout the period. An interesting observation in the data is that Wal-Mart seemed to open new stores in cities with smaller population (and lower population density) than Target did; Wal-Mart seemed comfortable expanding towards smaller markets while Target did not. Table 2.4 shows the number of markets where each firm has store advantage given that at least one firm exists in that market. Clearly, Wal-Mart is the dominant firm. There is also a significant portion of the markets with equal number of stores for each firm.

Table 2.3: Number of Markets under Different Market Structures

Year	Wal-Mart Only	Target Only	None	Both
2002	64	3	522	25
2003	65	2	518	29
2004	64	3	517	30
2005	64	4	515	31
2006	66	4	511	33

Table 2.4: Number of Markets Based on Store Advantage

Year	Wal-Mart>Target	Wal-Mart=Target	Wal-Mart<Target
2002	68	21	3
2003	69	24	3
2004	69	24	4
2005	70	24	5
2006	75	23	5

2.6 Results

Table 2.5 and 2.6 shows the estimates of the dynamic entry competition model between Wal-Mart and Target. $\theta_{Density}$, as expected, is positive as total profits would increase with greater population density. θ_{Rent} is positive for Target and negative for Wal-Mart, but it is not significant. θ_{GDP} and $\theta_{Distance}$ have small effects on the total profits of either firm. The effects of store number and entry decision on variable profits and fixed costs vary depending on the number of stores for each firm because the functional form is assumed to be quadratic. Most of the estimates for variable profits and fixed costs are significant at the 5% or 10% confidence level. The estimates for market characteristics are mostly insignificant but also have very small magnitudes.

Table 2.5: Estimates for Wal-Mart

Parameter	Estimate	Standard Error	t-Stat
θ_0^{VP}	4.504	1.442	3.123
θ_1^{VP}	0.001	0.183	0.005
θ_2^{VP}	0.040	0.055	0.725
θ_0^{FC}	0.086	0.051	1.686
θ_1^{FC}	0.144	0.085	1.703
θ_2^{FC}	0.017	0.018	0.981
$\theta_{Distance} \times Store$	0.002	0.000	-0.738
$\theta_{Density} \times Store$	0.010	0.010	1.004
$\theta_{GDP} \times Store$	0.001	0.001	0.240
$\theta_{Rent} \times Store$	-0.003	0.004	-0.776

Table 2.6: Estimates for Target

Parameter	Estimate	Standard Error	t-Stat
θ_0^{VP}	-12.152	8.457	-1.437
θ_1^{VP}	7.885	4.176	1.888
θ_2^{VP}	-0.663	0.351	-1.890
θ_0^{FC}	-4.976	3.122	-1.594
θ_1^{FC}	6.317	3.663	1.725
θ_2^{FC}	-0.853	0.522	-1.636
$\theta_{Distance} \times Store$	0.002	0.000	1.447
$\theta_{Density} \times Store$	0.010	0.014	0.707
$\theta_{GDP} \times Store$	0.001	0.002	0.177
$\theta_{Rent} \times Store$	0.008	0.008	1.073

Given the market characteristics, these estimates can be used to estimate the total profits for each firm under different market structures. The calculations can help firms make decisions by predicting the profits under different market structures and how profits would change given the rival's decisions. For example, Figure 2.1 shows the total profits normalized to the monopolist profit of one store for each firm under different market structures in Atlanta. Wal-Mart's profits increase as itself adds more stores, and decrease slightly as its rival, Target, adds more stores. A similar trend is observed for Target, but the decrease in profits is much greater in magnitude relative to the changes in Wal-Mart's profits. Also, Wal-Mart makes positive profits under all market structures, while Target must have more stores than Wal-Mart to make positive profits. The reason could be Wal-Mart's greater economies of scale that prevent it from being greatly affected by Target's presence in all market structures. This analysis can be run for any market under any market characteristics to shed light on how a specific market structure evolves over time and effects of firm entries on the profits on both firms.

Figure 2.1: Normalized Total Profits for Wal-Mart in Atlanta

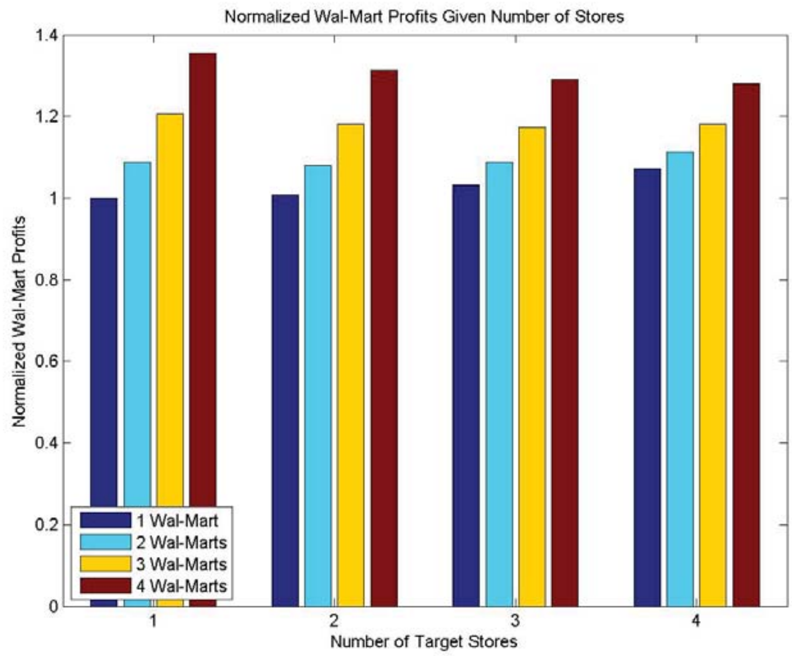
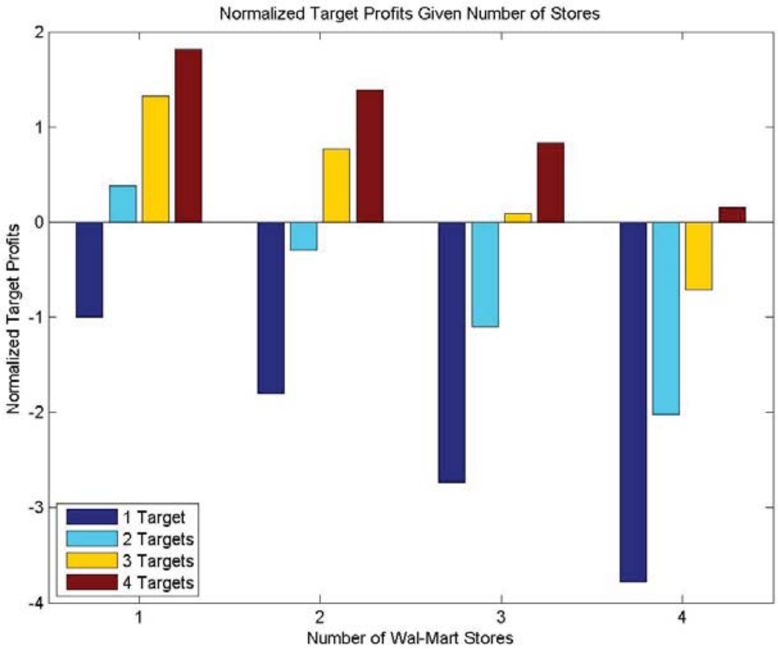


Figure 2.2: Normalized Total Profits for Target in Atlanta



2.7 Conclusion and Future Research

This paper shows how total profits, variable costs, and fixed costs of two main competitors in the retailing industry, Wal-Mart and Target, depend on the market structure and level of competition. Wal-Mart generally dominates the competition in all market structures; it is affected slightly by Target's presence. On the contrary, Target relies on having store advantage to make positive profits. This could also be the case for other smaller retailers and discount stores. These stores can only profit when they maintain a store advantage over the leader. This result can potentially help policy makers in charge of regulations on large retail establishments and store location, and zoning to maintain fair competition.

One limitation of this paper is that it only includes two competitors. Wal-Mart and Target are generally considered the two main competitors in retailing, but in other industries there may be more than two competitors. For example, in the fast food industry, the main players include at least McDonald's, KFC, and Burger King. It would insightful to extend this model into other industries with more than two competitors. However, modeling more than two competitors in a dynamic setting would be a challenge as the number of states increase exponentially due to the curse of dimensionality. Another limitation is that the model assumes that entry decisions are independent across markets; the presence of stores in other markets are assumed to be exogenous. In reality, there may be spillover effects from neighboring markets. Finally, the model assumes a market to be a city. There are advantages and disadvantages. The advantage is that there is less variance in the market characteristics compared to markets defined using counties (Jia, 2008) and states (Zhu, 2009). The disadvantage is that smaller markets may be affected more by spillover effects from nearby markets. Market definition varies, and there has been no consensus on the ideal market definition for entry games in retailing. In the future, one possibility could be to use data on consumer shopping behavior to infer consumer preferences for a more realistic market definition.

References

[1] Ailawadi, Kusum L., and Scott A. Neslin. "The Effect of Promotion on Consumption: Buying More and Consuming it Faster." Journal of Marketing Research 35.3 (1998): 390-8.

[2] Anderson, Rod W., and Ronald N. Johnson. "Antitrust and Sales-Below-Cost Laws: The Case of Retail Gasoline." Review of Industrial Organization 14.3 (1999): 189-204.

[3] Aguirregabiria, Victor and Pedro Mira, "Sequential Estimation of Dynamic Discrete Games", Econometrica 75 (2007), pp. 1-53.

[4] Bachmeier, Lance J., and James M. Griffin. "New Evidence on Asymmetric Gasoline Price Responses." The Review of Economics and Statistics 85.3 (2003): 772-776.

[5] Bell, David R., Jeongwen Chiang, and V. Padmanabhan. "The Decomposition of Promotional Response: An Empirical Generalization." Marketing Science 18.4 (1999): 504-26.

[6] Berry, Steven, James Levinsohn, and Ariel Pakes. "Automobile Prices in Market Equilibrium." Econometrica 63.4 (1995): 841-890.

[7] Borenstein, Severin, A. Colin Cameron, and Richard Gilbert. "Do Gasoline Prices Respond Asymmetrically to Crude Oil Price Changes?" The Quarterly Journal of Economics 112.1 (1997): 305-39.

[8] Brannon, J. Isaac. "The effects of resale price maintenance laws on petrol prices and station attrition: empirical evidence from Wisconsin," Applied Economics, vol. 35(3), (2003): 343-349.

[9] Bresnahan, Timothy F. "Competition and Collusion in the American Automobile Industry: The 1955 Price War." The Journal of Industrial Economics 35.4 (1987): 457-482.

[10] Bresnahan, Timothy F., and Reiss, Peter C. "Entry and competition in concentrated markets." Journal of Political Economy, 99(5), (1991): 977-1009.

[11] Browning, Martin, and M. Dolores Collado. "The response of expenditures to anticipated income changes: Panel data estimates." The American Economic Review 91 (3), (2001): 681-92.

[12] Byrne, David P. and Nicolas de Roos. "Search and Stockpiling in Retail Gasoline Markets." Working paper (2015).

[13] Chiang, Jeongwen. "A Simultaneous Approach to the Whether, what and how Much to Buy Questions." Marketing Science 10.4 (1991): 297-315.

[14] Ching, Andrew, Tulin Erdem, and Michael P. Keane. "A simple method to estimate the roles of learning, inventories and category consideration in consumer choice." Journal of Choice Modeling 13 (2014): 60-72.

[15] Cooper, Thomas E., and John T. Jones. "Asymmetric Competition on Commuter Routes: The Case of Gasoline Pricing." Southern Economic Journal 74.2 (2007): 483-504.

[16] Collard-Wexler, Allan. "Mergers and Sunk Costs: An Application to the Ready-Mix Concrete Industry." American Economic Journal: Microeconomics, 6(4), (2014): 407-47.

[17] Erdem, Tulin, Susumu Imai, and Michael P. Keane. "Brand and Quantity Choice Dynamics Under Price Uncertainty." Quantitative Marketing and Economics 1.1 (2003): 5-64.

[18] Gupta, Sunil. "Impact of Sales Promotions on when, what, and how Much to Buy." Journal of Marketing Research 25.4 (1988): 342-55.

[19] Heath, Chip, and Jack B. Soll. "Mental Budgeting and Consumer Decisions." Journal of Consumer Research 23.1 (1996): 40-52.

[20] Hendel, Igal, and Aviv Nevo. "Measuring the Implications of Sales and Consumer Inventory Behavior." Econometrica 74.6 (2006): 1637-73.

[21] Hendel, Igal, and Aviv Nevo. "Sales and Consumer Inventory." The RAND Journal of Economics 37.3 (2006): 543-61.

[22] Holmes, Thomas J. "The Diffusion of Wal-Mart and Economies of Density." Econometrica 79.1 (2011): 253-302.

[23] Houde, Jean-François. "Spatial Differentiation and Vertical Mergers in Retail Markets for Gasoline." The American Economic Review 102.5 (2012): 2147-82.

[24] Jia, Panle. "What Happens when Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry." Econometrica 76.6 (2008): 1263-316.

[25] Lewis, Matthew, and Michael Noel. "The Speed of Gasoline Price Response in Markets with and without Edgeworth Cycles." Review of Economics and Statistics 93.2 (2011): 672-82.

[26] Lewis, Matthew S. "Temporary Wholesale Gasoline Price Spikes have Long-Lasting Retail Effects: The Aftermath of Hurricane Rita." Journal of Law and Economics 52.3 (2009): 581-605.

[27] Parker, Jonathan A. "The Reaction of Household Consumption to Predictable Changes in Social Security Taxes." The American Economic Review 89.4 (1999): 959-73.

[28] Pesendorfer, Martin, and Philipp Schmidt-Dengler. "Asymptotic Least Squares Estimators for Dynamic Games." The Review of Economic Studies 75.3 (2008): 901-28.

[29] Poterba, James M. "Are Consumers Forward Looking? Evidence from Fiscal Experiments." The American Economic Review 78.2 (1988): 413-8.

[30] Sen, Anindya. "Does Increasing the Market Share of Smaller Firms Result in Lower Prices? Empirical Evidence from the Canadian Retail Gasoline Industry." Review of Industrial Organization 26.4 (2005): 371-89.

[31] Shepard, Andrea. "Contractual Form, Retail Price, and Asset Characteristics in Gasoline Retailing." The RAND Journal of Economics 24.1 (1993): 58-77.

[32] Slade, Margaret E. "Strategic Motives for Vertical Separation: Evidence from Retail Gasoline Markets." Journal of Law, Economics, & Organization 14.1 (1998): 84-113.

[33] Zhu, Ting, Vishal Singh, and Mark D. Manuszak. "Market Structure and Competition in the Retail Discount Industry." Journal of Marketing Research 46.4 (2009): 453-66.

[34] Zimmerman, Paul R. "The Competitive Impact of Hypermarket Retailers on Gasoline Prices." Journal of Law and Economics 55.1 (2012): 27-41.

[35] Zimmerman, Paul R., John M. Yun, and Christopher T. Taylor. "Edgeworth Price Cycles in Gasoline: Evidence from the United States." Review of Industrial Organization 42.3 (2013): 297-320.