UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

AN EMPIRICAL ANALYSIS OF RETAIL MARKET STRUCTURE: LOCATION, COMPETITION, AND CONSUMER OUT-SHOPPING

A DISSERTATION

SUBMMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

SI CHEN
Norman, Oklahoma
2020

AN EMPIRICAL ANALYSIS OF RETAIL MARKET STRUCTURE: LOCATION, COMPETITION, AND CONSUMER OUT-SHOPPING

A DISSERTATION APPROVED FOR THE
MICHAEL F. PRICE COLLEGE OF BUSINESS

BY THE COMMITTEE CONSISTING OF

Dr. Charles Ingene, Chair

Dr. Le Wang

Dr. Krish Muralidhar

Dr. Qiong Wang

Acknowledgements

I would like to express my deepest and sincere gratitude to the members of my dissertation committee: Dr. Charles Ingene, Dr. Le Wang, Dr. Krish Muralidhar, and Dr. Qiong Wang, for their support and help through this long journey. This dissertation would not have been possible without their generous help. Special thanks to my committee chair Dr. Charles Ingene, who is also my advisor, for his continued guidance, encouragement, and support. I learned from him what makes an excellent scholar.

I would also like to thank Dr. Rajiv Dant, who led me into academia, trained and mentored me how to be a good scholar. He will be forever missed. May his soul continue to rest in peace. In addition, I want to thank the professors who have helped me during my Ph.D. study: Dr. Pravin Nath, Dr. Kevin Grier, Dr. Hairong Song, and all the faculty members in the Division of Marketing & Supply Chain Management.

I would like to extend my thanks to my colleagues and friends – Dr. BJ Allen, Dr. Ola Seriki, and Yi Zheng, for their support and all the cheers and tears we shared during this long journey.

Finally, I wish to thank my parents for their understanding, support, and unconditional love. I could not have completed this dissertation without their support.

Table of Contents

List of Tables

List of Figures

Abstract

In this dissertation, I examine the U.S. retail market structure (*i.e.*, retail expenditure per household) across seven retail subsectors (*i.e.*, nine lines of retail trade), using data of 735 metropolitan and micropolitan areas obtained from 2012 Economic Census and 2012 American Community Survey (5-year estimates). I have shown that consumers' socioeconomic and demographic characteristics, two marketing mix variables (*i.e.*, quantity and quality of service), intertype and intratype competition, and consumers' out-of-town shopping behavior significantly influence retail expenditures across various lines of retail trade.

Specifically, the two service variables significantly increase household expenditures at all nine lines of retail trade. Moreover, with adequate and dedicated service employees, retailers can mitigate the negative impact incurred from local competitors as well as competitors located in nearby areas. Also, I have shown that key variables driving retail expenditures are slightly different in metropolitan vs. micropolitan areas. More importantly, I have examined consumers' out-of-town shopping behavior. Results indicate that consumers do out-of-town shopping for various lines of retail trade, including new cars, furniture, electronics and appliances, clothing, traditional department stores, discount department stores, and warehouse clubs and supercenters.

This dissertation contributes to the marketing literature and provides insights to retailers on decisions about store location, store size, competition, market segmentation, and consumer out-shopping.

Chapter 1 Introduction

This chapter presents the motivation and general background of this study, and an overview of the research. The focus of this dissertation is U.S. retail market structure and its possible influencers – consumers' socioeconomic and demographic characteristics, competitive forces, retailers' marketing strategies, and consumers' out-shopping behavior. The U.S. retail trade sector encompasses both brick-and-mortar businesses and e-commerce, but e-commerce shops do not occupy a physical presence, nor do they have concerns operating in a geographical area. In this dissertation, I focus on the retail market structure of brick-and-mortar businesses as they account for a large proportion of all U.S. retail sales.

1.1 Motivation of the Study

As defined by Bucklin (1972, p. 66), retail structure is "the manner in which the sale of a product is organized in a geographical market area". Retail structure has been empirically defined as the number of stores, the number of stores per capita, total sales, average sales per store, or average sales per capita or sales per household (Bucklin 1972; Ingene 1983). In this dissertation, I use sales per household (*i.e.*, household expenditure)[1] to define retail structure, for two reasons. First, households are the purchasing units for most consumer goods. Second, as argued by Ferber (1958), a big population always leads to a big sales revenue. Therefore, it is more reasonable and meaningful to examine the expenditure pattern of an average household. Furthermore, analyzing where and how an average household spends their money will provide insights for retailers as how they could tailor their marketing efforts to target these households.

[1] "Retail sales" and "household expenditures" are used interchangeably.

Understanding retail market structure is important for both theoretical and practical reasons. According to the 2012 Economic Census, the retail trade sector generates a total sales revenue of 4.22 trillion U.S. dollars, which accounts for 26.05% of 2012 U.S. GDP[2]. There are 14.7 million retail employees[3], which accounts for more than 10% of 142 million employed civilian population. It is the third largest work force by industry in the U.S., following education & health services, and professional & business services (Current Population Survey 2012, U.S. Bureau of Labor Statistics).

Retail sales revenues have been shown to vary across geographic areas in different lines of retail trade (e.g., Ingene & Yu 1981). Key determinants of retail sales examined in the literature include consumers' socioeconomic and demographic characteristics, competitive forces from retail industries that sell substitutable products, and managerial decisions on the marketing mix (e.g., Ingene 1983; Ingene & Brown 1987). Although the U.S. retail structure has been examined and the extant literature has shown strong evidence on the set of key determinants of retail sales, it lacks a few critical components that are worth investigating and reexamining.

First, the most recent literature regarding U.S. retail structure dated back to Miller et al. (1999) which examined the structure of sporting goods industry with data from the 1992 Economic Census of the retail trade sector. It is imperative for scholars and practitioners to understand if the key determinants of retail sales examined in the literature still hold true over twenty years, as it is certain that changes in the U.S. retail structure have occurred over time.

Second, retail trade has always been a competitive sector. However, competition has only been empirically examined in two studies regarding cross-industry comparisons with U.S. data

[2] The World Bank reports that U.S. GDP in 2012 was 16.197 trillion U.S. dollars.
[3] This figure of 14.7 million retail employees is from the 2012 Economic Census, and it is the number of employees paid in the first quarter of 2012.

(Ingene 1983; Miller et al 1999). It is most vital for practitioners to understand where competition forces come from and how to effectively respond to those forces.

Third, given the limitation and availability of secondary data in the nineteenth century, data on micropolitan areas was not available at all, thus, retail sales have only been examined at the city, state, and metropolitan area levels. According to the 2012 Economic Census, there are 381 Metropolitan Statistical Areas (MSAs) and 536 Micropolitan Statistical Areas[4] in the U.S. Retail establishments in the 536 micropolitan areas occupy 107,083 retail stores, generate nearly 341 billion dollars of sales revenue, and employ about 1.3 million retail workers. The micropolitan areas altogether contribute 8.5% of all retail sales, and they are a crucial and integral part of the country. It is important for both scholars and practitioners to understand if the key determinants of retail sales can be generalized to micropolitan areas.

Fourth, the extant literature has treated each geographic area as a “closed system”, which means retail sales of a geographic area only comes from residents who live and work in the area. A recent study by Chen et al. (2020) demonstrated that influx of population from nearby cities has a significant impact on sales at food and beverage stores in the focal city. It is possible that consumers’ geographic movement would impact the retail sales revenue of the focal area when there exists a population inflow or outflow to the focal area. Intuitively, residents in small areas would travel to nearby large areas for shopping purposes. Therefore, it is practically important to understand to what extent large areas draw sales away from surrounding small areas, or in other words, to what extent consumers do out-of-town shopping.

[4] “Metropolitan Statistical Areas have at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties. Micropolitan Statistical Areas have at least one urban cluster of at least 10,000 but less than 50,000 population, plus adjacent territory that has a high degree of social and economic integration with the core as measured by commuting ties.” (OMB Bulletin No. 18-04)

To address the abovementioned issues, I plan to examine the following research questions in this dissertation:

1. How does competition (intertype and intratype competition) influence retail expenditures per household?
2. What are the key variables driving retail expenditures per household across different lines of retail trade?
3. Do key variables driving household expenditures differ in metropolitan area vs. micropolitan areas?
4. Do large populated areas draw retail sales (*i.e.*, expenditure) from nearby small areas?

In this dissertation, I hope to examine the abovementioned research questions, provide a thorough examination of the U.S. retail market structure, and foreshadow the future of U.S. retailing.

1.2 General Background of the Study

In this dissertation, I use the 2012 Economic Census data to examine the retail trade sector in the United States. I focus on the U.S. retail market structure for two reasons. First, the United States represents the strongest economy in the world, and the retail trade sector is an important share of the U.S. economy. Any changes in the retail trade sector will likely lead to a significant impact on the U.S. economy. Second, the U.S. retail trade sector has received minimal attention in the marketing literature in the past twenty years (Chen et al., 2020). Considering that over 90% of all retail sales comes from brick-and-mortar retailers and less than 10% from online shopping in 2012[5], I focus on brick-and-mortar retailing due to its importance

[5] In 2012, the Economic Census reports that the retail trade sector has a total sales revenue of 4.22 trillion U.S. dollars, among which 9.14% comes from Nonstore Retailers (*i.e.*, electronic shopping, mail orders, vending machines, etc.), and the rest 90.86% comes from brick-and-mortar businesses.

in the U.S. retail sector and U.S. economy. Analyzing brick-and-mortar retailing in the U.S. not only enables us to understand a major dimension of the U.S. economy, but also lays a foundation for comprehending the factors that drive retail sales in the country.

Specifically, in this dissertation, I examine the factors that affect retail sales revenue (*i.e.*, household expenditure) in nine major lines of retail trade (*i.e.*, motor vehicle, furniture, electronics, food & beverage, gasoline, clothing, traditional department stores, discount department stores, and warehouse clubs and supercenters), using the 2012 U.S. Economic Census data on metropolitan and micropolitan areas. Previous research on household expenditures shows that a set of socioeconomic, demographic, and competitive factors, along with marketing mix variables are key determinants of household expenditures (*e.g.*, Ingene and Brown 1987, Chen et al. 2020). Building upon the extant literature, I derive a set of socioeconomic, demographic, marketing mix (*i.e.*, service), and competitive variables (*i.e.*, intertype and intratype competition across retail lines) that are applicable to this research study, given the availability and limitation of data from secondary sources. Also, I construct a variable that incorporates consumers' out-of-town shopping behavior.

This dissertation makes four major contributions to the literature. First, I consider and analyze both intertype and intratype competition among multiple retail lines. Second, I re-examine the set of determinants of retail sales derived from the literature with the most recent Economic Census data. Third, I examine retail sales at both metropolitan and micropolitan areas, which has never been done in the literature due to data limitation. Fourth, I examine the extent to which retailers in large areas draw sales from nearby small areas (*i.e.*, consumer's out-shopping behavior). These additions to the marketing literature not only help scholars and practitioners understand the dynamic nature of expenditure patterns at major lines of retail trade in the U.S., but also provide managerial insights to retailers regarding consumer characteristics, market segmentation, potential competitions, store locational decisions, and store employee management.

1.3 Overview of Research

The remainder of this dissertation is organized as follows. Chapter 2 presents a review of related literature on retail sales and its determinants. In Chapter 3, I synthesize a general theory that explains the factors driving retail sales. Chapter 4 provides an overview of the U.S. retail trade sector and the nine subsectors/industries examined in this dissertation. The description on research data and research methodology are presented in Chapter 5. Chapter 6 presents the empirical results. In Chapter 7, I discuss the results, theoretical contributions, and managerial implications. Chapter 8 concludes this dissertation with its limitations and avenues for future research.

Chapter 2 Literature Review

This chapter presents a review of the literature related to retail market structure and its determinants in the marketing domain. This dissertation is only concerned with retail sales to which geographical location is of most importance, although there are stores occupying non-spatial space (*i.e.*, electronic shopping & mail-order houses). Locational decisions are critical to retailers. First, a consumer is only willing to travel certain distance to shop at a brick-and-mortar store, because travel incurs both economic and temporal costs to the consumer. If a store is not located within the travel range, it may not be so attractive to consumers. Second, a store's external environment varies spatially. Consumers' socioeconomic and demographic characteristics differ in different geographic areas, as do their shopping preferences. Therefore, site selections affect the number of customers that a store may draw. In other words, demand varies geographically, which directly influences the size of a store and its product assortment, thus a store's sales revenue in the area. Third, operating costs (*i.e.*, leasing, utilities, taxes) vary geographically. Fourth, competition differs. In summary, location affects a store's potential profitability in the area.

Theoretically, the foundation of this literature is Reilly's *Law of Retail Gravitation* published in 1931 and Christaller's *Central Place Theory* published in 1933. Scholars in this research stream have attempted to extend their work in various ways. For example, Converse (1949) and Huff (1964) not only applied Reilly's gravity model, but also adjusted it to determine the division of sales between areas, communities, and even shopping centers. In the 1950s, scholars such as Reynolds (1953), Russell (1957), Ferber (1958), and others started to investigate retail patterns in American cities. In the 1980s, Ingene developed a series of works on the determinants of retail sales and added marketing mix variables as a new set of determinants in his studies (e.g., Ingene & Lusch 1980; Ingene & Yu 1981), which had never been done by previous studies in the literature. Ingene's works later became the theoretical foundation for many studies including this dissertation. His contribution to the literature was considered a milestone in the development of the retailing literature.

Among the abovementioned research studies, scholars examined different sets of determinants of retail sales empirically, however, most of them have ignored one important factor of sales – consumers who shop out-of-town. In other words, previous studies primarily focused on local shoppers, local competitions, and their influences on retail sales. A recent study by Chen et al. (2020) incorporated the concept of "consumer-out-shopping" in their study and showed that population from out-of-town affects retail sales at the focal city. Inspired by their research, I build upon the extant literature and extend it to consumer out-shopping. Figure 2.1 presents the theoretical model developed and examined in this dissertation.

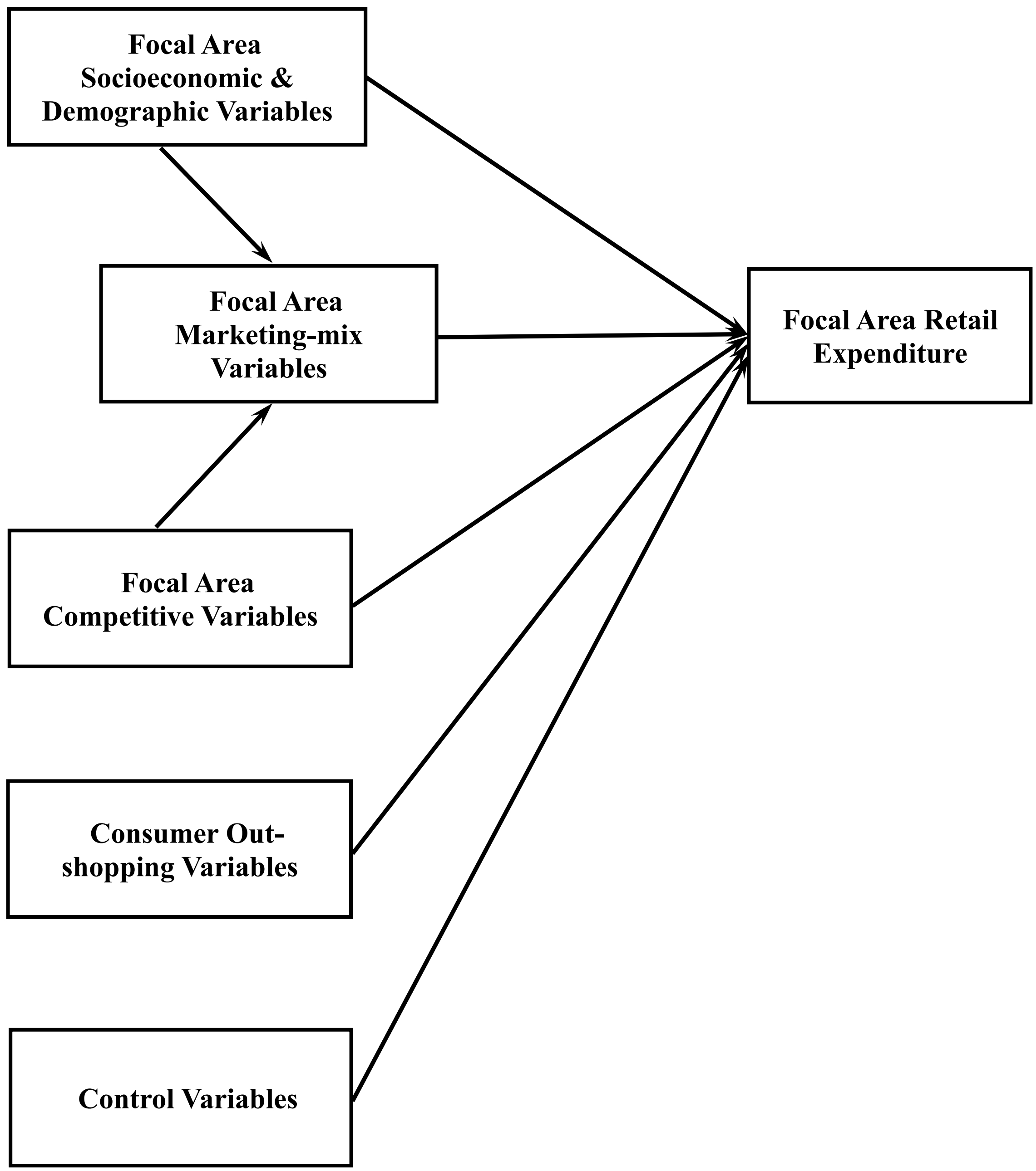

Figure 2.1. Theoretical Model Developed and Examined in This Dissertation

This chapter is arranged as follows. First, I discuss the foundation of this literature - retail gravity model. Second, I review the literature on the determinants of retail sales. Third, I extend the literature to consumer out-shopping.

2.1 Retail Gravitation

Reilly's "*The Law of Retail Gravitation*" is the earliest marketing literature that addressed the issue of retail sales in a geographic market (Reilly 1931). Reilly applied the gravitation law to explain the flow of retail trade among cities. He stated that two cities draw sales from any intermediate city or place in direct proportion to the population of the two cities and in inverse proportion to the square of their distances from the intermediate city. This is expressed mathematically as follows:

$$\frac{B_a}{B_b} = \left(\frac{P_a}{P_b}\right)\left(\frac{D_b}{D_a}\right)^2$$

where B_a is the proportion of the retail trade from the intermediate city drawn by city A; B_b is the proportion of the retail trade from the intermediate city drawn by city B; P_a is the population of city A; P_b is the population of city B; D_a is the distance from the intermediate city to city A; and D_b is the distance from the intermediate city to city B.

This idea is illustrated graphically in Figure 2.2.

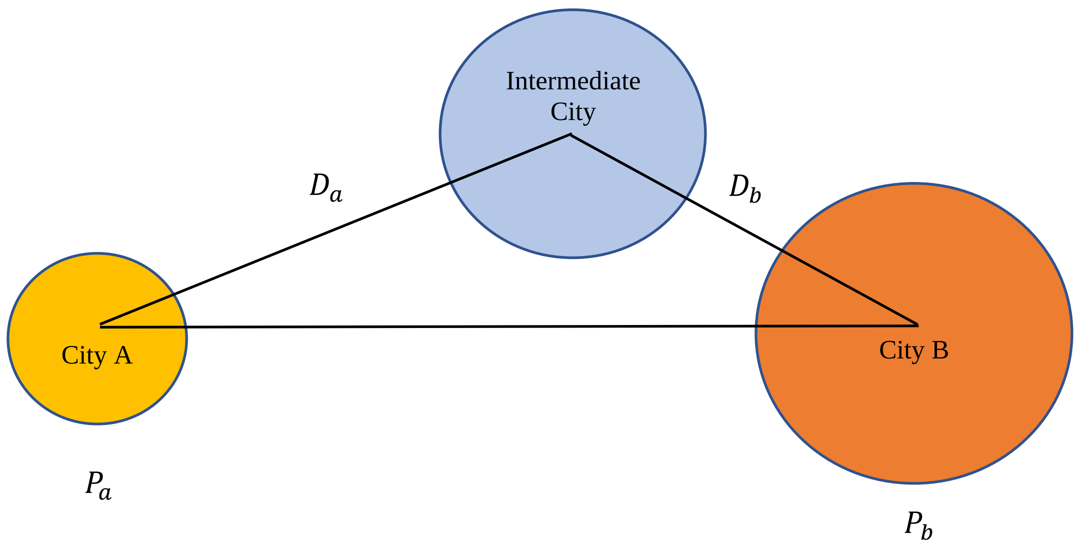

Figure 2.2. Law of Retail Gravitation (Reilly 1931)

The implicit conclusion from Reilly's work is that population and distance are two key explanatory factors of retail sales. More population indicates that potentially there are more goods available at the city, and short distance away from hometown indicates low temporal cost and low dollar cost on transportation (*i.e.*, gas).

Reilly pointed out that other than population and distance, factors such as transportation methods, ways of communication, consumer characteristics, population density, nature of the competition offered by other cities in the surrounding territory, among others, also influence the flow of retail trade from one city to another. More importantly, he noticed that with the change in consumer's mode of transportation (*i.e.*, from horses to automobiles), the out-of-town shopping and trading phenomenon had become popular and nation-wide. In other words, consumers were more willing to travel further for shopping purposes, given the ease and convenience of travelling with automobiles. Thus, retail businesses started to reshape their structure to accommodate this change in consumer shopping behavior. Reilly's point of view was confirmed by Ford (1935), who found due to the change in the transportation mode consumers were able to travel further to large but distant stores, thus the number of stores started to decline in a geographic area since each store was large enough to serve a large body of population in the area. This phenomenon is later referred as the “Ford Effect”. Details regarding this study is discussed in the following section of this chapter.

Reilly’s idea can be further illustrated in Figure 2.3.

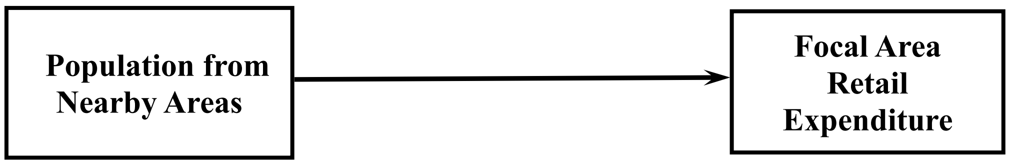

Figure 2.3. Theoretical Model by Reilly (1931)

Reilly’s *Retail Gravitation* provides us with two simple rules.

- Population: the larger the city, the more outside trade it draws, under similar circumstances.
- Distance: a city draws more trade from nearby towns than it does from more distant ones, *ceteris paribus*.

At about the same time, Christaller proposed the *Central Place Theory*. Similar with Reilly’s work, the *Central Place Theory* states that people living in small or rural areas would incur travel costs to shop at large urban areas, because not all goods are available in small areas. He argued that larger areas have more varieties of goods and smaller areas only carry a small amount of varieties, due to the differences in population. Therefore, the variety and range of goods increases as the size of the city. Reilly and Christaller both recognized that population is a proxy for the variety of goods available in a city.

The theoretical insights presented by Reilly and Christaller have influenced a great number of scholars. Thus, they have shaped the way that many scholars think of retail. Later in 1949, Converse attempted to redefine Reilly's mathematical model and proposed the *New Laws of Retail Gravitation*. His contribution lies in redefining the law such that the trading area of each community can be derived and the proportion of retail trade an area retains and the proportion it loses can be both predicted. Again, population and distance were used in the mathematical formula. Furthermore, based on Reilly and Converse's work, Huff (1964) was able to extend the model to determine the trading area of a shopping area within a community. He redefined the mathematical model and derived the trading areas of shopping centers. By combining the factors of shopping center size, consumer's travel time from home to shopping center, and consumer's willingness to travel, Huff derived a trading area.

Among the four studies reviewed above, there are three central issues being addressed. First, they all attempted to explain the geographical movement of retail trade among cities and areas, or in other words, they all attempted to explain consumers' spatial behavior, especially the out-of-town shopping phenomenon. Second, population or size of the shopping center, location or distance away from home, and consumers' travel cost are important explanatory variables of consumer's shopping decisions. Third, it is a dynamic process to shape the retail structure. The interaction among consumers, retail outlets, and the cities, determines and defines the trading area.

2.2 Retail Sales and its Determinants

While Reilly's model was further defined in its mathematical form, scholars in this research stream started to apply his work to empirical studies. Applied research in this field started with describing and explaining aggregate retail sales in various sizes of communities and cities in the United States. Different sets of structural determinants of retail sales were examined, and census data were used for most studies. Given the availability of secondary data at the time, early studies focus on examining the impacts of external environment (*i.e.*, consumer characteristics) on retail sales, and showing how different socioeconomic and demographic characteristics affect sales. Later, scholars recognized that the internal environment (*i.e.*, a retailer's marketing efforts) and the competition retailers face could potentially influence retail sales, thus, marketing mix and competitive variables were developed and integrated into the analysis to explain retail sales.

2.2.1 Socioeconomic & Demographic Determinants

Ford (1935) was among the first to examine retail structure with descriptive statistics in the literature. He explained the changes in the retail structure of UK between 1901 and 1931. Ford found that as population and the number of families and households increase in a city, the number of food and necessity shops declines in general. This was partly due to the change in transportation method at the time - consumers were able to travel further to large stores. This phenomenon is later referred as the "Ford Effect", which essentially states that the average store is serving more customers. Although the number of stores decreases, the size of store and merchandise assortment increase, leading stores to serve more people on average. Ford's study provides two major insights to our understanding of retail structure. First, due to the advances in transportation, consumers were willing to travel further to large stores. Second, retailers responded to the change in consumer shopping behavior with fewer but larger stores. Thus, the retail structure started to reshape and rebalance until it reached an equilibrium number and size of stores to serve customers in the area.

Beginning in the 1950s, scholars started to use U.S. Census data to empirically examine retail sales and its determinants. The first set of determinants was consumers' socioeconomic and demographic characteristics. Russell (1957) used the 1950 U.S. Census of Population data to examine the relationship between income and retail sales per capita in communities among 78 small cities in urbanized areas and 168 small cities outside urbanized areas in all regions across the United States. Population in these cities ranged from 25,000 to 49,999. She found that median income per family is positively related to retail sales outside urbanized areas but has no effect in urbanized areas. She inferred that the no effect in urbanized areas is due to population influx (e.g., residents from nearby areas, tourists, vacationists, or business visitors) to the focal area. Her results implied that when areas under consideration are relatively closed systems - residents earn and spend their income within the area, the positive effect of income on sales holds. However, the effect may disappear when residents frequently go out-of-town shopping. Furthermore, she compared results obtained from communities with those from 50 standard metropolitan statistical areas[6] and those obtained from states. She showed that the correlation between median income per family and retail sales per capita is highest in states, second highest in metropolitan areas, and lowest in communities. This further indicates that states are large enough to be self-sufficient economically (*i.e.*, closed systems), and metropolitan areas are to some extent less self-sufficient but more "closed" than communities. This study was among the first in the literature to recognize the existence of "out-of-town shopping" between different metropolitan areas.

Ferber (1958) investigated the factors influencing total retail sales in 51 Illinois cities with 10,000 or more population. Consistent with Russell's findings, income is a key determinant

[6] Standard Metropolitan Statistical Areas (SMSAs) were MSAs in the early years.

of total retail sales. Other than income, he found population, distance to the nearest larger city, the number of stores per capita in the city, and the relative number of stores in the city to the nearest larger city, are determinants of both total retail sales and per capita sales. In short, both Russell and Ferber realized that consumers do out-of-town shopping, and the "pulling power" from surrounding larger cities have substantial impact on sales in the focal area.

Hall, Knapp, and Winsten (1961) used UK and U.S. state data to compare the distribution of goods in the two countries. They found that both average sales per store and the number of stores per capita are influenced by income per capita, the growth rate of population, and population density. Van Tassel (1965) also used state data, but from four census years 1948, 1954, 1958, and 1963 to study the factors that affect retail sales in ten lines of trade. He extended the list of explanatory variables to include not only income, but also total population, race ratio (*i.e.*, white vs. non-white), gender ratio, occupation ratio (*i.e.*, agricultural vs. non-agricultural), educational attainment, age, unemployment rate, automotive ownership, population density, and average number of stores per capita. He found that in general income per capita and nonagricultural employment are two significant determinants of per capita retail sales across various retail industries, and these two variables explain a large proportion of the variance in per capita retail sales.

Bruce (1969) adopted a different approach to examine U.S. retail structure. He did a hierarchical cluster analysis with data of 79 large U.S. cities (*i.e.*, 150,000 or more population) from the 1960 Census. He used six variables - median income, median age, percent of residents in the same house in 1960 as in 1955, population density, per capita wholesale establishments, and per capita employed in manufacturing to group the 79 cities. His study shows that the clusters have significantly different retail structures in terms of the mean per capita sales. This implies that clusters with varying consumer characteristics adjust themselves to allow their retail systems (*i.e.*, marketing systems) adapt to their external environment (*i.e.*, consumers' socioeconomic and demographic characteristics).

Later, scholars realized that city or state is not the proper unit for analyzing aggregate retail sales (Ingene & Lusch 1980; Liu 1970). Cities are too small, and states are too large to be considered integrated and meaningful economic entities. Thus, scholars started to treat standard metropolitan statistical areas (SMSAs) as the unit of analysis. SMSAs are used for several reasons. First, retail stores in the same SMSA are in direct competition with each other. Second, SMSAs are meaningful labor markets. Third, marketing activities (e.g., store promotion) primarily target customers in the same SMSA. In short, SMSAs are highly integrated social and economic units.

Liu (1970) was among the first to examine retail sales in SMSAs. He used a sample of 38 large SMSAs from two census years 1954 and 1963 to investigate the determinants of retail sales. His results indicate that population, local government expenditures, population density, education, the number of stores, and income are important determinants of retail sales. Moreover, several other scholars used SMSA data for their research on retail structure. Schwartzman (1971) used SMSA data to examine total retail sales and sales per manhour in the U.S. In his study, he added two more variables - wage rate from the manufacturing sector and gasoline sales per household to explain total sales. He found that median family income, manufacturing wage, population growth, population density, and gasoline sales per household are significant determinants of total sales. His study further confirms that income and population indeed drive retail sales in SMSAs, and consumers are willing to incur travel costs (*i.e.*, gas, time) to travel further for shopping purposes in an SMSA. Forbes (1972) used 217 SMSA data from the 1963 Census and analyzed the structure of urban retailing. Forbes used population in the SMSAs as the sole explanatory variable of number of stores. He recognized that more specialty stores and fewer general stores per capita were found in larger metropolitan areas. His research provides very important theoretical insights – scale economy is easily attained for general stores (*i.e.*, therefore, fewer stores in larger SMSAs – "Ford Effect"), also, it requires certain threshold level of population and income to support more specialized retail institutions.

Later Bucklin (1972) and Takeuchi and Bucklin (1977) extended the literature on retail structure. Bucklin (1972) examined U.S. retailing in the aggregate. Other than the set of variables discussed earlier in this section, he added retail wage rate and the number of manufacturing establishments per capita (as a proxy for the percentage of families classified as working class) to explain store sales. Takeuchi and Bucklin (1977) extended Bucklin's 1972 study with a comparison of Japan and U.S. retailing. They used the share of total retail sales captured by department stores as a proxy for service in the analysis. Although it was not a good proxy for service, until then, scholars started to realize that marketing effort (*i.e.*, service) might potentially influence the level of retail sales in a geographic area.

In summary, scholars recognized that different socioeconomic and demographic variables influence retail sales at the aggregate level. Figure 2.4, which first appeared in Chen et al. (2020), presents this theoretical model. Table 2.1 provides an overview of some very important socioeconomic and demographic determinants examined in the literature.

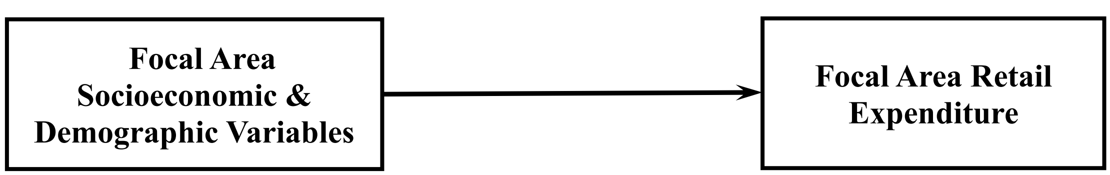

Figure 2.4. Theoretical Model with S&D[7] Variables (Chen et al. 2020)

[7] S&D refers to Socioeconomic & Demographic.

Table 2.1. Socioeconomic & Demographic Determinants of Retail Sales

Determinant	Level of Analysis	Method	Exemplars
Income	City	OLS Regression	Ferber (1958), Russell (1957),
	State	OLS Regression	Hall, Knapp, & Winsten (1961), Van Tassel (1965)
	MSA	OLS Regression	Liu (1970), Schwartzman (1971)
Population	City	OLS Regression	Ferber (1958)
	State	OLS Regression	Van Tassel (1965)
	MSA	OLS Regression	Forbes (1972)
Population Growth Rate	State	OLS Regression	Hall, Knapp, & Winsten (1961), Van Tassel (1965)
	MSA	OLS Regression	Takeuchi & Bucklin (1977)
Population Density	State	OLS Regression	Hall, Knapp, & Winsten (1961)
	MSA	OLS Regression	Liu (1970), Schwartzman (1971), Takeuchi & Bucklin (1977)
Automotive Ownership	State	OLS Regression	Van Tassel (1965)
	MSA	OLS Regression	Takeuchi & Bucklin (1977)
Race	State	OLS Regression	Van Tassel (1965)
Gender	State	OLS Regression	Van Tassel (1965)
Unemployment Rate	State	OLS Regression	Van Tassel (1965)
Education	MSA	OLS Regression	Liu (1970)
Age	MSA	OLS Regression	Liu (1970)

2.2.2 Marketing Mix & Competition

Over the following fifty years, as the world evolves, scholars expanded their thinking and understanding of the nature of retailing. Some realized that marketing mix can increase retail sales, thus, should be included in the analysis. Ingene and colleagues have a series of work on retail structure in the 1980s. Not only did they examine the impact of marketing mix variables on sales, but they also added competition in the analysis.

Ingene and Lusch (1980) examined department store sales in 213 SMSAs from the 1972 Census. They used a set of socioeconomic and demographic variables which include household income, the standard deviation and skewness of income, household size, automotive ownership, population density, plus a set of marketing mix variables: merchandise assortment, the level of service (*i.e*., quantity of service), quality of service, and store atmospherics. They realized that it was not only income, but also the distribution of income (*i.e.*, standard deviation and skewness of income) which impacts retail sales. Furthermore, their study was among the first to empirically test the effects of different marketing mix variables on retail sales in the literature. Given the availability of secondary data, these abovementioned marketing mix variables were measured indirectly with proxies. For instance, they used store square footage per capita as a proxy for merchandise assortment; the number of employees per 1000 square feet as a measure for the level of service; wage rate as a proxy for quality of service; and population growth rate as a proxy for store atmospherics. Their results showed that marketing mix significantly increases department store sales. Later, Ingene and Lusch (1981) used data of 210 SMSAs in another line of trade – grocery stores from the 1972 Census and constructed a theoretical framework to analyze retail structure. They examined the effects of a set of socioeconomic and demographic variables, plus a set of marketing mix variables on grocery sales per store and the number of grocery stores per household. Variables include household income, the standard deviation and skewness of income, automotive ownership, traffic congestion, household size, growth rate of population, merchandise assortment, quantity of service, quality of service, and the number of mom-and-pop grocery stores per household. This study confirmed that consumers' socioeconomic and demographic characteristics plus retailers' marketing efforts indeed influence retail sales.

Ingene and Yu (1981) reexamined the determinants of retail sales with data of 229 SMSAs from the 1972 Census across nine lines of trade. Their set of determinants include income, automotive ownership, unemployment rate, household size, percent of population living in an urbanized area, population density, and total population of the area. This study further indicated that SMSA is a proper unit of analysis for retail sales, as the set of determinants are generalizable to nine lines of trade.

Ingene (1983) added competition into the analysis of retail sales. He used the 1977 Census data on SMSAs and examined the intertype competition of food retailing in restaurants versus grocery stores. The set of determinants for this study included income, household size, automotive ownership, percent of young men, percent of whites, concentration ratios, assortment, quality of service, quantity of service, and store atmospherics. He showed that those socioeconomic, demographic, and marketing mix factors influence food expenditures at both restaurants and grocery stores. Furthermore, this study demonstrated that marketing efforts enable one line of retail trade to take sales away from an intertype competitor. Ingene (1984) also used the 1977 Census data on SMSAs but extended the analysis to eight merchandise lines.

In summary, scholars recognized that not only marketing mix variables but also competition influences the level of retail sales. Figure 2.5 presents this theoretical model.

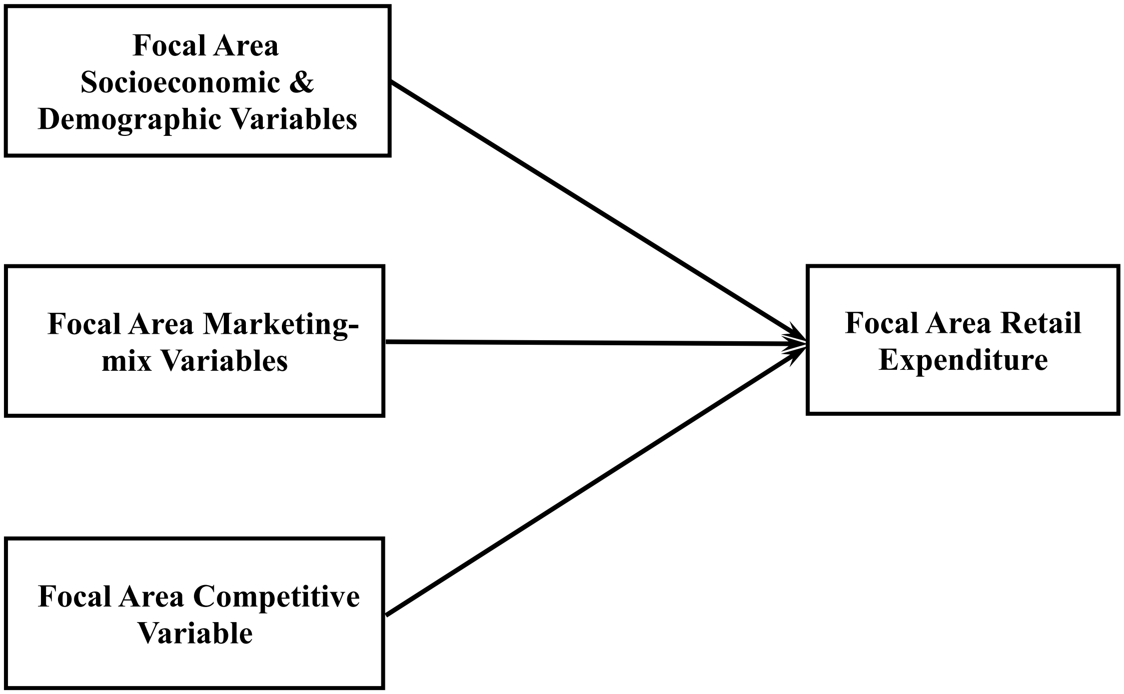

Figure 2.5. Theoretical Model with S&D, MM[8], & Competitive Variables (Ingene 1983)

Various studies had shown that marketing mix increases sales at different lines of trade. However, these studies treated marketing mix variables as exogenous and external to retailers' control. Ingene and Brown (1987) did a study on the structure of gasoline retailing and showed that marketing mix variables are determined and influenced by consumers' socioeconomic and demographic characteristics. This echoes the idea of consumer segmentation and targeting. They also demonstrated that socioeconomic, demographic, and marketing mix variables altogether impact the level of retail sales at gasoline stations. This was the first study to treat marketing mix variables as a set of mediators between socioeconomic & demographic variables and retail sales. This implies that marketing mix variables are managerially optimized in order to adapt to the socioeconomic and demographic characteristics from consumers living and working in the

[8] MM refers to Marketing mix.

geographical area. Miller et al. (1999) extended the work to examine the effects of different types of competition (*i.e.*, intratype, intertype, and intercategory competition) on the retail structure of different retailers. They showed that socioeconomic and demographic variables affect marketing mix variables (*i.e.*, quality and quantity of service level) and these marketing mix variables have significant impacts on the retail structure. Moreover, competition plays an important role in forming the retail structure.

In summary, scholars had realized that marketing mix is managerially determined and optimized to respond to the varying consumer characteristics. Figure 2.6 presents this theoretical model.

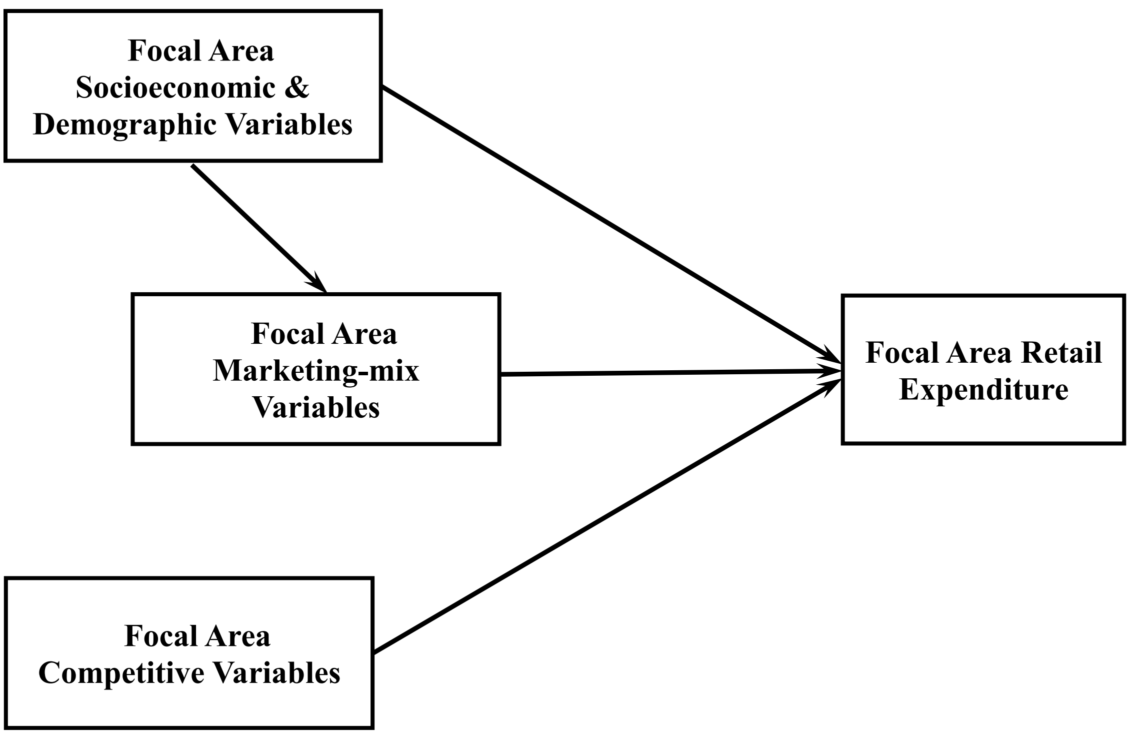

Figure 2.6. Theoretical Model 1 with S&D, Endogenous MM, & Competitive Variables (Ingene & Brown 1987)

A recent study by Chen et al. (2020) utilized the above framework and examined the retail structure of Japanese food and beverage retailers. The authors recognized that not only socioeconomic and demographic characteristics affect managerial decisions on the marketing mix, but also does competition. This implies that decisions on the marketing mix are driven by consumer characteristics as well as competition. It is wise for retailers to adapt and respond to the external environment with appropriate marketing effort. Figures 2.7 presents this theoretical model.

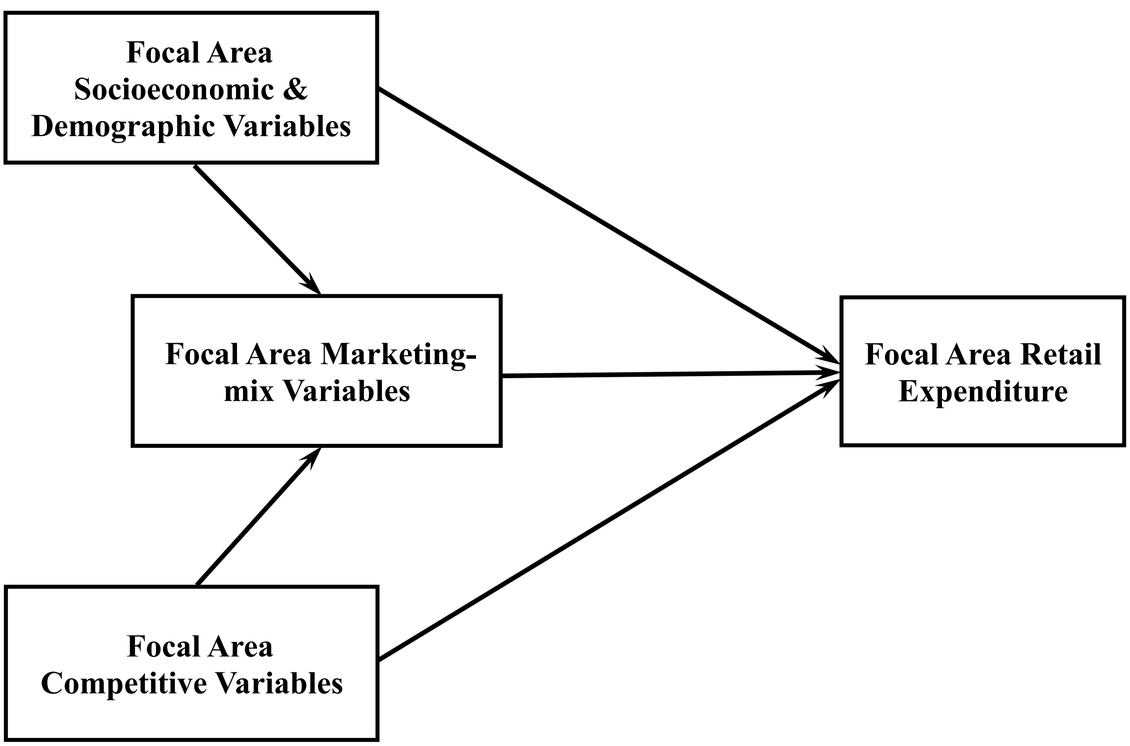

Figure 2.7. Theoretical Model 2 with S&D, Endogenous MM, & Competitive Variables (Chen et al. 2020)

Among all the research studies reviewed in this chapter thus far, the impacts of marketing mix variables on retail sales are most salient, compared to the set of socioeconomic, demographic, and competitive variables. This is partly due to the mediation effects of marketing mix variables.

In summary, scholars recognized that different marketing mix variables influence retail sales and they do so by responding to the socioeconomic and demographic characteristics as well as competition in the geographical area. Table 2.2 provides a summary of marketing mix variables and competitive variables examined in the literature.

Table 2.2. Marketing mix & Competitive Variables

Construct	Proxy Variable	Method	Exemplars
Merchandise Assortment	Floor Space per Capita	Ridge Regression	Ingene & Lusch (1980, 1981)
Quality of Service	Wage Rate	Ridge Regression	Ingene & Lusch (1980, 1981)
Quality of Service	Wage Rate	Factor Analysis & OLS Regression	Ingene (1983, 1984), Ingene & Brown (1987)
Quantity of Service	Number of Employees per Square Feet	Ridge Regression	Ingene & Lusch (1980, 1981)
Quantity of Service	Number of Employees per Square Feet	Factor Analysis & OLS Regression	Ingene (1983, 1984)
Quantity of Service	Number of Employees per Square Meter	OLS Regression	Chen et al. (2020)
Store Atmospherics	Growth Rate of Population	Ridge Regression	Ingene & Lusch (1980, 1981)
Store Atmospherics	Growth Rate of Population	Factor Analysis & OLS Regression	Ingene (1983, 1984)
Intratype Competition	Number of Intratype Competitors per Household	OLS Regression	Chen et al. (2020), Miller et al. (1999
Intertype Competition	Intertype Competitor's Market Share	OLS Regression	Ingene (1983)
Intertype Competition	Number of Intertype Competitors per Household	OLS Regression	Chen et al. (2020), Miller et al. (1999)
Intercategory Competition	Number of Intercategory Competitors per Household	OLS Regression	Miller et al. (1999)

2.3 Consumer Out-Shopping[9]

The phenomenon of consumer out-of-town shopping has been addressed in the retail literature, as discussed earlier in this chapter. However, it has never been examined empirically, due to the following reasons. First, due to the advances in technology and transportation, private automobile ownership has facilitated consumers' geographic mobility and out-shopping behavior in the past twenty decades. Therefore, this consumer out-shopping phenomenon is more notable in recent years. Second, only a handful of scholars in this field understand and conceptualize consumers' out-shopping behavior as a determining factor of retail sales – retailers located in consumers' focal area lose sales to retailers in nearby large areas. Third, it is a challenging task to operationalize Reilly's gravity model with data from secondary sources. Fourth, retail data on micropolitan areas was only available since 2012, and it was impossible for scholars to examine consumer out-shopping with both metropolitan and micropolitan data prior to 2012.

In the first section of this chapter, I have discussed Reilly's *Law of Retail Gravitation* (Reilly 1931) and it was the earliest literature addressing this phenomenon – sales are drawn from nearby cities or communities. This implies that consumers go beyond the boundaries of their own town to nearby cities for goods and services. Although those pioneer scholars such as Reilly, Converse, and Huff addressed this shopping phenomenon mathematically, empirical researchers in this field have never attempted to examine it. As discussed earlier, retail systems were treated as closed systems in previous studies, and most scholars suggested that retail sales are only influenced by consumers who live in the area, except a recent study by Chen et al. (2020). In their study, the authors examined the effect of daytime population (*i.e.*, who come to the city during daytime for work or education but reside in another city) on food and beverage

[9] "Consumer out-shopping" and "consumer out-of-town shopping" are used interchangeably.

sales. They found that the influx of population significantly increases sales at both food and beverage retailers and restaurants. The results are not surprising, given that although people go out-of-town for purposes other than shopping, they will shop for goods while they are there.

The findings from the retail literature (*i.e.*, at the aggregate level) are consistent with those from the consumer behavior literature (*i.e.*, at the micro level). In a quick review of the literature, several key factors emerge: consumers tend to out-shop for wider selections of merchandise in large cities or due to the inadequacy of assortment at local stores; consumers out-shop for higher quality of service at stores in large cities or due to their dissatisfaction of price, quality, or service at local stores; consumers who out-shop tend to have better education and higher income; out-shopping behavior is not product-specific, but rather generalized to many types of shopping goods; certain products such as personal or convenience goods are mostly purchased from local stores; and social interaction with friends facilitates out-shopping (e.g., Herrmann & Beik 1968; Reidenbach et al. 1984; Reynolds & Darden 1972; Reynolds & Martin 1974; Thompson 1971).

In this dissertation, I build upon the retail literature and extend it to examine this consumer out-shopping phenomenon. Population has been examined as a determinant of retail sales in the literature. It was also used by Reilly (1931) as a proxy for the "drawing power". Therefore, I use population from nearby areas in certain radiuses as proxy for their influence on retail sales at the focal area. Figure 2.8 presents this theoretical model.

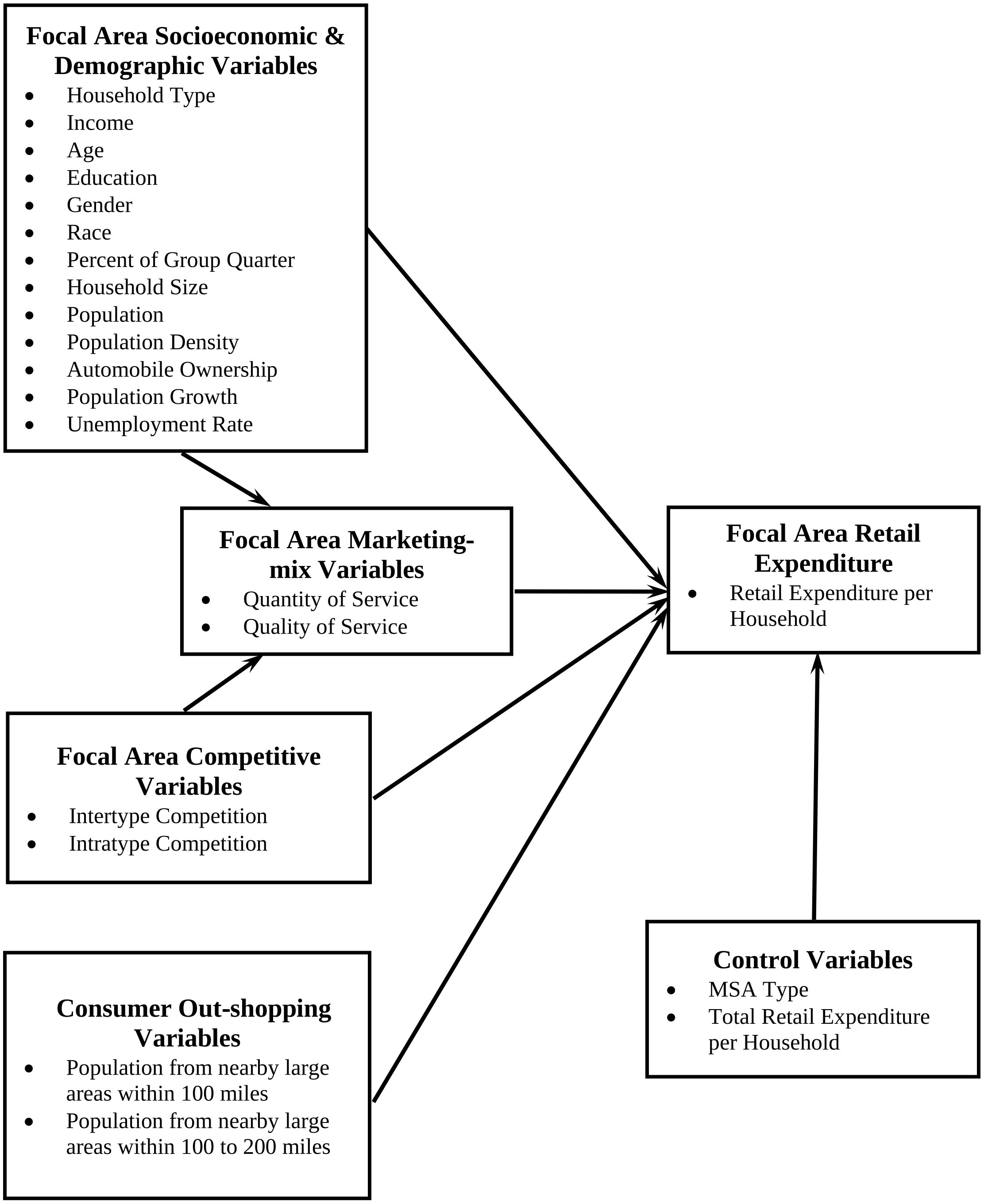

Figure 2.8. Theoretical Model Developed and Examined in This Dissertation (Expanded)[10]

[10] Details regarding these variables are provided in Chapter 5. Dependent variable "Retail Expenditure per Household" refers retail expenditure per household in one specific retail subsector or industry. Control variable "Total Retail Expenditure per Household" refers to the total expenditure per household in all twelve subsectors of the retail sector (NAICS 44-45).

Chapter 3 A Theory of Retail Market Structure

In this chapter I develop a theory of retail market structure to synthesize current understanding and thinking of retailing from both the consumer's and firm's perspectives. The goal is to provide a theoretical foundation to integrate and extend existing research. Consumers and firms are basic economic entities because they perform activities of production and consumption in an economy. Drawing on basic economic theories and marketing literature, I set forth this theory of retail market structure.

This chapter is arranged as follows. First, I discuss the theory from a consumer's perspective (*i.e.*, demand), and show how consumers make decisions on choosing retailers regarding what to buy and where to buy. Second, I discuss the theory from a retailer's perspective (*i.e*., supply), and show how retailers make managerial decisions on store locations, product assortments, service, etc. Lastly, I integrate the two perspectives and show how they interact and impact each other.

3.1 A Consumer's Perspective

Consumers are economic entities, whom typically try to maximize their welfare, given their limited economic and temporal resources (Becker 1965). To do so, consumers evaluate the benefits of shopping against the costs of shopping incurred in the process.

3.1.1 Consumer Shopping Benefits

Consumers shop for many reasons. As Tauber (1972) indicated in his study on why people shop, there are four broad categories of shopping motives. First, consumers shop for the need of products or services. By doing so, consumers gain utility from acquiring goods or experiencing services. Second, consumers shop for psychological and social benefits which are associated with and thus obtained via shopping. Psychological benefits include self-gratification, role-playing, psychic diversion, and sensory stimulation, while social benefits consist of social experiences outside the home, social interaction with others having similar interests, peer group attraction, status and authority, and the pleasure of bargaining (Tauber 1972). Third, consumers shop for informational benefits. Consumers can learn about new trends, browse for product information, and compare prices in a shopping trip. Fourth, consumers shop for physical benefits. Not only is shopping a social behavior, but it is also a physical activity. Shopping provides consumers with opportunity to exercise, but at a leisurely pace.

One should note that there is a significant difference between shopping and buying. Shopping, in a broad sense, involves consumers going on a shopping trip to retail stores, but they are not committed to make any purchases on products or services, except for those expenses that would not have been incurred without the shopping trip, such as gasoline, food, beverages, etc. Buying, on the other hand, involves the ultimate outcome of shopping - consumers' dollars spent on products or services. This dissertation is only concerned with consumers' buying behavior, because the primary interest is in consumers' (*i.e.*, households') expenditure patterns at different retail lines. Therefore, I do not consider "window shopping" and its associated benefits and costs in this dissertation, because, first, they do not trigger the actual buying behavior nor the expenditures on products or services; second, this "window shopping" cannot be measured with secondary data. In summary, I focus on the first category of benefits - consumers' utility derived from acquiring goods or experiencing services.

3.1.2 Consumer Shopping Costs

There are four types of cost associated with consumer shopping and buying. Costs consist of economic cost (*i.e.*, dollars), temporal cost (*i.e.*, time), physical cost, and psychic cost. First, consumers incur economic cost in their shopping and buying. This cost includes dollars spent on products and services, and expenses incurred with the shopping trip itself such as gasoline, food, beverages, etc. Second, shopping and buying requires consumers to voluntarily and involuntarily spend their time (Ingene & Ghosh 1990). Therefore, this temporal cost includes the time needed to and from the store, plus the time needed inside the store. Third, shopping is a physical activity, and therefore requires physical energy, or physical cost. This cost includes traveling (*i.e*., driving, riding a bus, walking) to and from a store, walking inside a store, picking up products from store shelves, and all other physical activities involved in a shopping trip. Fourth, as a social behavior, shopping requires consumers to inevitably socialize with others in a shopping trip. This may involve talking with salespeople, interacting with shopping companion, and all other psychic energy needed during a shopping trip.

Due to data availability, I focus on the first two types of cost – economic cost and temporal cost, due to data availability. Economic cost – how much a consumer is willing to pay for a product or service, is largely determined by his or her disposable income. Temporal cost – voluntary time and involuntary time, is largely determined by the store size (*i.e.*, time incurred inside the store), the mode of transportation, and the distance away from a store. For simplicity, I assume consumers and households to be rational decision makers, in that they would try to minimize their shopping costs as much as possible, while maximizing their utility derived from a shopping trip.

3.1.3 Consumer Buying Decisions – What to Buy and Where to Buy

As stated earlier in this chapter, consumers try to maximize their utility, given their economic and temporal constraints. The tradeoff between shopping benefits and shopping costs can be demonstrated as follows. Consumers would:

$$max \ \{ \sum_1^n U\ (p_i) - \sum_1^n U\ [e(\ p_i\ )] - \sum_1^n U\ [t(\ p_i\ )]\}^{11},\ i = 1, 2, \ldots, n \qquad (3.1)$$

where U refers to utility; p_i refers to product i; $e(\ p_i\ )$ refers to the economic cost associated with product i; $t(\ p_i\ )$ refers to the temporal cost associated with product i; $\sum_1^n U\ (p_i)$ refers to the total utility derived from buying n products; $\sum_1^n\ U\ [e(\ p_i\ )]$ refers to the total disutility derived from the economic costs incurred in buying n products; and $\sum_1^n\ U\ [t(\ p_i\ )]$ refers to the total disutility derived from the temporal costs incurred in buying n products.

One should note that this expression 3.1 is on a per product basis. Each utility, economic cost, and temporal cost is for per product. Therefore, if a consumer is on a shopping trip buying multiple products at once, it is very likely he or she will save some economic and temporal costs by doing a comprehensive multi-purpose one-stop shopping.

In order to maximize the value of (3.1), consumers would evaluate the utility versus the disutility derived from buying products or services. Utility is determined by the level of pleasure or satisfaction that a consumer obtains from the acquisition and consumption of products and services. This is a question of *what* to buy to satisfy the needs. What to buy and how much to spend depend on consumer characteristics which include income, household size, and possibly their transportation mode. Furthermore, it is a matter of personal choice and individuals may have different preferences, given their socioeconomic status. Therefore, *what* to buy is considered as internally determined by the consumer.

For the disutility derived from economic cost which primarily consists of dollars spent on products or services, I take price as given and assume there is no significant pricing difference at a given retail trade line across different metropolitan and micropolitan areas, due to data availability. However, one should note that a product may be sold at different retail lines and

[11] I use this linear expression of utility maximization for simplicity and demonstration purposes. The utility function may not be linear in reality.

prices of the product may be different. This will impact consumers' choice of *where* to buy. For example, grocery prices may be lower at a supercenter (*i.e.*, General Merchandise Store), compared with prices at a grocery store (*i.e.*, Food and Beverage Store).

For the disutility derived from temporal cost, three factors may mitigate this cost and therefore lower the disutility. First, the number of stores and the location of a store may impact the time needed for consumers to travel to and from a store. Intuitively, the closer a store locates within a geographic area to a consumer, the lower the time cost is incurred by a consumer traveling to and from a store. However, consumers' time cost depends on the retail structure consumers face in a geographic area. The more retail stores in an area, the shorter the distance an average consumer must travel; on the other hand, the fewer the number of retail stores, the greater the distance and therefore the greater the time cost is incurred in traveling. Second, the level (*i.e.*, quantity) and quality of in-store service impact the time needed inside a store. The great the in-store service provided by store employees, the lower the time cost is incurred by a consumer needing to spend inside the store. Third, the mode of transportation is likely to influence the time needed to travel to and from a store. In general, in most areas the greater the number of automobiles owned by a consumer or household, the less the time cost incurred in traveling to and from a store. However, in large metropolitan areas where public transportation is widely available, it may be more time efficient to take public transportation (*i.e.*, subway) when there exists high degree of traffic congestion.

To summarize, although I do not consider retailers' influence on consumers' choice of *what* to buy, it is certain that retailers are attempting to influence consumers' decision on *where* to buy, as they can provide tailored marketing efforts to lower consumers' disutility of temporal cost. In this dissertation, I discuss the use of service (*i.e.*, both quantity and quality) and its potential impact on consumer expenditures.

3.2 A Retailer's Perspective

In this dissertation, I assume retailers are profit-driven, and therefore they will try to maximize their profit in their trade area[12]. Profit is calculated as the difference between the total sales revenue and the total cost incurred with operating a store. If expressed in another easy to follow formula, profit equals the net revenue minus fixed costs. Profit-driven retailers will only carry products that are profitable. Thus, assuming there are n products sold at a store, the i^{th} product's profit equals its per-unit margin (*i.e.*, price minus the per-unit cost of goods sold) times its quantity sold minus its related fixed cost such as the cost of the shelf space. To sum up the total profit at a store level, one should deduct the store level fixed costs (e.g., rent, utilities, salaries, etc.) from the aggregated profit (Chen et al. 2020). The mathematical formula is shown as follows.

$$\Pi = \sum_{1}^{n} (m_i Q_i - f_i) - F, \; i = 1, 2, \ldots, n \tag{3.2}$$

where Π refers to profit; m_i refers to the per-unit margin of product i; Q_i refers to the quantity sold for product i; f_i refers to product i's fixed cost; and F refers to the store-level fixed costs. One thing worth mentioning is that when the i^{th} product's sales decline to which it will not cover its fixed cost (f_i), then it is likely that the i^{th} product will be dropped from the product portfolio, because it will no longer contribute to profit.

To understand each component that makes up equation 3.2, I decompose Q_i into the following parts. Firstly, the i^{th} product's quantity sold at a retailer store can be expressed as the total quantity sold at all retail stores in trade area times the focal retailer's market share of

[12] A retailer's trade area may be smaller than the area of the city where the retailer is located. The trade area is partly determined by how far a consumer is willing to travel (Reilly 1931).

product *i* in the area. Secondly, the total quantity sold at all retail stores can be further expressed as the trade area's population times the average quantity a representative consumer would purchase. Thirdly, a trade area's population equals the population density times the trade area. Assuming the trade area is circular for simplicity, Q_i is expressed as follows.

$$Q_i = s_i \, q_i \, \Phi \, \pi \, d^2, \; i = 1, 2, \ldots, n \tag{3.3}$$

where s_i is product *i*'s market share acquired by the focal retailer in the trade area; q_i is the average quantity of product *i* a consumer would purchase in the trade area; Φ refers to the population density in the area; π is a constant number which approximately equals to 3.1415926; and *d* refers to the radius of the circular area, which also indicates the distance an average consumer is willing to travel and shop at the focal store.

Plugging in Q_i from equation 3.3 into equation 3.2, I obtain the following:

$$\Pi = \sum_1^n \left(m_i \, s_i \, q_i \, \Phi \, \pi \, d^2 - f_i\right) - F, \; i = 1, 2, \ldots, n \tag{3.4}$$

Thus, a retailer would try to

$$max \; \Pi = \sum_1^n \left(m_i \, s_i \, q_i \, \Phi \, \pi \, d^2 - f_i\right) - F, \; i = 1, 2, \ldots, n$$

In the following of this section, I discuss each component of the profit.

Per-unit margin m_i. As stated earlier, per-unit margin is determined by price and per-unit cost of goods. Pricing and cost strategies are internally determined by the firm, and by competition. This is beyond the scope of this dissertation, and I take per-unit margin as taken.

Market share s_i. Market share is determined by demand of the product at the focal store. Other than consumers' socioeconomic characteristics such as income, household size, and transportation mode, a retailer's marketing efforts and its competitors' offerings will also influence consumers' buying decisions. In other words, a retailer is capable of increasing market share by properly utilizing elements in the marketing mix. For example, the greater the in-store service, the more likely a consumer would be able to find a product he or she wants and thus make a purchase at the focal store. In the meantime, customer service would lower a consumer's temporal cost of locating a product and checking out, thus, consumers' disutility of temporal cost is reduced.

Average quantity sold per capita q_i. How many an average consumer would buy from a retailer depends on their socioeconomic characteristics, as well as the retailer's marketing effort. A wise retailer is not only capable of increasing its market share, but also capable of increasing the quantity a consumer is willing to buy, through various marketing strategies, such as providing excellent in-store and customer service. Also, if a product comes only with a bundle, or a big pack, at an affordable price, then consumers may be willing to purchase, given a lower per-unit price. Examples can be seen at warehouse clubs and supercenters. Furthermore, large stores are more attractive to value-driven consumers, due to their large selection of products (*i.e.*, product assortment) and the convenience of one-stop shopping. By going to large stores, consumers can stockpile and save their temporal cost. Unfortunately, I do not have information on product assortment, but this will be discussed later in this dissertation.

Population density Φ. I take population density as given in this dissertation. However, one should note that population density impacts a retailer's decision on store location and store size. For example, if a retailer looks to open a store in a low density and less populated area, then from equation 3.4, one would know that the store needs to be large and large enough to serve as many consumers in the area as possible, because only a large store with greater product assortment will be able to draw consumers located further away from the store.

Travel distance d. How far a consumer is willing to travel depends on their socioeconomic and demographic characteristics, the store location, the transportation mode, competitors' offerings, and the focal retailer's offerings (*i.e.*, marketing mix). Marketing mix, if used properly by managers, can increase consumers' travel distance.

Competition. In this dissertation, I consider both intertype and intratype competition. Intertype competition occurs between retailers selling same or substitutable products but in different retail industries located in the same area. Intratype competition refers to competition between retailers selling same or substitutable products within the same retail subsector. In this dissertation, specifically, intratype competition occurs between retailers in the subsector NAICS 452 General Merchandise Stores. For instance, Supermarkets and Other Grocery (except convenience) Stores is in intertype competition with Warehouse Clubs and Supercenters, whereas Warehouse Clubs and Supercenter is in intratype competition with Discount Department Stores.

Fixed costs f_i and F. The fixed costs are partially determined by the location and thus the rent of the store, utilities, overhead, and the number of employees working at the store. Intuitively, if a retail store hires more employees, then it will incur a higher fixed cost. However, on the other hand, the store will be able to provide more services to customers, and potentially the store will generate a higher sales revenue. The same logic applies to paying to higher wages to employees. If a store pays higher than average wage to employees, then it will incur a higher fixed cost, but the store will be able to provide greater services and thus generate more sales.

A retailer's trade area is illustrated in Figure 3.1. Assuming a retailer faces a circular market for simplicity, then the retailer is located at the center point of the circle, and the maximum distance a consumer is willing to travel and shop at this retailer is the radius d. In other words, the furthest consumers this retailer can reach are located on the edge of this circle.

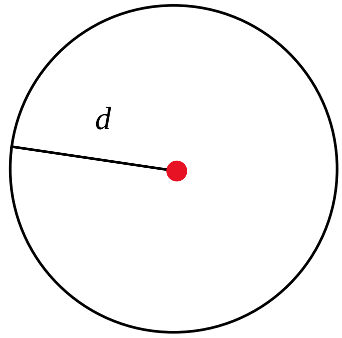

Figure 3.1. A Retailer's Trade Area

3.3 A Theory of Retail Market Structure

To understand how consumers and retailers interact with each other, I integrate the two perspectives from consumer's and retailer's, respectively, then I discuss how retailers may influence consumers' decision on *where* to buy, and lastly, I consider the possibility of consumer out-of-town shopping.

As Bucklin (1967) stated, shopping is a function of store location, product assortment, and the store image (*i.e.*, a mix of all the store can offer to consumers). Consumer evaluate the benefits and costs of shopping, and they do so by maximizing { $\sum_1^n U\ (p_i) - \sum_1^n U\ [e(\ p_i\ )] - \sum_1^n U\ [t(\ p_i\ )]$} (expression 3.1). From a retailer's perspective, retailers are profit-driven, and therefore they maximize their profit in a given geographic area. Profit is calculated as Π = $\sum_1^n\ (m_i\ s_i\ q_i\ \Phi\ \pi\ d^2$ - f_i) – F (equation 3.4).

Recall Figure 2.8 in Chapter 2, one should note that consumers' socioeconomic and demographic characteristics, and competition from rivals are external to the focal retailer, the focal retailer cannot control who shop at its store, nor what its rivals offer to consumers. However, the focal retailer can adapt and respond to its external environment by using different marketing mix efforts. For instance, retailers can provide greater in-store service to lower consumers' temporal cost, and thus leading to more sales. However, consumers who cannot afford this extra service may choose not to shop at the focal retailer. Thus, marketing mix efforts need to tailor to consumers with different socioeconomic and demographic characteristics, and this is what we call "market segmentation" in marketing.

Figure 3.1 shows a retailer's local trade area. When consumers do out-of-town shopping, the trade area is expanded to reach and include consumers who live outside of the focal area. Figure 3.2 illustrates a retailer's trade area with out-of-town consumers. The original trade area has a radius of *d*, but with consumers from other areas, the trade area has been expanded and now has a new radius of *d'*. The shaded area indicates the extra area and consumers that the focal retailer covers.

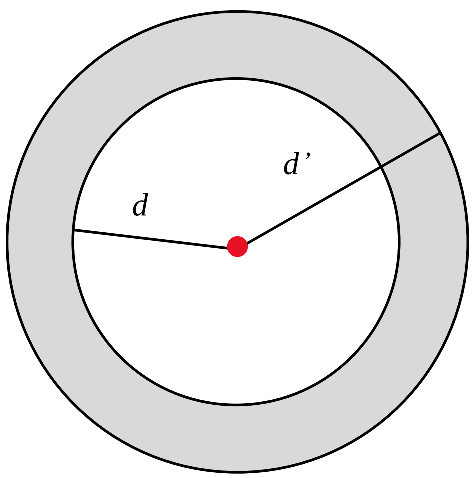

Figure 3.2. A Retailer's Trade Area Expanded with Out-of-town Consumers

To summarize, a wise retailer would respond to its external environment by adopting differentiated marketing mix efforts to serve its customers and to compete with rivals. Furthermore, retailers should pay special attention to attract and retain out-of-town shoppers, because they incur higher economic and temporal costs to shop at their stores. Moreover, retailers should note that their major sources of profitability are still the consumers living in the focal area. This expanded ring area will have a much lower market share compared to its original market. However, exceptions may apply when multiple metropolitan areas are in close proximity or clusters with each other, such as those in New York and New Jersey.

Chapter 4 U.S. Retail Trade Sector

This chapter provides an overview of the U.S. retail trade sector and details regarding the seven subsectors that I examine in this dissertation. As discussed earlier, I use the 2012 Economic Census data of U.S. metropolitan and micropolitan areas for this dissertation.

4.1 Overview of U.S. Retail Trade Sector (NAICS 44-45)

4.1.1 NAICS 44-45 Retail Trade: Definition

The retail trade sector is classified using the North American Industry Classification System (NAICS). The NAICS is the standard used by governmental agencies in the United States, Canada, and Mexico to classify business establishments for collecting, analyzing, and publishing statistical data related to their economy.

The retail trade sector is defined as the industrial sector that "comprises establishments engaged in retailing merchandise, generally without transformation, and rendering services incidental to the sale of merchandise" (NAICS). This sector consists of twelve subsectors, including:

- NAICS 441 – Motor Vehicle and Parts Dealers
- NAICS 442 – Furniture and Home Furnishing Stores
- NAICS 443 – Electronics and Appliance Stores
- NAICS 444 – Building Material and Garden Equipment and Supplies Dealers
- NAICS 445 – Food and Beverage Stores
- NAICS 446 – Health and Personal Care Stores
- NAICS 447 – Gasoline Stations
- NAICS 448 – Clothing and Clothing Accessories Stores
- NAICS 451 – Sporting Goods, Hobby, Musical Instrument, and Book Stores

- NAICS 452 – General Merchandise Stores
- NAICS 453 – Miscellaneous Store Retailers
- NAICS 454 – Nonstore Retailers

There is one important note that the retail trade sector sells untransformed merchandise from automobiles to food for personal or household consumption, at both brick-and-mortar stores and non-store retailers (*i.e.*, electronic shopping, mail-orders, & vending machines). There is another sector – NAICS 72 Accommodation and Food Services which involves the sales of transformed products (*i.e.*, prepared food). In this dissertation, I focus on seven of the abovementioned brick-and-mortar retail subsectors which break down to nine lines of retail trade, which will be discussed later in this chapter.

4.1.2 NAICS 44-45 Retail Trade: 2012 Economic Census

The United States has a well-established distribution network for all types of retail businesses. In order to keep track of these businesses and understand their impact to the U.S. economy, the U.S. Census Bureau collects extensive information and data about these retail businesses every five years. This official statistical release, known as the Economic Census, serves as the basis for the measurement of U.S. businesses and the economy. The data produced from the Economic Census covers businesses across all geographic areas in the United States, including large, medium, and small size businesses. The 2012 Economic Census contains data across 18 industrial sectors classified using NAICS (see Appendix A for details). In this dissertation, I use data of the retail trade sector (NAICS 44-45).

The key statistics produced from the Economic Census include Total Number of Retail Establishments, Total Sales Revenue, Annual Payroll, First-quarter Payroll, and Total Number of Paid Employees for Pay Period Including March 12 (*i.e.*, number of employees in the first quarter). In 2012, the retail trade sector generated a total sales revenue of 4.22 trillion U.S. dollars (U.S. Census Bureau). Among the twelve retail subsectors, Motor Vehicle and Parts Dealers (NAICS 441) contributes the highest share of revenue – 0.87 trillion U.S. dollars, followed by General Merchandise Stores (NAICS 452) – 0.64 trillion U.S. dollars and Food and Beverage Stores (NAICS 445) – 0.62 trillion U.S. dollars. Table 4.1 provides the key statistics and brief descriptions of all twelve subsectors. To save space, first-quarter payroll is not included in Table 4.1.

Table 4.1. U.S. Retail Trade Sector – 2012 Economic Census

NAICS Code (3-digit level)	Sector Name	Number of Retail Stores	Number of Employees in the 1st Quarter	Annual Payroll ($M)	Total Sales ($M)	Share of Total Retail Sales	Included sub-industries (NAICS 4-digit level)	Data Used
44-45	**Retail Trade**	**1,062,083**	**14,703,529**	**369,001**	**4,219,822**	**100.00%**		
441	Motor Vehicle & Parts	116,482	1,709,998	70,165	868,805	20.59%	4411 Automobile Dealers; 4412 Other Motor Vehicle Dealers; 4413 Automotive Parts, Accessories, & Tire Stores	Yes
442	Furniture & Home Furnishings	51,635	424,651	11,485	89,058	2.11%	4421 Furniture Stores; 4422 Home Furnishing Stores	Yes
443	Electronics & Appliance	48,826	431,616	10,202	102,597	2.43%	4431 Electronics & Appliance Stores	Yes
444	Building Material & Garden Supplies	78,282	1,144,716	33,542	279,140	6.61%	4441 Building Materials & Supplies Dealers; 4442 Lawn & Garden Equipment and Supplies Stores	No
445	Food & Beverage	147,579	2,864,650	60,539	620,024	14.69%	4451 Grocery Stores; 4452 Specialty Food Stores; 4453 Beer, Wine, & Liquor Stores	Yes
446	Health & Personal Care	90,959	1,002,165	30,995	270,461	6.41%	4461 Health & Personal Care Stores	No
447	Gasoline	114,474	856,538	15,572	554,256	13.13%	4471 Gasoline Stations	Yes

Table 4.1. U.S. Retail Trade Sector – 2012 Economic Census (Cont.)

NAICS Code (3-digit level)	Sector Name	Number of Retail Stores	Number of Employees in the 1st Quarter	Annual Payroll ($M)	Total Sales ($M)	Share of Total Retail Sales	Included sub-industries (NAICS 4-digit level)	Data Used
448	Clothing & Accessories	147,709	1,664,114	28,000	233,812	5.54%	4481 Clothing Stores; 4482 Shoes Stores; 4483 Jewelry, Luggage, & Leather Goods Stores	Yes
451	Sporting Goods, Hobby, Music, & Book	46,353	518,826	9,136	77,649	1.84%	4511 Sporting Goods, Hobby, & Musical Instrument Stores; 4512 Book Stores & News Dealers	No
452	General Merchandise	49,147	2,772,612	58,048	640,627	15.18%	4521 Department Stores; 4529 Other General Merchandise Stores	Yes
453	Miscellaneous	107,040	713,315	14,176	97,588	2.31%	4531 Florists; 4532 Office Supplies, Stationary, & Gift Stores; 4533 Used Merchandise Stores; 4539 Other Miscellaneous Store Retailers	No
454	Nonstore Retailers	63,597	600,328	27,140	385,804	9.14%	4541 Electronic Shopping & Mail-order Houses; 4542 Vending Machine Operators; 4543 Direct Selling Establishments	No

In this dissertation, I examine seven out of the twelve retail subsectors, given their importance in the U.S. economy and relevance of the research subject. They are Motor Vehicle and Parts Dealers (NAICS 441), Furniture and Home Furnishing Stores (NAICS 442), Electronics and Appliance Stores (NAICS 443), Food and Beverage Stores (NAICS 445), Gasoline Stations (NAICS 447), Clothing and Clothing Accessories Stores (NAICS 448), and General Merchandise Stores (NAICS 452). These seven subsectors generate 73.68% of the total retail sales revenue, employ nearly 73%[13] of all retail workers, and comprise 63.63% of all retail stores.

The five retail subsectors which are not included in this dissertation are Building Material & Garden Supplies (NAICS 444), Health & Personal Care (NAICS 446), Sporting Goods, Hobby, Music, & Book (NAICS 451), Miscellaneous (NAICS 453), and Nonstore Retailers (NAICS 454). Nonstore retailers (*i.e.*, electronic shopping, mail-order houses, vending machines) are excluded from this dissertation because I only focus on brick-and-mortar retailers. The other four subsectors are excluded from this dissertation for two reasons. First, their shares of total retail sales are relatively small among the twelve subsectors. Second, studies have shown that consumers do out-of-town shopping for more style related products such as clothing, furniture, etc., and they buy certain products which possess a personal or convenience nature almost exclusively from local stores (e.g., Reynolds & Martin 1974). Garden supplies, health care products, music, book, and miscellaneous (*i.e.*, flowers, office supplies, gifts) are examples of these products that consumers rarely buy from out-of-town. Therefore, given their irrelevance to the topic of this dissertation, these five subsectors are excluded from the analysis.

[13] This figure is calculated with numbers of first-quarter employees reported in the 2012 Economic Census.

In the following seven subsections of this chapter, I discuss the seven retail subsectors included in this dissertation, respectively. In the last subsection, I provide a general observation of the seven retail subsectors.

4.2 NAICS 441 Motor Vehicle and Parts Dealers

4.2.1 NAICS 441 Motor Vehicle and Parts Dealers: Definition

The Motor Vehicle and Parts Dealers subsector is the largest component of the retail trade sector. NAICS defines 441 as follows.

> "Industries in the motor vehicle and parts dealers subsector retail motor vehicles and parts from fixed point-of-sale locations. Establishments in this subsector typically operate from a showroom and/or an open lot where the vehicles are on display. The display of vehicles and the related parts require little by way of display equipment. The personnel generally include both the sales and sales support staff familiar with the requirements for registering and financing a vehicle as well as a staff of parts experts and mechanics trained to provide repair and maintenance services for the vehicles. Specific industries have been included in this subsector to identify the type of vehicle being retailed."

This subsector consists of three industry groups:

- NAICS 4411 - Automobile Dealers include all new car and used car dealers such as Ford, Chevy, Toyota, etc.
- NAICS 4412 - Other Motor Vehicle Dealers sell RV, motorcycle, and all other motor vehicles.
- NAICS 4413 - Automotive Parts, Accessories, and Tire Stores sell parts, accessories, and tires; e.g., AutoZone Auto Parts.

4.2.2 NAICS 441 Motor Vehicle and Parts Dealers: 2012 Economic Census

In 2012, the Motor Vehicle and Parts Dealers subsector operates 116,482 retail establishments and is reported to have an annual sales revenue of nearly 0.87 trillion U.S. dollars, which accounts for 20.59% of all sales in the retail trade sector. Among the three industries, Automobile Dealers generates about 0.74 trillion U.S. dollars sales, which is 85.15%

of the total sales within the subsector. Other Motor Vehicle Dealers has the smallest share of sales – 5.44% in the subsector. Automotive Parts, Accessories, and Tire Stores has a sales revenue of 81,734 million U.S. dollars, which accounts for 9.41% of total sales.

In the Automobile Dealers industry, New Car Dealers accounts for over 90% of the sales, while Used Car Dealers has only less than 10% of the market share. In the Other Motor Vehicle Dealers industry, Motorcycle, Boat, and Other Motor Vehicle Dealers contributes about 70% of the sales, while Recreational Vehicle Dealers has 30% of the market share. In the Automotive Parts, Accessories, and Tire Stores industry, Automotive Parts and Accessories takes about 59% of the market share, which leaves Tire Dealers the rest 41% of the sales.

In terms of the labor market, the Economic Census only provides data on annual payroll, first-quarter payroll, and the number of employees paid in the first quarter. To understand each retail subsector through the entire census year, I have estimated the number of employees paid in the year 2012, employee' projected annual salary, and employee's relative salary to the entire retail trade sector annual salary. See Appendix B for details.

Motor Vehicle and Parts Dealers employs approximately 1.8 million workers in total, which accounts for 11.79% of all retail employees. The average annual salary for the entire subsector is $39,499 and it is 1.6 times the annual salary of the entire retail trade sector. Among the three industries within the subsector, workers in the Automobile Dealers industry make the highest annual salary.

Key statistics from the 2012 Economic Census are summarized in Table 4.2. As indicated in Table 4.2, Automotive Parts, Accessories, and Tire Stores has more stores than Automobile Dealers but fewer employees. This evidence suggests that Automobile Dealers requires a greater number of employees to perform services to customers, compared to Automotive Parts, Accessories, and Tire Stores. Moreover, annual salary indicates that employees at Automobile Dealers especially New Car Dealers receive higher wages than others, and therefore they are expected to perform highest quality of service (Ingene & Lusch 1980). After carefully reviewing the numbers below in Table 4.2, one will notice that the number of stores, number of employees, and total sales revenue are not in proportion in Other Motor Vehicle Dealers and Automotive Parts, Accessories, and Tire Stores, compared with Automobile Dealers. Furthermore, Used Car Dealers encompasses more stores than New Car Dealers but only generates a small portion of the total sales in the automobile industry. For the abovementioned reasons, I will only use data from NAICS 44111 – New Car Dealers for this study, because the inclusion of data from other industries in this subsector may distort information on the key variables (*i.e.*, number of employees per store and relative wage rate).

Table 4.2. NAICS 441 Motor Vehicle and Parts Dealers – 2012 Economic Census

NAICS Code	Subsector/Industry Name	# of Retail Stores	Estimated # of Employees	Estimated Annual Salary ($)	Relative Salary	Total Sales ($M)	% of Total Sales
441	**Motor vehicle and parts dealers**	**116,482**	**1,776,399**	**39,499**	**1.61**	**868,805**	**100%**
4411	*Automobile dealers*	*45,223*	*1,151,007*	*45,181*	*1.84*	*739,777*	*85.15%*
44111	**New car dealers**	**21,292**	**1,022,019**	**46,500**	**1.90**	**672,550**	**77.41%**
44112	Used car dealers	23,931	128,988	34,725	1.42	67,227	7.74%
4412	*Other motor vehicle dealers*	*14,249*	*144,486*	*32,517*	*1.33*	*47,294*	*5.44%*
44121	Recreational vehicle dealers	2,605	36,894	36,119	1.47	14,245	1.64%
44122	Motorcycle, boat, and other motor vehicle dealers	11,644	107,592	31,282	1.28	33,050	3.80%
4413	*Automotive parts, accessories, and tire stores*	*57,010*	*480,906*	*27,997*	*1.14*	*81,734*	*9.41%*
44131	Automotive parts and accessories stores	36,710	313,329	24,347	0.99	48,308	5.56%
44132	Tire Dealers	20,300	167,577	34,820	1.42	33,426	3.85%

4.3 NAICS 442 Furniture and Home Furnishing Stores

4.3.1 NAICS 442 Furniture and Home Furnishing Stores: Definition

Furniture and Home Furnishing Stores is the second smallest subsector in the retail trade sector, according to the statistics of the 2012 Economic Census. NAICS defines 442 as follows.

> "Industries in the Furniture and Home Furnishings Stores subsector retail new furniture and home furnishings from fixed point-of-sale locations. Establishments in this subsector usually operate from showrooms and have substantial areas for the presentation of their products. Many offer interior decorating services in addition to the sale of products."

This subsector consists of two industry groups:

- NAICS 4421 - Furniture Stores include stores selling all kinds of furniture; e.g., IKEA, Sur La Table.
- NAICS 4422 - Home Furnishings Stores consist of stores selling new home furnishings (except furniture); e.g., Kohler Co.

4.3.2 NAICS 442 Furniture and Home Furnishing Stores: 2012 Economic Census

In 2012, the Furniture and Home Furnishing Stores subsector operates 51,635 retail stores and is reported to generate an annual sales revenue of 89,058 million U.S. dollars, which accounts for 2.11% of all sales in the retail trade sector. Of the two industries, Furniture Stores generates 48,784 million U.S. dollars sales, which equals 54.78% of the total sales within the subsector. Home Furnishing Stores is reported to have an annual sales revenue of 40,274 million U.S. dollars, of which 63.4% comes from the subindustry Other Home Furnishings Stores and 36.6% from Floor Covering Stores.

In terms of the labor market, there are approximately 434,759 people working in this subsector and it is about 2.89% of all retail employees. The average annual salary is $26,418 for this subsector, and it is about the same rate as the annual salary for the retail trade sector.

Key statistics from the 2012 Economic Census are summarized in Table 4.3. As indicated in Table 4.3, Furniture Stores employees receive a higher wage than employees from Home Furnishing Stores, on average. This is consistent with the theory that employees are paid more to provide more services in terms of both quantity and quality.

Table 4.3. NAICS 442 Furniture and Home Furnishing Stores – 2012 Economic Census

NAICS Code	Subsector/Industry Name	# of Retail Stores	Estimated # of Employees	Estimated Annual Salary ($)	Relative Salary	Total Sales ($M)	% of Total Sales
442	**Furniture and home furnishings stores**	**51,635**	**434,759**	**26,418**	**1.08**	**89,058**	**100%**
4421	*Furniture stores*	*23,657*	*193,450*	*32,847*	*1.34*	*48,784*	*54.78%*
44211	Furniture stores	23,657	193,450	32,847	*1.34*	48,784	54.78%
4422	*Home furnishings stores*	*27,978*	*241,309*	*21,264*	*0.87*	*40,274*	*45.22%*
44221	Floor covering stores	11,373	64,039	33,730	1.38	14,744	16.55%
44229	Other home furnishings stores	16,605	177,270	16,761	0.68	25,530	28.67%

4.4 NAICS 443 Electronics and Appliance Stores

4.4.1 NAICS 443 Electronics and Appliance Stores: Definition

The Electronics and Appliance Stores subsector is a small component of the retail trade sector. According to the 2012 Economic Census, it accounts for 2.43% of all retail sales. NAICS defines 443 as follows.

> “Industries in the Electronics and Appliance Stores subsector retail new electronics and appliances from point-of-sale locations. Establishments in this subsector often operate from locations that have special provisions for floor displays requiring special electrical capacity to accommodate the proper demonstration of the products. The staff includes sales personnel knowledgeable in the characteristics and warranties of the line of goods retailed and may also include trained repair persons to handle the maintenance and repair of the electronic equipment and appliances. The classifications within this subsector are made principally on the type of product and knowledge required to operate each type of store.”

This subsector consists of only one industry group, which is Electronics and Appliance Stores (NAICS 4431), however it is divided up into two subindustries at the 6-digit level of NAICS. These two subindustries are as follows.

- NAICS 443141 - Household Appliance Stores sell new household appliances, such as refrigerators, dishwashers, vacuum cleaners etc.; e.g., Whirlpool, GE Appliances.
- NAICS 443142 - Electronics Stores sell consumer electronic products such as TV, computers, etc.; e.g., Best Buy, Samsung.

4.4.2 NAICS 443 Electronics and Appliance Stores: 2012 Economic Census

In 2012, Electronics and Appliance Stores operates 48,826 retail stores and generates an annual sales revenue of 10,202 million U.S. dollars, which accounts for 2.43% of all sales in the retail trade sector. Of the two subindustries, Electronics Stores dominates the industry and contributes 84.24% of total sales revenue.

In terms of the labor market in this subsector, there are approximately 431,209 workers in total, which accounts for 2.86% of all retail workers. The average annual salary for the entire subsector is $23,660 and it is a little less than the average annual salary of the retail trade sector.

Key statistics from the 2012 Economic Census are summarized in Table 4.4. As indicated from average annual salary in Table 4.4, Household Appliance Stores require higher quality of service than Electronics Stores.

Table 4.4. NAICS 443 Electronics and Appliance Stores – 2012 Economic Census

NAICS Code	Subsector/Industry Name	# of Retail Stores	Estimated # of Employees	Estimated Annual Salary ($)	Relative Salary	Total Sales ($M)	% of Total Sales
443	**Electronics and appliance stores**	**48,826**	**431,209**	**23,660**	**0.97**	**102,597**	**100%**
4431	*Electronics and appliance stores*	*48,826*	*431,209*	*23,660*	*0.97*	*102,597*	*100%*
44314	Electronics and appliance stores	48,826	431,209	23,660	0.97	102,597	100%
443141	Household appliance stores	8,305	63,006	30,193	1.23	16,168	15.76%
443142	Electronics stores	40,521	368,203	22,542	0.92	86,429	84.24%

4.5 NAICS 445 Food and Beverage Stores

4.5.1 NAICS 445 Food and Beverage Stores: Definition

Food and Beverage Stores is the third largest subsector in the retail trade sector in terms of sales revenue, given the statistics in the 2012 Economic Census. NAICS defines 445 as follows.

> "Industries in the Food and Beverage Stores subsector usually retail food and beverages merchandise from fixed point-of-sale locations. Establishments in this subsector have special equipment (e.g., freezers, refrigerated display cases, refrigerators) for displaying food and beverage goods. They have staff trained in the processing of food products to guarantee the proper storage and sanitary conditions required by regulatory authority."

This subsector consists of three industry groups:

- NAICS 4451 - Grocery Stores consist of supermarkets and convenience stores; e.g., Kroger, Winco Foods, 7 Eleven.
- NAICS 4452 - Specialty Food Stores include stores selling meat, fish and seafood, fruit and vegetables, and other specialty food including baked goods; e.g., Godiva Chocolatier.
- NAICS 4453 - Beer, Wine, and Liquor Stores sell packaged alcoholic beverages; e.g., a local wine and spirits shop.

4.5.2 NAICS 445 Food and Beverage Stores: 2012 Economic Census

In 2012, Food and Beverage Stores operates 147,579 retail stores and generates an annual sales revenue of 620,024 million U.S. dollars, which accounts for 14.69% of all sales in the retail trade sector. Among the three industry groups, Grocery Stores dominates this subsector with a 90.30% industry share.

Food and Beverage Stores provides the largest labor market in the retail trade sector in 2012. This subsector employs nearly 2.9 million workers, which is equivalent to 19.19% of all retail employees. These workers make an average annual salary of $20,944, which is only 86% of the average retail annual salary.

Key statistics from the 2012 Economic Census are summarized in Table 4.5. As indicated in Table 4.5, Supermarkets and Other Grocery (except Convenience) Stores dominate this subsector with over 80% of the retail sales and on average they have the highest number of employees per store and the highest wage among all the industries. This further demonstrates that service (*i.e.*, employees) does matter. Since this is the dominant industry in this retail subsector, and the inclusion of data from other industries may distort information on the key variables (*i.e.*, number of employees per store and relative wage rate) needed for this study, I will only use data from NAICS 44511 – Supermarkets and Other Grocery (except Convenience) Stores for the analysis.

Table 4.5. NAICS 445 Food and Beverage Stores – 2012 Economic Census

NAICS Code	Subsector/Industry Name	# of Retail Stores	Estimated # of Employees	Estimated Annual Salary ($)	Relative Salary	Total Sales ($M)	% of Total Sales
445	**Food and beverage stores**	**147,579**	**2,890,552**	**20,944**	**0.86**	**620,024**	**100%**
4451	*Grocery stores*	*92,849*	*2,584,068*	*21,309*	*0.87*	*559,838*	*90.30%*
44511	**Supermarkets and other grocery (except convenience) stores**	**66,343**	**2,464,589**	**21,569**	**0.88**	**537,322**	**86.66%**
44512	Convenience Stores	26,506	119,479	15,961	0.65	22,516	3.64%
4452	*Specialty food stores*	*22,105*	*146,960*	*16,514*	*0.67*	*17,560*	*2.83%*
44521	Meat markets	5,365	37,050	19,476	0.80	5,795	0.93%
44522	Fish and seafood markets	1,964	11,335	19,703	0.80	2,105	0.34%
44523	Fruit and vegetable markets	2,761	21,990	18,281	0.75	3,584	0.58%
44529	Other specialty food stores	12,015	76,585	14,101	0.58	6,075	0.98%
4453	*Beer, wine, and liquor stores*	*32,625*	*159,524*	*19,102*	*0.78*	*42,626*	*6.87%*
44531	Beer, wine, and liquor stores	32,625	159,524	19,102	0.78	42,626	6.87%

4.6 NAICS 447 Gasoline Stations

4.6.1 NAICS 447 Gasoline Stations: Definition

Gasoline Stations is the fourth largest subsector in the retail trade sector in 2012. NAICS defines 447 as the following.

> "Industries in the Gasoline Stations subsector retail automotive fuels (e.g., gasoline, diesel fuel, gasohol, alternative fuels) and automotive oils or retail these products in combination with convenience store items. These establishments have specialized equipment for storing and dispensing automotive fuels."

The Gasoline Stations subsector has only one industry at the 4-digit NAICS code level, which is Gasoline Stations, however, it is divided into two subindustries at the 5-digit level. The two subindustries are:

- NAICS 44711 - Gasoline Stations with Convenience Stores sell gasoline and food products.
- NAICS 44719 - Other Gasoline Stations only sell gasoline and do not include convenience stores.

4.6.2 NAICS 447 Gasoline Stations: 2012 Economic Census

In 2012, there are 114,474 gasoline stations in the United States. The whole subsector contributes an annual sales revenue of 554,256 million U.S. dollars, which accounts for 13.13% of all retail sales. Of the two subindustry groups, Gasoline Stations with Convenience Stores dominates the industry with 77.06% of sales in this subsector.

This subsector employs nearly 0.89 million workers in 2012, and this figure accounts for 5.9% of all retail employees. The average annual salary is $17,509, and it is about 72% of the average retail annual salary.

Key statistics from the 2012 Economic Census are summarized in Table 4.6. As indicated in Table 4.6, Other Gasoline Stations (*i.e.*, with no convenience stores) has a higher sales revenue per store, higher number of employees per store, and higher wage than Gasoline Stations with Convenience Stores. This observation indicates that gasoline stations may be better off to focus only on gasoline business and not include convenience stores within their business.

Table 4.6. NAICS 447 Gasoline Stations – 2012 Economic Census

NAICS Code	Subsector/Industry Name	# of Retail Stores	Estimated # of Employees	Estimated Annual Salary ($)	Relative Salary	Total Sales ($M)	% of Total Sales
447	**Gasoline stations**	**114,474**	**889,378**	**17,509**	**0.71**	**554,256**	**100%**
4471	*Gasoline stations*	114,474	889,378	17,509	0.71	554,256	100%
44711	Gasoline stations with convenience stores	97,394	743,978	16,633	0.68	427,090	77.06%
44719	Other gasoline stations	17,080	145,400	21,995	0.90	127,166	22.94%

4.7 NAICS 448 Clothing and Clothing Accessories Stores

4.7.1 NAICS 448 Clothing and Clothing Accessories Stores: Definition

Clothing and Clothing Accessories Stores is a small part of the retail trade sector. NAICS defines 448 as follows.

> "Industries in the Clothing and Clothing Accessories Stores subsector retail new clothing and clothing accessories from fixed point-of-sale locations. Establishments in this subsector have similar display equipment and staff that is knowledgeable regarding fashion trends and the proper match of styles, colors, and combinations of clothing and accessories to the characteristics and tastes of the customer."

There are three industry groups within the subsector:

- NAICS 4481 - Clothing Stores sell men's, women's, children's & family clothing, and other clothing including sportwear, workwear, etc.; e.g., REI, Gap, Zara, Lululemon, Carhartt.
- NAICS 4482 - Shoes Stores sell footwear; e.g., DSW, Skechers.
- NAICS 4483 - Jewelry, Luggage, and Leather Goods Stores sell jewelry, luggage, and leather goods and accessories; e.g., Reeds Jewelers, Samsonite, Coach.

4.7.2 NAICS 448 Clothing and Clothing Accessories Stores: 2012 Economic Census

In 2012, Clothing and Clothing Accessories operates 147,709 retail stores in the United States. This subsector is reported to generate an annual sales revenue of 233,812 million U.S. dollars, which accounts for 5.54% of all retail sales. Among the three industry groups, Clothing Stores dominates with 73.73% of total sales revenue in this subsector.

This subsector employs nearly 1.7 million workers in 2012, which accounts for 11.28% of all retail employees. The average annual salary for this subsector is $16,473, which is 67% of the average retail annual salary.

Key statistics from the 2012 Economic Census are summarized in Table 4.7. As indicated in Table 4.7, Jewelry, Luggage, and Leather Goods Stores employees have the highest wage among all industries. This observation indicates that for high-priced items such as jewelry, luggage, and leather goods, retailers tend to provide higher quality of service.

Table 4.7. NAICS 448 Clothing and Clothing Accessories Stores – 2012 Economic Census

NAICS Code	Subsector/Industry Name	# of Retail Stores	Estimated # of Employees	Estimated Annual Salary ($)	Relative Salary	Total Sales ($M)	% of Total Sales
448	**Clothing and clothing accessories stores**	**147,709**	**1,699,734**	**16,473**	**0.67**	**233,812**	**100%**
4481	*Clothing stores*	*97,721*	*1,360,288*	*15,117*	*0.62*	*172,385*	*73.73%*
44811	Men's clothing stores	7,354	52,842	24,859	1.01	8,040	3.44%
44812	Women's clothing stores	36,344	375,875	14,541	0.59	40,983	17.53%
44813	Children's and infants' clothing stores	6,962	85,201	11,795	0.48	9,368	4.01%
44814	Family clothing stores	27,770	678,461	14,710	0.60	92,136	39.41%
44815	Clothing accessories stores	8,259	54,686	18,858	0.77	9,031	3.86%
44819	Other clothing stores	11,032	113,223	15,621	0.64	12,825	5.49%
4482	*Shoe stores*	*25,551*	*205,498*	*16,852*	*0.69*	*30,759*	*13.15%*
44821	Shoe stores	25,551	205,498	16,852	0.69	30,759	13.15%
4483	*Jewelry, luggage, and leather goods stores*	*24,437*	*133,948*	*29,662*	*1.21*	*30,668*	*13.12%*
44831	Jewelry stores	23,477	128,736	29,498	1.20	28,295	12.10%
44832	Luggage and leather goods stores	960	5,212	33,689	1.38	2,374	1.02%

4.8 NAICS 452 General Merchandise Stores

4.8.1 NAICS 452 General Merchandise Stores: Definition

General Merchandise Stores is the second largest subsector in the retail trade sector, according to the statistics in the 2012 Economic Census. NAICS defines 452 as the following.

> "Industries in the General Merchandise Stores subsector retail new general merchandise from fixed point-of-sale locations. Establishments in this subsector are unique in that they have the equipment and staff capable of retailing a large variety of goods from a single location. This includes a variety of display equipment and staff trained to provide information on many lines of products."

General Merchandise Stores consists of two distinct industry groups – Department Stores (NAICS 4521) and Other General Merchandise Stores (NAICS 4529).

- NAICS 4521 - Department Stores sell general merchandise including apparel, jewelry, home furnishings, etc. with no one merchandise line predominating. Department stores may sell perishable groceries, but such sales are insignificant. Department stores encompass both traditional department stores and discount department stores; e.g., Nordstrom, Macy's, Nordstrom Rack, T.J. Maxx, Marshall's.
- NAICS 4529 - Other General Merchandise Stores sell general merchandise in general merchandise stores other than Department Stores. Merchandise includes apparel, home furnishing, groceries, and even automotive parts, but with no one merchandise line predominating. Retail establishments known as warehouse clubs, supercenters, and dollar stores are included in this industry; e.g., Costco, Sam's Club, Walmart Supercenter, Dollar Tree.

There is a huge difference between Department Stores and Other General Merchandise Stores, in terms of the sales of groceries. Sales of groceries are insignificant at Department Stores, while for Other General Merchandise Stores, especially Warehouse Clubs and Supercenters, sales of groceries are important and significant. There is one important note that the General Merchandise Stores subsector retails various lines of merchandise and obviously it is in competition with many other subsectors in retail trade, such as Furniture and Home Furnishing Stores (NAICS 442), Electronics and Appliance Stores (NAICS 443), Food and Beverage Stores (NAICS 445), Clothing and Clothing Accessories Stores (NAICS 448), and even Motor Vehicle and Parts Dealers (NAICS 441) and Gasoline Stations (NAICS 447).

4.8.2 NAICS 452 General Merchandise Stores: 2012 Economic Census

In 2012, General Merchandise Stores is the second largest subsector in the retail trade sector. This subsector operates 49,147 retail establishments in the United States and generates an annual sales revenue of 640,627 million U.S. dollars, which accounts for 15.18% of all retail sales. Department Stores takes 27.68% of the market share, while Other General Merchandise Stores accounts for 72.32% of the market. In the Department Stores industry, sales of Discount Department Stores are about twice the size of traditional Department Stores. In the Other General Merchandise Stores industry, Warehouse Clubs and Supercenters dominates the industry with 87.70% of the market share.

General Merchandise Stores is also the second largest subsector in terms of the labor market. In total, this subsector employs 2.81 million workers in 2012, which accounts for 18.68% of all retail employees. On average, workers make an annual salary of $20,630, which is below the average retail annual salary.

Key statistics of the 2012 Economic Census are summarized in Table 4.8. As indicated in Table 4.8, Warehouse Clubs and Supercenters dominates this retail subsector with over 60% of total sales revenue and over 50% of retail employees but only occupy about 10% of retail stores. Moreover, they have the highest number of employees per store and highest wage rate in this subsector, even higher than traditional department stores. Another important note is that All Other General Merchandise Stores (*i.e.*, dollar stores) occupies the most of retail stores but employs the smallest number of employees with the lowest wage rate. This observation is not surprising because these dollar stores are almost self-serviced with minimum service level and quality from the employees. Since the inclusion of this industry will distort the data on key variables (*i.e.*, number of employees per store and relative wage rate), I will not include this industry in the analysis. Furthermore, traditional department store, discount department stores, and warehouse clubs and supercenters present significantly different types of businesses, and therefore I use the data on these three industries respectively in the analysis, rather than aggregating them into one single retail line.

Table 4.8. NAICS 452 General Merchandise Stores – 2012 Economic Census

NAICS Code	Subsector/Industry Name	# of Retail Stores	Estimated # of Employees	Estimated Annual Salary ($)	Relative Salary	Total Sales ($M)	% of Total Sales
452	**General merchandise stores**	**49,147**	**2,813,737**	**20,630**	**0.84**	**640,627**	**100%**
4521	*Department stores*	*8,064*	*1,066,298*	*18,193*	*0.74*	*177,313*	*27.68%*
45211	Department stores	8,064	1,066,298	18,193	0.74	177,313	27.68%
452111	**Department stores (except discount department stores)[14]**	**3,339**	**454,631**	**19,028**	**0.78**	**62,961**	**9.83%**
452112	**Discount department stores**	**4,725**	**611,667**	**17,573**	**0.72**	**114,352**	**17.85%**
4529	*Other general merchandise stores*	*41,083*	*1,747,439*	*22,117*	*0.90*	*463,314*	*72.32%*
45291	**Warehouse clubs and supercenters**	**5,114**	**1,400,826**	**23,757**	**0.97**	**406,309**	**63.42%**
45299	All other general merchandise stores	35,969	346,613	15,490	0.63	57,005	8.90%

4.9 General Observation: 2012 Economic Census

Given the conceptual theorizing provide in Chapter 3 and a lengthy review of the literature in Chapter 2, one can infer that the number of retail stores is driven by demand (*i.e.*, population, population density) and store size. For instance, Food and Beverage and Clothing subsectors have the highest number of retail stores, while General Merchandise has the second smallest number of retail stores but generates the second highest sales revenue among all retail subsectors. This indicates that on average General Merchandise Stores serves a large body of

[14] Department stores (except discount department stores) refers to traditional department stores such as Macy's, Nordstrom, Neiman Marcus, etc. Therefore, for simplicity, I will use the wording "traditional department stores" to refer to this industry in the following discussions.

population in the area with potentially large assortments. With a simple visual inspection of the data, one can also infer that average annual salary is a good indicator of the quality of service provided by retailers. For instance, employees at New Car Dealers get paid with the highest average wage in the subsector. Intuitively, they are expected to perform higher than average quality of service for customers. This is also evident for Department Stores. Traditional department stores (e.g., Macy’s) pay their employees with a higher wage rate than discount department stores (e.g., TJ Maxx), because service does matter for traditional department stores.

In summary, the 2012 Economic Census provides rich information on the retail trade sector. Not only am I able to incorporate intertype and intratype competition in the analysis, but I can also examine the extent to which consumers’ out-of-town shopping occurs across these subsectors. As discussed earlier, to avoid the distortion on key marketing mix variables, I exclude some of the industries from the seven subsectors in the analysis.

Chapter 5 Research Methodology

This chapter presents information regarding the research design of this dissertation. I discuss details on the research data and the research methods used for the empirical analysis.

5.1 Research Data

This section of the chapter provides information about the data used in this dissertation, including data collection, data description and basic statistics, and data processing procedures.

5.1.1 Data Collection

The data used for this dissertation primarily comes from two sources: the 2012 Economic Census and the 2012 American Community Survey (5-year estimates). Information regarding retail stores across the seven subsectors are obtained from the 2012 Economic Census, while details on household characteristics are from the 2012 American Community Survey 5-year estimates (2008 – 2012).

The Economic Census gathers data via three methods:

- All large- and medium-size firms and all multi-establishment firms receive report forms to be completed for each of their establishments[15] and returned to the Census Bureau.
- A sample of small employers (those with paid employees) also receive report forms. This sample consists of single-establishment firms with payroll below a specified cut off.
- For those very small firms, data from existing administrative records of other federal agencies are used. These records provide basic information on the firms including location, type of business, receipts, payroll, number of employees etc.

[15] "Establishments" and "stores" are used interchangeably.

The American Community Survey is an ongoing survey which provides important information regarding the nation and its people on a yearly basis. The methodology of American Community Survey is as follows.

> "It uses a series of monthly samples to produce annually updated estimates for the same small areas (census tracts and block groups) formerly surveyed via the decennial census long-form sample. Initially, five years of samples were required to produce these small-area data. Once the Census Bureau, released its first 5-year estimates in December 2010; new small-area statistics now are produced annually. The Census Bureau also will produce 3-year and 1-year data products for larger geographic areas." (U.S. Census Bureau)

5.1.2 Data Description & Basic Statistics

In this subsection, I set forth the variables used in the empirical analysis. Note that, the unit of analysis is each metropolitan or micropolitan area, and therefore all the dependent variables and independent variables are an average for each of the areas. There are 917 metropolitan and micropolitan areas (*i.e.*, 381 metropolitan & 536 micropolitan areas) in total from the 2012 Economic Census. However, since the 2012 American Community Survey does not cover all 917 metropolitan and micropolitan areas, after merging data from the two sources, I obtain a sample consisting of 735 metropolitan and micropolitan areas (*i.e.*, 354 metropolitan & 381 micropolitan areas).

Some of the metropolitan areas are excluded because there is no data in the 2012 ACS regarding those areas. For instance, Los Angeles-Long Beach-Anaheim, CA Metro Area, Santa Maria-Santa Barbara, CA Metro Area, Urban Honolulu, HI Metro Area, Lafayette-West Lafayette, IN Metro Area are not included in the analysis, possibly due to the change in delineation of those areas. Since the 2012 ACS (5-year estimates) uses data from 2008 to 2012, it is based on the delineation as of 2008, and the 2012 Economic Census is based on the delineation as of 2012. Therefore, it is likely that changes were made between 2008 and 2012 such that some counties were added into a metropolitan or micropolitan area, while some might be removed from one area. For a complete list of the missing 182 areas, please refer to Appendix C. Furthermore, many micropolitan areas are excluded due to confidentiality and privacy concerns, since retail data that may lead to identify an individual business are not disclosed in certain areas. These excluded areas have a wide range on the number of stores starting from 1 to over 100. Therefore, the sample size varies across each of the seven retail subsectors (*i.e.*, nine retail lines).

Dependent variables, which are of interest to this dissertation, are household expenditures (*i.e.*, total sales divided by total number of households) in each of the nine retail lines[16] in the United States.

Independent variables consist of socioeconomic and demographic characteristics, marketing mix variables, competition variables, consumer out-shopping variables, and control variables. Socioeconomic and demographic characteristics include household type, household income distribution, household median income, age distribution, educational attainment, gender, race, percentage of group quarters (*i.e.*, the percentage of people living or staying in group living arrangements such as university housing, prisons, nursing homes), household size, total population, population density[17], household mobility (*i.e.*, automobiles per household), population growth rate, and unemployment rate.

For the marketing mix variables, given the limitation of secondary data, I examine only one aspect of the marketing mix – service. I use the number of employees per retail store as a proxy for the quantity of service, and the relative wage rate of a specific industry to the total

[16] Certain exclusions of data apply to NAICS 441 Motor Vehicle and Parts Dealers, NAICS 445 Food and Beverage Stores, and NAICS 452 General Merchandise Stores.

[17] Population density = Total Population/Total Land Area. The 2012 land areas of metropolitan and micropolitan areas are calculated from the 2012 county land areas with 2009 delineation of metropolitan and micropolitan areas (U.S. Census Bureau). Note that the delineation remains the same from Dec. 2009 to Feb. 2013.

retail average wage as a proxy for the quality of service. Scholars have used the number of employees per square footage (e.g., Ingene & Lusch 1980) as a proxy for the quantity of service, however, the Economic Census no longer publishes data on store size, therefore, I adopt another approach – using number of employees per store as a proxy for quantity of service. Although this approach of using proxies from secondary sources is the optimal solution at hand with secondary data, one should note that this approach may not provide accurate measures for this marketing mix variable – service. As shown by Teas (1993) and Cronin et al. (2000), consumers' perceived service quality is a multi-dimensional construct, and it is unlikely to reflect all dimensions with proxies from secondary sources.

Competition here refers to both intertype and intratype competition. Intertype competition exists among different retail subsectors, while intratype competition occurs within the retail subsector – NAICS 452 General Merchandise Stores, where Department Stores (except Discount Department Stores), Discount Department Stores, and Warehouse Clubs and Supercenters are in competition with each other. Details about intertype and intratype competition are provided below in Table 5.1. For this set of variables, I use the average number of stores per 1000 households in the competing industry as the competitive variable for the focal retail subsector/industry. This is consistent with the retail literature (Chen et al. 2020).

Table 5.1. Intertype and Intratype Competition

Retail Subsector/Industry	Intertype Competition with	Intratype Competition with
NAICS 44111 New Car Dealers	None	None
NAICS 442 Furniture Stores	NAICS 45291	None
NAICS 443 Electronics	NAICS 45291	None
NAICS 44511 Supermarkets and Other Grocery (except Convenience) Stores	NAICS 45291	None
NAICS 447 Gasoline Stations	None	None
NAICS 448 Clothing & Accessories Stores	NAICS 452111, NAICS 452112, & NAICS 45291	None
NAICS 452111 Department Stores (except Discount Department Stores)	NAICS 448	NAICS 452112 & NAICS 45291
NAICS 452112 Discount Department Stores	NAICS 448	NAICS 452111 & NAICS 45291
NAICS 45291 Warehouse Clubs & Supercenters	NAICS 442, NAICS 443, NAICS 44511, & NAICS 448	NAICS 452111 & NAICS 452112

To examine consumers' out-shopping phenomenon, I construct a variable named "neighbor" to measure whether one area is far away or close enough to be a neighbor area of the focal area so that consumers are willing to travel out-of-town for shopping purposes. To construct the variable, I follow the procedures from Holian & Kahn (2015) and Gardner & Hendrickson (2018). First, I obtain the 2012 list of principal cities of all 735 metropolitan and micropolitan areas from the Census Bureau. Second, the central location of each metropolitan and micropolitan area's principal city is obtained by recording the latitude and longitude returned when entering the principal city name in the Batch Geocode Tool from Google Map Developers[18]. Third, I then calculate the travel distance between each combination of all 735

[18] The geocode obtained from Google Map Developers is the same as from Google Earth.

areas using the *georoute* command in Stata (Weber & Péclat 2017). This command utilizes HERE Maps API. Lastly, I set distances of 100 miles and 200 miles as cut-off points of nearby areas, and then I calculate the population from areas within these two different radiuses. Specifically, I differentiate whether the population is from a nearby area where the population is larger than the focal area, or smaller than or equal to the focal area. According to Reilly (1931), if there exists a nearby area with large population and easy to travel to, it is highly possible that consumers would choose to go and shop in this large area, rather than the focal area or even a smaller area. For interpretation purposes, this set of population variables are in millions.

To control for the size of the area and different levels of household expenditures in the retail sector, I add MSA type (*i.e.*, whether it is a metropolitan or micropolitan area) and total retail expenditure per household as control variables in the analysis. The latter is in thousand dollars.

Table 5.2 provides variable definitions. Descriptive statistics (*i.e.*, number of observations, mean, standard deviation, minimum, and maximum) are presented in Table 5.3.

Table 5.2. Variable Definitions

	Name	Symbol	Definition	Data Source
Dependent Variables	Sales per Household (*i.e.*, Household Expenditures)	sales_car	Average expenditure per household at New Car Dealers.	Calculated from 2012 Economic Census & 2012 ACS (5-year)
		sales_fur	Average expenditure per household at Furniture Stores.	
		sales_elec	Average expenditure per household at Electronics Stores.	
		sales_food	Average expenditure per household at Supermarkets and Other Grocery (except convenience) Stores.	
		sales_gas	Average expenditure per household at Gasoline Stations.	
		sales_cloth	Average expenditure per household at Clothing & Accessories Stores.	
		sales_dept	Average expenditure per household at Department Stores (except Discount Department Stores).	
		sales_ddept	Average expenditure per household at Discount Department Stores.	
		sales_super	Average expenditure per household at Warehouse Clubs and Supercenters	
Socioeconomic & Demographic Variables	Household Type	hh_mc	Percentage of married couple households.	2012 ACS (5-year)
		hh_sh	Percentage of single head households.	
		hh_nf	Percentage of non-family households.	
	Income	income1	Percentage of households with income between $0-24,999.	
		income2	Percentage of households with income between $25,000-49,999.	

Table 5.2. Variable Definitions (Cont.)

	Name	Symbol	Definition	Data Source
Socioeconomic & Demographic Variables		income3	Percentage of households with income between $50,000-74,999.	
		income4	Percentage of households with income between $75,000-99,999.	
		income5	Percentage of households with income between $100,000-149,999.	
		income6	Percentage of households with income between $150,000-199,999.	
		income7	Percentage of households with income $200,000 and above.	
	Household Median Income	mdincome	Median household income.	
	Age	age1	Percentage of population age between 0-4.	
		age2	Percentage of population age between 5-14.	
		age3	Percentage of population age between 15-19.	
		age4	Percentage of population age between 20-24.	
		age5	Percentage of population age between 25-34.	
		age6	Percentage of population age between 35-44.	
		age7	Percentage of population age between 45-59.	
		age8	Percentage of population age between 60-64.	
		age9	Percentage of population age between 65-74.	
		age10	Percentage of population age between 75-84.	
		age11	Percentage of population age 85 and above.	
	Education	edu1	Percentage of population age 25+ with less than high school degree.	

Table 5.2. Variable Definitions (Cont.)

	Name	Symbol	Definition	Data Source
Socioeconomic & Demographic Variables		edu2	Percentage of population age 25+ with high school diploma.	
		edu3	Percentage of population age 25+ with some college or Associate degree.	
		edu4	Percentage of population age 25+ with Bachelor's degree.	
		edu5	Percentage of population age 25+ with Graduate or professional degree.	
	Gender	male	Percentage of male.	
		female	Percentage of female.	
	Race	whitealone	Percentage of Whitealone population.	
		black	Percentage of Black population.	
		asian	Percentage of Asian population.	
		hispanic	Percentage of Hispanic population.	
	Group Quarters	gq	Percentage of population living in group quarters.	
	Household Size	hhsize	Average population per household.	
	Total Population	pop	2012 Estimated Total Population	
	Population Density	density	Average population per square mile.	
	Mobility	auto	Average automobiles per household.	
	Population Growth	popgrowth	Average population growth rate from 2010 to 2012.	Calculated from 2010 Census & 2012 ACS (5-year)
	Unemployment Rate	unempl	Unemployment rate in the area.	

Table 5.2. Variable Definitions (Cont.)

	Name	Symbol	Definition	Data Source
Marketing mix Variables	Employees per Store	labor_car	Average number of employees at New Car Dealers.	Calculated from 2012 Economic Census
		labor_fur	Average number of employees at Furniture Stores.	
		labor_elec	Average number of employees at Electronic Stores.	
		labor_food	Average number of employees at Supermarkets and Other Grocery (except convenience) Stores.	
		labor_gas	Average number of employees at Gasoline Stations.	
		labor_cloth	Average number of employees at Clothing & Accessories Stores.	
		labor_dept	Average number of employees at Department Stores (except Discount Department Stores).	
		labor_ddept	Average number of employees at Discount Department Stores.	
		labor_super	Average number of employees at Warehouse Clubs and Supercenters.	
	Relative Wage Rate	wage_car	Average Annual Wage of New Car Dealers Employees/Average Annual Wage of Retail Employees	Calculated from 2012 Economic Census
		wage_fur	Average Annual Wage of Furniture Stores Employees/Average Annual Wage of Retail Employees	
		wage_elec	Average Annual Wage of Electronics Stores Employees/Average Annual Wage of Retail Employees	

Table 5.2. Variable Definitions (Cont.)

	Name	Symbol	Definition	Data Source
Marketing mix Variables		wage_food	Average Annual Wage of Supermarkets and Other Grocery (except convenience) Stores Employees/Average Annual Wage of Retail Employees	
		wage_gas	Average Annual Wage of Gasoline Stations Employees/Average Annual Wage of Retail Employees	
		wage_cloth	Average Annual Wage of Clothing & Accessories Stores Employees/Average Annual Wage of Retail Employees	
		wage_dept	Average Annual Wage of Department Stores (except Discount Department Stores) Employees/Average Annual Wage of Retail Employees	
		wage_ddept	Average Annual Wage of Discount Department Stores Employees/Average Annual Wage of Retail Employees	
		wage_super	Average Annual Wage of Warehouse Clubs and Supercenters Employees/Average Annual Wage of Retail Employees	
Intertype Competition Variables	Stores per Household	stores_fur	Average Number of Furniture Stores per Thousand Households	Calculated from 2012 Economic Census & 2012 ACS (5-year)
		stores_elec	Average Number of Electronics Stores per Thousand Households	
		stores_food	Average Number of Supermarkets and Other Grocery (except convenience) Stores per Thousand Households	

Table 5.2. Variable Definitions (Cont.)

	Name	Symbol	Definition	Data Source
Intertype Competition Variables		stores_cloth	Average Number of Clothing and Accessories Stores per Thousand Households	
		stores_dept	Average Number of Department Stores (except Discount Department Stores) per Thousand Households	
		stores_ddept	Average Number of Discount Department Stores per Thousand Households	
		stores_super	Average Number of Warehouse Clubs and Supercenters per Thousand Households	
Consumer Out-Shopping Variables	Population in Nearby Large Areas	pop_large100	Total population from nearby large areas within 100 miles of the focal area where their population are greater than the focal area	Calculated from 2012 ACS (5-year)
		pop_large200	Total population from nearby large areas within 200 miles of the focal area where their population are greater than the focal area	
Control Variables	MSA Type	msa_type	Metropolitan or micropolitan area.	Calculated from 2012 Economic Census
	Total Retail Expenditure per Household	sales_total	Total Retail Expenditure in Thousands per Household in Each Area	Calculated from 2012 Economic Census & 2012 ACS (5-year)

Table 5.3. Descriptive Statistics[19]

Variable	Number of Observations	Mean	Std. Dev.	Min	Max
sales_car (per HH)	558	$5,595.25	$2,298.593	$1,435.038	$19,719.76
sales_fur (per HH)	613	$618.664	$362.412	$28.670	$3,786.813
sales_elec (per HH)	607	$641.186	$377.723	$28.847	$3,541.415
sales_food (per HH)	505	$4,276.285	$1,894.199	$905.129	$28,232.42
sales_gas (per HH)	731	$6,000.302	$2,673.980	$1,216.077	$29,082.170
sales_cloth (per HH)	663	$1,338.490	$1,091.463	$26.775	$10,722.920
saels_dept (per HH)	213	$638.500	$251.872	$305.316	$2,866.249
saels_ddept (per HH)	225	$1,260.52	$693.960	$286.180	$6,764.542
saels_super (per HH)	57	$4,636.421	$2,083.501	$1,105.19	$12,223.61
hh_mc	735	0.502	0.045	0.334	0.705
hh_sh	735	0.166	0.036	0.075	0.342
hh_nf	735	0.332	0.043	0.159	0.486
income1	735	0.271	0.061	0.112	0.506
income2	735	0.267	0.028	0.148	0.347
income3	735	0.188	0.019	0.112	0.243
income4	735	0.117	0.020	0.054	0.188
income5	735	0.102	0.029	0.040	0.202
income6	735	0.030	0.015	0.006	0.114
income7	735	0.024	0.017	0.005	0.168
mdincome	735	$46,381.970	$8,657.586	$24,653	$89,940
age1	735	0.063	0.010	0.024	0.111
age2	735	0.129	0.016	0.049	0.194
age3	735	0.072	0.011	0.026	0.130
age4	735	0.073	0.028	0.027	0.221
age5	735	0.122	0.016	0.069	0.171
age6	735	0.124	0.011	0.077	0.161
age7	735	0.209	0.019	0.119	0.266
age8	735	0.059	0.010	0.028	0.120
age9	735	0.080	0.020	0.037	0.290

[19] As discussed earlier, the seven retail subsectors (*i.e.*, nine retail lines) have varying number of observations due to confidentiality and privacy concerns that retail data that may lead to identify an individual business are not disclosed at certain areas.

Table 5.3. Descriptive Statistics (Cont.)

Variable	Number of Observations	Mean	Std. Dev.	Min	Max
age10	735	0.048	0.013	0.016	0.126
age11	735	0.019	0.006	0.005	0.049
edu1	735	0.147	0.058	0.039	0.551
edu2	735	0.329	0.067	0.127	0.521
edu3	735	0.303	0.046	0.157	0.429
edu4	735	0.141	0.046	0.049	0.324
edu5	735	0.080	0.036	0.023	0.290
male	735	0.495	0.014	0.460	0.599
female	735	0.505	0.014	0.401	0.540
whitealone	735	0.757	0.183	0.013	0.976
black	735	0.093	0.119	0.001	0.709
asian	735	0.018	0.028	0.000	0.353
hispanic	735	0.101	0.144	0.007	0.983
gq	735	0.035	0.028	0.006	0.312
hhsize	735	2.626	0.227	2.165	4.104
pop	735	362,258.100	1,066,998	22,344	18.9 million
density	735	177.339	233.039	2.387	2,829.781
auto	735	1.844	0.129	1.218	2.284
popgrowth	735	0.000	0.002	-0.008	0.008
unempl	735	0.093	0.027	0.028	0.209
labor_car	558	41.716	14.149	7.188	86.706
labor_fur	613	7.407	2.489	1.9	21.648
labor_elec	607	7.642	2.617	2.080	16.816
labor_food	505	41.272	14.236	12.605	123.756
labor_gas	731	8.423	2.381	2.614	21.168
labor_cloth	663	9.065	2.629	2.559	16.475
labor_dept	213	123.201	40.824	51.819	327.805
labor_ddept	225	115.570	29.933	56.589	206.641
labor_super	57	247.009	45.655	136.091	321.238
wage_car	558	1.822	0.220	1.041	2.503
wage_fur	613	1.108	0.204	0.517	2.312
wage_elec	607	0.958	0.165	0.514	2.114

Table 5.3. Descriptive Statistics (Cont.)

Variable	Number of Observations	Mean	Std. Dev.	Min	Max
wage_food	505	0.877	0.128	0.523	1.338
wage_gas	731	0.742	0.088	0.494	1.099
wage_cloth	663	0.644	0.090	0.390	1.352
wage_dept	213	0.748	0.077	0.378	0.926
wage_ddept	225	0.717	0.076	0.534	1.046
wage_super	57	1.030	0.099	0.802	1.213
stores_fur (per 1000 HH)	734	0.466	0.216	0.095	2.755
stores_elec (per 1000 HH)	734	0.424	0.153	0.069	1.884
stores_food (per 1000 HH)	735	0.491	0.220	0.088	3.878
stores_cloth (per 1000 HH)	735	1.153	0.727	0.043	8.442
stores_dept (per 1000 HH)	541	0.052	0.027	0.010	0.279
stores_ddept (per 1000 HH)	584	0.058	0.031	0.010	0.279
stores_super (per 1000 HH)	710	0.060	0.026	0.010	0.222
pop_large100 (million)	735	1.573	2.695	0	26.643
pop_large200 (million)	735	6.227	7.031	0	39.548
msa_type	735	0.482	0.500	0	1
sales_total ($000)	735	35.081	11.205	13.747	197.638

5.1.3 Data Processing – Factor Analysis

In this subsection, I discuss the use of factor analysis on the set of socioeconomic and demographic variables. Since data from the American Community Survey (5-year estimates) represent the statistical facts about consumer and household in those metropolitan and micropolitan areas, factor analysis can be used on these variables to reveal the underlying consumer and household characteristics (*i.e.*, latent variables – factors) that drive these socioeconomic and demographic variables. This data reduction technique provides important consumer insights in those metropolitan and micropolitan areas, which serves as the basis of discussions on market segmentation in the following chapters.

Following Ingene's previous studies (e.g., Ingene 1983, 1984), I exclude one variable from each of the following variable groups: household type, income, age, education, gender, and race. Since variables in each of the abovementioned groups are percentages of the corresponding characteristics representing each area in the sample, thus they sum up to 100%, which means any variable can be expressed as a linear combination of all other variables in the variable group. Therefore, I choose to exclude the following variables in the factor analysis: *hh_sh*, *income3*, *age7*, *edu2*, *female*, and *whitealone*. There is no rule of thumb as to which variables should be excluded in the factor analysis, but one should note that the exclusion will yield a consumer profile with those selected variables and this consumer profile will become the baseline of comparisons in the succeeding data analysis and interpretation. Table 5.4 presents details about the excluded variables. From this table, one can tell that this is a typical type of an average American consumer/household: a female (single head of the household), with middle income ($50,000 – 74,999) and a high school degree, in her age of 45 to 59. For simplicity, I label this baseline consumer as "female single-head".

Table 5.4. Variables Excluded in Factor Analysis

Variable Group	Variable Symbol	Definition
Household Type	hh_sh	Percentage of single head households.
Income	income3	Percentage of households with income between $50,000-74,999.
Age	age7	Percentage of population age between 45-59.
Education	edu2	Percentage of population age 25+ with high school diploma.
Gender	female	Percentage of female.
Race	whitealone	Percentage of whitealone population.

A full list of the remaining socioeconomic and demographic variables for the factor analysis is provided in Table 5.5. Table 5.6 presents results of the Kaiser-Meyer-Olkin (KMO) test for sampling adequacy on those variables. The KMO statistic is a measure of how well the data are suitable for a factor analysis, and it is a number between 0 and 1. The closer it is to 1, the better the statistic and thus the better the data are for factor analysis.

Table 5.5. Variables Included in Factor Analysis

Socioeconomic & Demographic Variables	Symbol	Definition
Household Type	hh_mc	Percentage of married couple households.
	hh_nf	Percentage of non-family households.
Income	income1	Percentage of households with income between $0-24,999.
	income2	Percentage of households with income between $25,000-49,999.
	income4	Percentage of households with income between $75,000-99,999.
	income5	Percentage of households with income between $100,000-149,999.
	income6	Percentage of households with income between $150,000-199,999.
	income7	Percentage of households with income $200,000 and above.
Household Median Income	mdincome	Median household income.
Age	age1	Percentage of population age between 0-4.
	age2	Percentage of population age between 5-14.
	age3	Percentage of population age between 15-19.
	age4	Percentage of population age between 20-24.
	age5	Percentage of population age between 25-34.
	age6	Percentage of population age between 35-44.
	age8	Percentage of population age between 60-64.
	age9	Percentage of population age between 65-74.
	age10	Percentage of population age between 75-84.
	age11	Percentage of population age 85 and above.
Education	edu1	Percentage of population age 25+ with less than high school degree.
	edu3	Percentage of population age 25+ with some college or Associate degree.

Table 5.5. Variables Included in Factor Analysis (Cont.)

Socioeconomic & Demographic Variables	Symbol	Definition
Education	edu4	Percentage of population age 25+ with Bachelor degree.
	edu5	Percentage of population age 25+ with Graduate or professional degree.
Gender	male	Percentage of male.
Race	black	Percentage of Black population.
	asian	Percentage of Asian population.
	hispanic	Percentage of Hispanic population.
Group Quarters	gq	Percentage of population living in group quarters.
Household Size	hhsize	Average population per household.
Total Population	pop	2012 Estimated Total Population
Population Density	density	Average population per square mile.
Mobility	auto	Average automobiles per household.
Population Growth	popgrowth	Average population growth rate from 2010 to 2012.
Unemployment Rate	unempl	Unemployment rate in the area.

Table 5.6. Results of KMO Test for Sampling Adequacy

Variable	KMO
hh_mc	0.5947
hh_nf	0.7478
income1	0.755
income2	0.7625
income4	0.8354
income5	0.9023
income6	0.9512
income7	0.896
mdincome	0.8429
age1	0.8462
age2	0.782
age3	0.7001
age4	0.7035
age5	0.8153
age6	0.6411
age8	0.8286
age9	0.857
age10	0.8572
age11	0.8528
edu1	0.8117
edu3	0.5626
edu4	0.9014
edu5	0.9031
male	0.5652
black	0.5706
asian	0.9124
hispanic	0.6672
gq	0.6449
hhsize	0.8782
pop	0.8401
density	0.8529
auto	0.7917
popgrowth	0.8988
unempl	0.7758
Overall	**0.8098**

Table 5.6 suggests that the overall KMO statistic for all variables is 0.8098, which indicates the sampling is adequate and suitable for factor analysis (Kaiser 1974). Then, I use Stata to conduct the factor analysis with the extraction method of principle factors. After carefully inspecting the eigenvalues of all factors and reviewing the scree plot, five factors are retained in this analysis. The results of this five-factor model with a varimax (*i.e.*, orthogonal) rotation are shown in Table 5.7. Factor loadings with an absolute value greater than 0.3000 are highlighted in bold.

As indicated in Table 5.7, these five factors show distinct consumer characteristics drawn from the sample. Interpretation of these five factors – five consumer groups is provided below.

- Consumer Group 1 (Consumer1): highly educated (at least Bachelor's degree), wealthy (at least $75,000 annual household income) young professionals in their age of 35 to 44, living in highly populated and dense areas where population is growing. For simplicity, I label consumer group 1 as "rich young professionals".
- Consumer Group 2 (Consumer2): an average large full-nest household with children and parents under age 44. For simplicity, I label consumer group 2 as "large full-nest households".
- Consumer Group 3 (Consumer3): typical college students age from 15 to 24 with Bachelor's or above degree, living primarily in university dorms. For simplicity, I label consumer group 3 as "college students".
- Consumer Group 4 (Consumer4): typical middle-class family, with annual household income from $75,000 to 99,999, living in less populated and less dense areas with low unemployment rate, where autos are needed as means of transportation. For simplicity, I label consumer group 4 as "middle-class families".

- Consumer Group 5 (Consumer5): large households characterized with children under age 14, less than high school degree, living in areas with high unemployment rate. For simplicity, I label consumer group 5 as “large families with low education”.

Table 5.7. Rotated Factor Loadings and Unique Variances

Variable	Consumer1	Consumer2	Consumer3	Consumer4	Consumer5	Uniqueness
hh_mc	0.0373	-0.0881	-0.2986	**0.7778**	0.2364	0.2408
hh_nf	0.1535	-0.249	**0.5505**	**-0.3541**	**-0.5953**	0.1316
income1	**-0.7825**	-0.006	0.1177	**-0.5005**	0.1902	0.0871
income2	**-0.7727**	-0.2391	-0.092	0.2007	0.005	0.297
income4	**0.6202**	0.0757	-0.0645	**0.5125**	**-0.302**	0.2516
income5	**0.9059**	0.1517	-0.0249	0.176	-0.0985	0.115
income6	**0.9319**	0.1301	-0.0032	-0.0589	0.0584	0.1078
income7	**0.8922**	0.035	-0.0016	-0.1643	0.0899	0.1676
mdincome	**0.9313**	0.0572	-0.0727	0.2941	-0.0971	0.0282
age1	-0.0949	**0.7869**	**-0.3771**	0.1202	0.2531	0.1511
age2	-0.0918	**0.6653**	**-0.5677**	0.1455	**0.3321**	0.0952
age3	-0.0689	**0.529**	**0.6212**	-0.1002	0.0282	0.3187
age4	0.0152	**0.4328**	**0.7884**	-0.1452	-0.1955	0.1316
age5	0.2905	**0.7947**	0.0907	-0.019	0.0008	0.2756
age6	**0.3467**	0.2949	**-0.3782**	-0.0818	0.2825	0.5634
age8	-0.1057	**-0.8595**	-0.1916	-0.0083	-0.1289	0.1967
age9	-0.2376	**-0.8842**	-0.1505	0.0247	0.0343	0.1373
age10	-0.2032	**-0.8998**	-0.1114	0.0659	-0.0341	0.1312
age11	-0.0503	**-0.7634**	-0.0437	0.1035	-0.1771	0.3706
edu1	**-0.4394**	0.1926	-0.1823	-0.2326	**0.738**	0.1378
edu3	0.0341	0.0403	-0.1074	**0.3481**	**-0.3241**	0.7595
edu4	**0.7428**	0.1472	0.2066	-0.0579	**-0.3841**	0.233
edu5	**0.6807**	0.0674	**0.411**	-0.2316	-0.2849	0.2284
male	0.0128	0.0455	**0.4573**	**0.4885**	0.2414	0.4916
black	-0.1082	0.2388	-0.1313	**-0.6595**	-0.0109	0.479
asian	**0.5759**	0.1221	0.0978	-0.0841	0.1343	0.6187
hispanic	0.0889	0.2676	-0.0463	0.0954	**0.6265**	0.5168
gq	-0.0744	0.0198	**0.8073**	0.0691	0.1038	0.3268
hhsize	-0.0075	**0.5834**	0.1096	0.0829	**0.6933**	0.1601
pop	**0.5287**	0.0592	-0.1081	**-0.3155**	0.1867	0.571
density	**0.6567**	-0.0317	-0.098	**-0.3926**	0.2029	0.3628
auto	-0.0272	0.1881	-0.1269	**0.6651**	-0.129	0.4888
popgrowth	**0.33**	0.111	-0.0308	-0.0528	-0.0936	0.8663
unempl	-0.2823	-0.0983	-0.1428	**-0.3484**	**0.3149**	0.6698

5.1.4 Data Processing – Transformation of the Dependent Variables

After a visual inspection of the data and a plot of the residuals along the fitted values of the dependent variables ($\hat{y}$), I notice there exists slight heteroskedasticity in the data. Therefore, I decide to transform the dependent variables to mitigate the effects of heteroskedasticity. After adopting the Box-Cox transformation method which provides the optimal transformation function for the data, the original dependent variables are transformed with a *ln* function such that the new dependent variables equal the natural logarithm of the original dependent variables.

Table 5.8 provides details about these transformed dependent variables, the five consumer groups, and all other variables used in the succeeding data analysis. Table 5.9 presents the correlation matrix. This table is created using *asdoc*, a Stata program written by Shah (2018).

Table 5.8. Descriptive Statistics of Variables Used in Regressions

Variable	Number of Observations	Mean	Std. Dev.	Min	Max
Dependent Variables					
ln_car	558	8.548	0.4123	7.269	9.889
ln_fur	613	6.272	0.5820	3.356	8.239
ln_elec	607	6.269	0.6742	3.362	8.172
ln_food	505	8.287	0.3815	6.808	10.248
ln_gas	731	8.623	0.3806	7.103	10.278
ln_cloth	663	6.928	0.7764	3.287	9.280
ln_dept	213	6.405	0.3108	5.721	7.961
ln_ddept	225	7.034	0.4437	5.657	8.819
ln_super	57	8.335	0.4866	7.008	9.411
Consumer Groups					
Consumer1	735	0	0.994	-1.679	5.520
Consumer2	735	0	0.991	-6.722	3.527
Consumer3	735	0	0.980	-1.447	6.121
Consumer4	735	0	0.977	-4.663	3.247
Consumer5	735	0	0.968	-2.139	6.191
Marketing mix Variables					
labor_car	558	41.716	14.149	7.188	86.706
labor_fur	613	7.407	2.489	1.9	21.648
labor_elec	607	7.642	2.617	2.080	16.816
labor_food	505	41.272	14.236	12.605	123.756
labor_gas	731	8.423	2.381	2.614	21.168
labor_cloth	663	9.065	2.629	2.559	16.475
labor_dept	213	123.201	40.824	51.819	327.805
labor_ddept	225	115.570	29.933	56.589	206.641
labor_super	57	247.009	45.655	136.091	321.238
wage_car	558	1.822	0.220	1.041	2.503
wage_fur	613	1.108	0.204	0.517	2.312
wage_elec	607	0.958	0.165	0.514	2.114
wage_food	505	0.877	0.128	0.523	1.338
wage_gas	731	0.742	0.088	0.494	1.099

Table 5.8. Descriptive Statistics of Variables Used in Regressions (Cont.)

Variable	Number of Observations	Mean	Std. Dev.	Min	Max
Marketing mix Variables					
wage_cloth	663	0.644	0.090	0.390	1.352
wage_dept	213	0.748	0.077	0.378	0.926
wage_ddept	225	0.717	0.076	0.534	1.046
wage_super	57	1.030	0.099	0.802	1.213
Competition Variables					
stores_fur (per 1000 HH)	734	0.466	0.216	0.095	2.755
stores_elec (per 1000 HH)	734	0.424	0.153	0.069	1.884
stores_food (per 1000 HH)	735	0.491	0.220	0.088	3.878
stores_cloth (per 1000 HH)	735	1.153	0.727	0.043	8.442
stores_dept (per 1000 HH)	541	0.052	0.027	0.010	0.279
stores_ddept (per 1000 HH)	584	0.058	0.031	0.010	0.279
stores_super (per 1000 HH)	710	0.060	0.026	0.010	0.222
Consumer Out-shopping Variables					
pop_large100 (million)	735	1.573	2.695	0	26.643
pop_large200 (million)	735	6.227	7.031	0	39.548
Control Variables					
msa_type	735	0.482	0.500	0	1
sales_total ($000)	735	35.081	11.205	13.747	197.638

Table 5.9. Correlation Matrix

Variables	ln_car	ln_fur	ln_elec	ln_food	ln_gas	ln_cloth	ln_dept	ln_ddept
ln_car	1							
ln_fur	0.520***	1						
ln_elec	0.482***	0.651***	1					
ln_food	0.275***	0.389***	0.347***	1				
ln_gas	0.173***	0.016	-0.025	-0.071	1			
ln_cloth	0.444***	0.669***	0.631***	0.445***	0.014	1		
ln_dept	0.346***	0.353***	0.367***	0.082	0.347***	0.460***	1	
ln_ddept	-0.055	-0.011	0.128*	0.375***	0.071	0.019	0.181**	1
ln_super	-0.154	-0.275**	-0.320**	-0.240*	0.09	-0.443***	-0.055	-0.508***
Consumer1	0.298***	0.444***	0.489***	0.509***	-0.296***	0.434***	-0.124*	0.119*
Consumer2	0.148***	0.120***	0.223***	-0.095**	0.154***	0.209***	0.078	-0.024
Consumer3	-0.027	-0.031	0.006	0.05	-0.034	0.004	0.038	0.286***
Consumer4	-0.087**	-0.107***	-0.090**	-0.011	0.067*	-0.245***	-0.08	0.229***
Consumer5	-0.073*	-0.114***	-0.162***	0.078*	-0.006	0.085**	0.187***	0.102
labor_car	0.499***	0.483***	0.535***	0.211***	-0.166***	0.485***	0.002	-0.198***
labor_fur	0.304***	0.549***	0.387***	0.088*	-0.026	0.403***	0.018	-0.016
labor_elec	0.309***	0.455***	0.729***	0.129***	-0.105***	0.449***	0.164**	0.075
labor_food	0.259***	0.316***	0.324***	0.365***	-0.041	0.182***	-0.07	0.01
labor_gas	0.042	0.024	0.139***	0.039	0.291***	-0.043	0.033	0.264***
labor_cloth	0.301***	0.459***	0.566***	0.258***	-0.111***	0.704***	0.038	0.007
labor_dept	0.023	0.297***	0.364***	0.274***	-0.334***	0.382***	0.051	0.184**
labor_ddept	-0.157**	0.150**	0.218***	0.414***	-0.428***	0.346***	0.098	0.441***
labor_super	0.576***	0.404***	0.434***	-0.201	0.366***	0.407***	0.156	-0.384**
wage_car	0.208***	0.238***	0.303***	0.003	-0.054	0.340***	0.206***	-0.203***
wage_fur	0.006	0.149***	0.048	-0.090*	0.076*	-0.067*	0.057	-0.138**
wage_elec	-0.077*	-0.001	0.189***	0.003	0.036	0.003	-0.130*	0.006
wage_food	-0.173***	0.038	0.103**	0.219***	-0.187***	0.177***	-0.077	0.078
wage_gas	-0.086**	0.032	0.062	0.117***	0.210***	0.212***	0.269***	0.029
wage_cloth	-0.156***	-0.195***	-0.146***	-0.079*	0.041	-0.140***	0.035	0.021
wage_dept	0.069	0.218***	0.211***	0.028	0.012	0.382***	0.277***	-0.106
wage_ddept	-0.115*	0.035	0.049	0.225***	-0.071	0.176***	0.170**	0.433***
wage_super	-0.446***	-0.317**	-0.379***	0.209	-0.302**	-0.233*	-0.09	0.235
stores_fur	0.289***	0.554***	0.261***	0.412***	0.155***	0.426***	0.358***	0.172***
stores_elec	0.302***	0.276***	0.418***	0.334***	0.270***	0.235***	0.282***	0.277***
stores_food	0.077*	0.129***	0.054	0.501***	0.122***	0.162***	0.234***	0.261***
stores_cloth	0.292***	0.473***	0.370***	0.434***	0.155***	0.739***	0.481***	0.148**
stores_dept	0.093**	-0.059	-0.111**	-0.189***	0.318***	-0.131***	0.572***	0.279***
stores_ddept	0.05	0.033	0.065	0.199***	0.243***	-0.076*	0.226***	0.688***
stores_super	0.048	0.028	-0.033	-0.095**	0.181***	-0.052	0.200***	-0.115*
pop_large100	-0.131***	-0.181***	-0.201***	-0.022	-0.076**	-0.174***	-0.127*	0.208***
pop_large200	-0.120***	-0.198***	-0.152***	0.005	-0.034	-0.180***	0.074	0.263***
msa_type	0.277***	0.467***	0.593***	0.218***	-0.194***	0.496***	-0.176**	-0.285***
sales_total	0.560***	0.524***	0.473***	0.460***	0.445***	0.522***	0.486***	0.205***

Table 5.9. Correlation Matrix (Cont.)

Variables	ln_super	Consumer1	Consumer2	Consumer3	Consumer4	Consumer5	labor_car	labor_fur
ln_super	1							
Consumer1	-0.522***	1						
Consumer2	0.315**	0.005	1					
Consumer3	-0.263**	0	0.006	1				
Consumer4	0.674***	0.011	-0.001	-0.018	1			
Consumer5	-0.370***	-0.012	0.017	-0.026	-0.005	1		
labor_car	-0.077	0.468***	0.194***	-0.047	-0.145***	-0.035	1	
labor_fur	-0.258*	0.244***	0.237***	0.003	-0.082**	0.028	0.387***	1
labor_elec	-0.145	0.363***	0.250***	-0.007	-0.092**	-0.092**	0.460***	0.369***
labor_food	-0.079	0.320***	-0.049	0.122***	0.147***	-0.258***	0.239***	0.196***
labor_gas	0.146	0.046	0.034	0.141***	0.369***	-0.185***	-0.021	0.106***
labor_cloth	-0.518***	0.461***	0.225***	0.102***	-0.195***	0.103***	0.482***	0.450***
labor_dept	-0.559***	0.528***	0.128*	-0.161**	-0.297***	0.388***	0.309***	0.295***
labor_ddept	-0.603***	0.537***	-0.104	0.127*	-0.357***	0.480***	0.083	0.108
labor_super	0.029	0.221*	0.504***	-0.205	0.061	0.007	0.447***	0.320**
wage_car	-0.224*	0.208***	0.164***	-0.016	-0.162***	0.022	0.234***	0.172***
wage_fur	0.192	-0.097**	0.067*	-0.041	0.023	-0.067*	-0.027	0.064
wage_elec	-0.408***	0.012	-0.067*	0.021	-0.031	-0.190***	-0.118***	-0.003
wage_food	0.372***	0.079*	-0.039	0.04	-0.005	0.223***	0.135***	-0.026
wage_gas	-0.144	0.092**	0.019	0.054	-0.058	0.141***	0.047	0.115***
wage_cloth	-0.056	-0.084**	-0.069*	-0.015	0.01	-0.009	-0.177***	-0.174***
wage_dept	-0.18	0.039	0.024	0.087	-0.137**	0.266***	0.121*	0.067
wage_ddept	-0.045	-0.054	-0.062	0.362***	0.016	0.067	-0.180***	-0.04
wage_super	0.163	-0.241*	-0.278**	0.1	0.330**	0.174	-0.417***	-0.291**
stores_fur	0.092	0.164***	-0.069*	0.02	-0.024	-0.171***	0.084**	-0.091**
stores_elec	-0.059	0.094**	0.037	0.044	0.130***	-0.155***	0.039	-0.017
stores_food	-0.328**	0.128***	-0.121***	-0.016	-0.108***	0.139***	-0.119***	-0.058
stores_cloth	-0.330**	0.240***	0.073**	-0.002	-0.134***	0.008	0.210***	0.162***
stores_dept	0.347***	-0.384***	-0.109**	0.087**	0.123***	-0.146***	-0.248***	-0.107**
stores_ddept	-0.031	-0.077*	-0.080*	0.053	0.320***	-0.167***	-0.226***	-0.076*
stores_super	0.790***	-0.295***	0.014	0.013	0.147***	-0.095**	-0.095**	-0.048
pop_large100	-0.271**	0.086**	-0.146***	-0.059	0.089**	0.186***	-0.191***	-0.137***
pop_large200	-0.491***	-0.007	-0.244***	0.085**	0.059	0.098***	-0.253***	-0.160***
msa_type	-0.276**	0.483***	0.272***	-0.014	-0.178***	-0.073**	0.538***	0.354***
sales_total	-0.018	0.255***	0.153***	0.003	0	-0.090**	0.212***	0.195***

Table 5.9. Correlation Matrix (Cont.)

Variables	labor_elec	labor_food	labor_gas	labor_cloth	labor_dept	labor_ddept	labor_super	wage_car
labor_elec	1							
labor_food	0.257***	1						
labor_gas	0.064	0.267***	1					
labor_cloth	0.500***	0.237***	0.038	1				
labor_dept	0.112	-0.182**	-0.126*	0.527***	1			
labor_ddept	0.081	-0.095	-0.156**	0.433***	0.610***	1		
labor_super	0.399***	0.184	-0.1	0.328**	-0.091	-0.388***	1	
wage_car	0.253***	0.185***	-0.077*	0.304***	0.172**	0.068	0.303**	1
wage_fur	-0.035	0.013	0.045	0	-0.035	-0.237***	0.185	-0.045
wage_elec	-0.018	0.081*	0.121***	0.039	-0.026	-0.032	-0.096	-0.058
wage_food	0.053	-0.215***	-0.116***	0.094**	0.186**	0.06	-0.137	-0.024
wage_gas	0.092**	0.110**	0.019	0.161***	-0.011	0.186***	-0.125	0.114***
wage_cloth	-0.143***	-0.052	0.082**	-0.281***	0.200***	0.115*	-0.107	-0.072*
wage_dept	0.184***	0.061	-0.027	0.201***	0.025	0.180**	0.188	0.298***
wage_ddept	0.06	0.052	0.02	0.029	-0.046	0.292***	-0.522***	-0.052
wage_super	-0.305**	-0.1	0.088	-0.288**	-0.082	0.250*	-0.453***	-0.201
stores_fur	0.051	0.107**	-0.028	0.056	-0.118*	-0.029	-0.062	0.051
stores_elec	-0.069*	0.085*	0.215***	0.03	0.008	-0.033	-0.006	-0.022
stores_food	-0.108***	-0.309***	-0.071*	-0.048	0.314***	0.376***	-0.231*	-0.168***
stores_cloth	0.147***	0.074*	0.019	0.243***	0.037	0.084	0.310**	0.228***
stores_dept	-0.120***	-0.092*	0.162***	-0.344***	-0.569***	-0.378***	0.002	-0.100**
stores_ddept	-0.082*	0.106**	0.275***	-0.190***	-0.219***	-0.211***	0.03	-0.213***
stores_super	-0.125***	-0.187***	0.081**	-0.203***	-0.362***	-0.509***	-0.461***	-0.144***
pop_large100	-0.202***	-0.06	-0.046	-0.100***	0.061	0.203***	-0.052	-0.049
pop_large200	-0.182***	-0.012	0.013	-0.109***	-0.164**	0.185***	-0.083	-0.153***
msa_type	0.553***	0.193***	-0.032	0.564***	0.213***	0.232***	0.421***	0.294***
sales_total	0.222***	0.161***	0.176***	0.264***	-0.002	-0.088	0.361***	0.053

Table 5.9. Correlation Matrix (Cont.)

Variables	wage_fur	wage_elec	wage_food	wage_gas	wage_cloth	wage_dept	wage_ddept	wage_super
wage_fur	1							
wage_elec	0.119***	1						
wage_food	-0.110**	-0.127***	1					
wage_gas	-0.080**	0.014	0.045	1				
wage_cloth	0.121***	0.101**	-0.174***	0.02	1			
wage_dept	-0.134*	-0.120*	0.053	0.334***	0.133*	1		
wage_ddept	-0.147**	0.067	0.086	0.164**	0.062	0.194**	1	
wage_super	0.002	0.052	0.184	0.106	0.055	-0.035	0.473***	1
stores_fur	-0.102**	0.011	-0.034	0.044	0.021	0.042	0.197***	0.132
stores_elec	0.066*	0.100**	-0.039	-0.066*	-0.002	-0.068	0.035	-0.094
stores_food	-0.029	0.071*	0.01	0.024	0.102***	0.022	0.187***	0.213
stores_cloth	-0.105***	0.032	0.143***	0.222***	-0.008	0.222***	0.236***	-0.131
stores_dept	0.06	-0.047	-0.071	-0.03	0.112**	-0.08	0.170**	0.127
stores_ddept	0.105**	0.161***	-0.047	-0.097**	0.137***	-0.168**	0.217***	0.07
stores_super	-0.002	-0.05	0.104**	-0.067*	0.023	-0.099	0.008	0.248*
pop_large100	0.047	0.044	0.004	0.064*	0.055	-0.078	0.096	0.217*
pop_large200	0	0.093**	-0.108**	0.036	0.079**	-0.05	0.217***	0.109
msa_type	-0.022	0.007	0.159***	0.109***	-0.291***	0.175**	-0.027	-0.184
sales_total	0.014	-0.025	-0.094**	0.026	-0.073*	0.063	0.005	-0.285**

Table 5.9. Correlation Matrix (Cont.)

Variables	stores_fur	stores_elec	stores_food	stores_gas	stores_dept	stores_ddept	stores_super	pop_large100
stores_fur	1							
stores_elec	0.466***	1						
stores_food	0.404***	0.332***	1					
stores_cloth	0.647***	0.380***	0.335***	1				
stores_dept	0.167***	0.251***	0.044	0.116***	1			
stores_ddept	0.248***	0.441***	0.281***	0.175***	0.450***	1		
stores_super	0.234***	0.289***	0.164***	0.129***	0.388***	0.213***	1	
pop_large100	-0.133***	-0.117***	0.033	-0.117***	-0.05	0.004	-0.113***	1
pop_large200	-0.049	-0.053	0.104***	-0.087**	0.153***	0.113***	-0.096**	0.309***
msa_type	0.093**	-0.008	-0.084**	0.194***	-0.342***	-0.294***	-0.161***	-0.158***
sales_total	0.574***	0.576***	0.480***	0.571***	0.191***	0.327***	0.272***	-0.174***

Table 5.9. Correlation Matrix (Cont.)

Variables	pop_large200	msa_type	sales_total
pop_large200	1		
msa_type	-0.104***	1	
sales_total	-0.080**	0.199***	1

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

5.2 Research Methods

In this section, I discuss the research methods and econometric models used for the analysis. Recall the research questions of this dissertation: I plan to examine

1. how intertype and intratype competition influence retail expenditures per household.
2. what key variables drive retail expenditures per household across different retail lines.
3. whether those key variables differ in metropolitan areas vs. micropolitan areas.
4. whether large areas draw retail sales (*i.e.*, expenditures) from nearby small areas.

To address the abovementioned research questions, I use the method of OLS regression. Specially, I run seven different models to address the research questions. Details about these seven models are provided in Table 5.10.

Table 5.10. Model Specification

Models / Variables	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Consumer Groups			✓	✓	✓		
Marketing mix		✓	✓	✓	✓		✓
Competition[20]	✓	✓	✓	✓	✓		
Control	✓	✓	✓	✓	✓		
Consumer Out-shopping						✓	✓
Sample[21]	Full	Full	Full	Metropolitan Areas Only	Micropolitan Areas Only	Full	Full
Number of Retail Subsectors/Industries Examined	Seven	Seven	Nine	Nine	Six	Seven	Seven
Research Question Addressed	#1	#1	#2	#3	#3	#4	#4

Below I discuss the model specification and how each model is applied in the nine retail subsectors/industries.

- Model 1

In model 1, I use all the observations in the sample to run the below regression analysis for seven of the nine retail subsectors/industries. For NAICS 44111 New Car Dealers and

[20] For Model 3, Model 4, and Model 5, competition variable(s) is added when there exists competition in the focal retail subsector/industry.

[21] For NAICS 452111 Department Stores (except Discount Department Stores), NAICS 452112 Discount Department Stores, and NAICS 45291 Warehouse Clubs and Supercenters, the number of observations is very small in micropolitan areas. Therefore, for these industries, I only use the subsample of metropolitan areas for Model 1, Model 2, and Model 3. Also, I am not able to run Model 5 for these industries.

NAICS 447 Gasoline Stations, they do not compete with other retail subsectors/industries, therefore, this model is not applicable to these two retail lines. For the three retail industries under subsector NAICS 452 General Merchandise Stores, since there are only 13 observations of micropolitan areas in NAICS 452111 Department Stores (except Discount Department Stores), 20 in NAICS 452112 Discount Department Stores, and 4 in NAICS 45291 Warehouse Clubs and Supercenters, given the relatively small size of micropolitan areas in these three industries, I decide to exclude them in the analysis and only use observations of metropolitan areas for these industries. The econometric model is shown below.

$ln\ (industry_{1i}) = \beta_{10i} + \beta_{1i}Competition_i + \sum_1^2 \gamma_{1j}\ Control_j + \varepsilon_{1i}$ (i = 1, 2, …, 7; j = 1, 2)

where Competition refers to either intertype or intratype the focal retail subsector/industry face; and Control refers to control variables (MSA type and total retail expenditure per household). For the three industries in subsector NAICS 452 General Merchandise Stores, only total retail expenditure per household is added as the control variable, since observations on micropolitan areas are not included and thus the control variable – MSA type is not applicable.

- Model 2

Model 2 is built upon model 1, but with the addition of marketing mix variables. I use the same retail subsectors/industries which are examined in model 1. The addition of marketing mix variables demonstrates how marketing mix effort mitigates the harm of competition to the focal subsector/industry. Model 1 and Model 2 both address research question #1: how intertype and intratype competition influence retail expenditures per household. The econometric model is shown below.

$ln\ (industry_{2i}) = \beta_{20i} + \beta_{2i}Competition_i + \sum_1^2 \gamma_{2j}\ Control_j + \sum_1^2 \delta_{2ki}\ MarketingMix_{ki} +$

ε_{2i} (i = 1, 2, …, 7; j = 1, 2; k = 1, 2)

where MarketingMix refers to the two marketing mix variables (quantity and quality of service).

Note that I only add competition variables, marketing mix variables, and control variables in this model, excluding the consumer groups, so that competition is allowed to vary across different consumer segments. Same as in Model 1, for the three industries in NAICS 452 General Merchandise Stores, only observations on metropolitan areas are used for the analysis and only one control variable – total retail expenditure per household is added in the model.

- Model 3

Model 3 is built upon model 2, but with the addition of the five consumer groups. I run this model with all the nine retail subsectors/industries. The purpose of this model is to examine all the variables and address research question #2: what key variables drive retail expenditures per household across different retail lines. The econometric model is shown below.

$ln\ (industry_{3i}) = \beta_{30i} + (\beta_{3i}Competition_i) + \sum_1^2 \gamma_{3j}\ Control_j + \sum_1^2 \delta_{3ki}\ MarketingMix_{ki}$

$+ \sum_1^5 \theta_{3li}\ Consumer_l + \varepsilon_{3i}$ (i = 1, 2, …, 9; j = 1, 2; k = 1, 2; l = 1, 2, …, 5)

where Consumer refers to the five consumer groups.

Note that same as in Model 1 and Model 2, for the three industries in NAICS 452 General Merchandise Stores, only observations on metropolitan areas are used for the analysis and only one control variable – total retail expenditure per household is added in the model.

- Model 4

Model 4 includes the same variables of model 3 but is only examined with observations of metropolitan areas. The econometric model is shown below.

$ln\ (industry_{4i})= \beta_{40i}$ + ($\beta_{4i}Competition_i$) + $\sum_1^2 \gamma_{4j}\ Control_j$ + $\sum_1^2 \delta_{4ki}\ MarketingMix_{ki}$ + $\sum_1^5 \theta_{4li}\ Consumer_l$ + ε_{4i} (i = 1, 2, …, 9; j = 1, 2; k = 1, 2; l = 1, 2, …, 5)

where Control refers to only one control variable – total retail expenditure per household.

- Model 5

Model 5 also includes the same variables of model 3 but is only examined with observations of micropolitan areas. I run this model for all nine retail subsectors/industries, except for NAICS 452111 Department Stores (except Discount Department Stores), NAICS 452112 Discount Department Stores, and NAICS 45291 Warehouse Clubs and Supercenters, because their sample sizes in micropolitan areas are too small to run regression analysis. Model 4 and model 5 together address research question #3: whether key variables differ in metropolitan areas vs. micropolitan areas. The econometric model is shown below.

$ln\ (industry_{5i})= \beta_{50i}$ + ($\beta_{5i}Competition_i$) + $\sum_1^2 \gamma_{5j}\ Control_j$ + $\sum_1^2 \delta_{5ki}\ MarketingMix_{ki}$ + $\sum_1^5 \theta_{5li}\ Consumer_l$ + ε_{5i} (i = 1, 2, …, 9; j = 1, 2; k = 1, 2; l = 1, 2, …, 5)

where Control refers to only one control variable – total retail expenditure per household.

- Model 6

In model 6, I examine the effects of consumer out-shopping variables only, and I run this model with all the nine retail subsectors/industries, except NAICS 44511 Supermarkets and Other Grocery (except Convenience) Stores and NAICS 447 Gasolines Stations because consumers primarily shop in the focal area and barely go out-of-town shopping for these needs. The econometric model is shown below.

$ln\ (industry_{6i})=\beta_{60i}+\sum_{1}^{4}\lambda_{6mi}\ OutShopping_{m}+\varepsilon_{6i}$ (i = 1, 2, …, 7; m = 1, 2, 3, & 4)

where OutShopping refers to the two consumer out-shopping variables (*i.e.*, population from nearby large areas).

- Model 7

Model 7 is built upon model 6 but with the addition of marketing mix variables. I use the same retail subsectors/industries which are examined in model 6. The addition of marketing mix variables demonstrates 1) how marketing mix effort mitigates the harm of competition from nearby large areas, and 2) enhance the attractiveness of the focal area to nearby small areas. Model 6 and Model 7 both address research question #4: whether large areas draw retail sales (*i.e.*, expenditures) from nearby small areas. The econometric model is shown below.

$ln\ (industry_{7i})=\beta_{70i}+\sum_{1}^{4}\lambda_{7mi}\ OutShopping_{m}+\varepsilon_{7i}$ (i = 1, 2, …, 7; m = 1, 2, 3, & 4)

Other than OLS regression, I do a mediation analysis using the PROCESS macro for SPSS by Hayes (2017) on Model 3 for all the nine retail subsectors/industries. Results of OLS regressions and mediation analyses are provided in Chapter 6.

Chapter 6 Results

This chapter presents the empirical results of data analysis. Results of the seven models, where applicable, are shown for each of the nine retail subsectors/industries in the following sections. Furthermore, a mediation analysis of the direct effects and indirect effects is conducted on Model 3 for all the nine retail subsectors/industries.

6.1 NAICS 44111 New Car Dealers

The results of five OLS regressions are shown in Table 6.1. Since New Car Dealers does not compete with any of the remaining subsector or industry included in this dissertation, Model 1 and Model 2 are not applicable for this industry. Only results of Model 3, Model 4, Model 5, Model 6, and Model 7 are available.

Results of Model 3, Model 4, and Model 5 indicate that both quantity and quality of service significantly increase household expenditure at New Car Dealers, which is true for both metropolitan and micropolitan areas. Among the five consumer groups, only consumer group 5 – large families with low education, has a significant and negative effect on new car expenditure in metropolitan areas. This shows that on average this consumer group spends less on new cars than the baseline consumers – female single-head.

Results of Model 6 and Model 7 indicate that large areas indeed take sales away from small areas, especially when they are located within 100 miles. Furthermore, consumers located in small areas within 100 to 200 miles are willing to travel and shop for new cars in large areas. Model 6 and Model 7 show that consumers do shop out-of-town for new cars, however, this behavior is influenced by the marketing effort made by local new car dealers. When local new car dealers provide adequate and quality service, consumers may choose not to shop out-of-town.

This further demonstrates that marketing mix effort in the focal area can mitigate the negative impact from out-of-town competitors.

Table 6.2 presents the results of a mediation analysis on Model 3. This table shows that consumer group 1 and consumer group 2 have significant and positive indirect effects on new car expenditure through quantity of service. In other words, quantity of service is important and influential for rich young professionals and large full-nest households. Moreover, consumer group 4 – middle-class families do not need high quality service when buying new cars. They may weigh more on the value of the car.

Table 6.1. NAICS 44111 New Car Dealers – OLS Results

Variables	**Model 3**	**Model 4 (Metro)**	**Model 5 (Micro)**	**Model 6**	**Model 7**
Consumer1	0.009 (0.015)	0.028 (0.015)	0.031 (0.047)		
Consumer2	0.007 (0.014)	-0.015 (0.015)	0.050 (0.028)		
Consumer3	-0.005 (0.014)	-0.019 (0.016)	-0.003 (0.024)		
Consumer4	-0.008 (0.014)	0.010 (0.017)	-0.022 (0.027)		
Consumer5	-0.016 (0.014)	-0.040** (0.014)	0.033 (0.030)		
labor_car	0.011*** (0.001)	0.008*** (0.001)	0.015*** (0.002)		0.014*** (0.001)
wage_car	0.175** (0.062)	0.168* (0.081)	0.204* (0.096)		0.186** (0.071)
msa_type	-0.021 (0.033)				
sales_total	0.017*** (0.001)	0.018*** (0.002)	0.016*** (0.002)		
pop_large100				-0.016* (0.007)	-0.007 (0.006)
pop_large200				-0.005* (0.002)	0.002 (0.002)
_cons	7.177*** (0.122)	7.290*** (0.164)	7.046*** (0.194)	8.602*** (0.023)	7.632*** (0.132)
N	558	320	238	558	558
Adj. R-sq	0.467	0.423	0.451	0.021	0.254

Standard errors in parentheses

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

Model 1 and Model 2 are not applicable since this industry does not compete with other subsectors/industries included in this dissertation.

Table 6.2. NAICS 44111 New Car Dealers – Mediation Analysis

	Indirect Effects				***Direct Effect***	***Total Effect***
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_car)		Quality of Service (wage_car)		Household Expenditure (ln_car)	Household Expenditure (ln_car)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.042*** (0.009)	(0.026, 0.059)	0.004 (0.002)	(0.001, 0.010)	0.009 (0.015)	0.055*** (0.015)
Consumer2	0.014* (0.007)	(-0.000, 0.029)	0.005 (0.003)	(0.001, 0.011)	0.007 (0.014)	0.026 (0.015)
Consumer3	-0.006 (0.005)	(-0.017, 0.004)	-0.000 (0.002)	(-0.004, 0.003)	-0.005 (0.014)	-0.012 (0.015)
Consumer4	-0.010 (0.005)	(-0.022, -0.000)	-0.005* (0.002)	(-0.010, -0.001)	-0.008 (0.014)	-0.023 (0.015)
Consumer5	-0.008 (0.006)	(-0.021, 0.004)	0.000 (0.002)	(-0.004, 0.005)	-0.016 (0.014)	-0.024 (0.015)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

6.2 NAICS 442 Furniture and Home Furnishing Stores

The results of seven OLS regressions are presented in Table 6.3. Results of Model 1 and Model 2 indicate that there exists intertype competition between Furniture and Home Furnishing Stores and Warehouse Clubs and Supercenters. As the number of Warehouse Clubs and Supercenters increases in the focal area, household expenditure at Furniture and Home Furnishing Stores will decrease. However, service provided at local furniture stores can mitigate the negative effect from competition and increase household expenditure at local furniture stores.

Results of Model 3, Model 4, and Model 5 indicate that household expenditure at furniture stores is significantly higher in metropolitan areas. In general, consumer group 1 – rich young professionals spend more money on furniture than the baseline consumers – female single-head. On average, consumer group 2 – large full-nest households, consumer group 4 – middle-class families, and consumer group 5 – large families with low education spend less on furniture than baseline consumers.

Results of Model 6 and Model 7 indicate a very interesting phenomenon that consumers would go 100 to 200 miles to large areas to shop for furniture. Service provided at local furniture stores mitigates the harm from outside competition but is not sufficient. Since there exists a high degree of customization for furniture and home furnishing products, which is strongly related with one's personal lifestyle, consumers are willing to incur money and time costs to travel and shop out-of-town for this type of more personalized products.

Table 6.4 presents the results of a mediation analysis on Model 3. This table shows that although consumer group 2 – large full-nest households spend less at furniture stores, providing adequate service with this group of consumers can increase their expenditure. Furthermore, quality of service may not work well for consumer group 1 – rich young professionals at furniture stores.

Table 6.3. NAICS 442 Furniture and Home Furnishing Stores – OLS Results

Variables	Model 1	Model 2	Model 3	Model 4 (Metro)	Model 5 (Micro)	Model 6	Model 7
Consumer1			0.097*** (0.019)	0.091*** (0.017)	0.152* (0.061)		
Consumer2			-0.032* (0.016)	-0.024 (0.017)	-0.038 (0.031)		
Consumer3			-0.015 (0.016)	-0.021 (0.017)	-0.014 (0.027)		
Consumer4			-0.033* (0.016)	-0.035 (0.018)	-0.058 (0.033)		
Consumer5			-0.046** (0.016)	-0.040* (0.016)	-0.058 (0.032)		
labor_fur		0.083*** (0.007)	0.084*** (0.007)	0.064*** (0.009)	0.096*** (0.010)		0.121*** (0.008)
wage_fur		0.367*** (0.077)	0.406*** (0.075)	0.383*** (0.107)	0.436*** (0.110)		0.344*** (0.095)
stores_super	-1.687* (0.749)	-1.198 (0.655)	0.164 (0.692)	0.285 (0.816)	-0.138 (1.140)		
msa_type	0.464*** (0.037)	0.339*** (0.034)	0.257*** (0.038)				
sales_total	0.025*** (0.002)	0.022*** (0.002)	0.019*** (0.002)	0.024*** (0.002)	0.017*** (0.002)		
pop_large100						-0.028** (0.009)	-0.019* (0.007)
pop_large200						-0.013*** (0.003)	-0.007* (0.003)
_cons	5.239*** (0.065)	4.360*** (0.108)	4.356*** (0.107)	4.643*** (0.142)	4.366*** (0.161)	6.396*** (0.031)	5.072*** (0.121)
N	593	593	593	321	272	613	613
Adj. R-sq	0.437	0.570	0.597	0.548	0.439	0.052	0.331

Standard errors in parentheses

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

Table 6.4. NAICS 442 Furniture and Home Furnishing Stores – Mediation Analysis

	Indirect Effects				***Direct Effect***	***Total Effect***
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_fur)		Quality of Service (wage_fur)		Household Expenditure (ln_fur)	Household Expenditure (ln_fur)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.018 (0.009)	(0.001, 0.036)	-0.010* (0.005)	(-0.022, -0.002)	0.097*** (0.019)	0.105*** (0.022)
Consumer2	0.033*** (0.009)	(0.015, 0.052)	0.005 (0.005)	(-0.004, 0.016)	-0.032* (0.016)	0.006 (0.019)
Consumer3	0.002 (0.008)	(-0.013, 0.018)	-0.004 (0.004)	(-0.014, 0.003)	-0.015 (0.016)	-0.017 (0.018)
Consumer4	-0.009 (0.008)	(-0.026, 0.004)	0.002 (0.004)	(-0.005, 0.009)	-0.033* (0.016)	-0.040* (0.018)
Consumer5	0.008 (0.014)	(-0.019, 0.034)	-0.006 (0.005)	(-0.016, 0.002)	-0.046** (0.016)	-0.044* (0.018)
stores_super	-0.216 (0.318)	(-0.822, 0.420)	-0.226 (0.210)	(-0.711, 0.114)	0.164 (0.692)	-0.278 (0.800)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

6.3 NAICS 443 Electronics and Appliances Stores

The results of seven OLS regressions are presented in Table 6.5. Results of Model 1 and Model 2 indicate that Electronics and Appliances Stores faces intertype competition with Warehouse Clubs and Supercenters. As the number of Warehouse Clubs and Supercenters increases in the focal area, household expenditure at Electronics and Appliances Stores significantly decreases. However, service provided by Electronics and Appliances Stores mitigates the negative effect of intertype competition.

Results of Model 3, Model 4, and Model 5 indicate that household expenditure at electronics stores is significantly higher in metropolitan areas. In general, consumer group 1 – rich young professionals spend more money on electronics than the baseline consumers. Since this group of consumers live in large populated areas, primarily in metropolitan areas, this result is also true for the subsample of metropolitan areas. Furthermore, on average consumer group 4 – middle-class families living in metropolitan areas spend less on electronics; and consumer group 5 – large families with low education especially those living in micropolitan areas spend less on electronics, compared with baseline consumers.

Results of Model 6 and Model 7 indicate that consumers do out-of-town shopping for electronics and appliances. They are willing to travel 200 miles for this type of products. The rationale behind this consumer behavior is that electronics and appliances are high-price high-tech items, the utilities from buying these products may be much higher than the associated costs. Moreover, service provided by local stores facilitates consumer shopping locally and thus reduces their out-of-town shopping frequency and its negative effect on sales.

Table 6.6 presents the results of a mediation analysis on Model 3. This table shows that effects of consumer group 1 – rich young professionals and consumer group 2 – large full-nest households on household expenditure are mediated through quantity of service. In other words, these two types of consumers are very sensitive to the quantity of service at electronics and appliances stores. On the other hand, quantity and quality of service may not work well for consumer group 5 – large families with low education because they cannot afford, or they are not willing to pay for the extra cost of service at electronics and appliances stores.

Table 6.5. NAICS 443 Electronics and Appliances Stores – OLS Results

Variables	Model 1	Model 2	Model 3	Model 4 (Metro)	Model 5 (Micro)	Model 6	Model 7
Consumer1			0.097*** (0.017)	0.080*** (0.015)	0.065 (0.054)		
Consumer2			0.023 (0.016)	0.023 (0.015)	0.061 (0.032)		
Consumer3			0.002 (0.014)	0.001 (0.015)	0.013 (0.024)		
Consumer4			-0.004 (0.015)	-0.043** (0.016)	0.040 (0.031)		
Consumer5			-0.044** (0.015)	0.008 (0.014)	-0.143*** (0.031)		
labor_elec		0.130*** (0.007)	0.124*** (0.007)	0.101*** (0.007)	0.142*** (0.011)		0.185*** (0.007)
wage_elec		0.768*** (0.087)	0.723*** (0.086)	0.560*** (0.101)	0.767*** (0.133)		0.843*** (0.109)
stores_super	-2.051** (0.787)	-0.517 (0.599)	0.539 (0.636)	0.527 (0.707)	0.950 (1.028)		
msa_type	0.695*** (0.039)	0.352*** (0.035)	0.275*** (0.037)				
sales_total	0.024*** (0.002)	0.019*** (0.001)	0.017*** (0.001)	0.017*** (0.002)	0.016*** (0.002)		
pop_large100						-0.042*** (0.010)	-0.014* (0.007)
pop_large200						-0.009* (0.004)	-0.002 (0.003)
_cons	5.174*** (0.071)	3.706*** (0.109)	3.857*** (0.110)	4.476*** (0.142)	3.691*** (0.164)	6.391*** (0.036)	4.087*** (0.120)
N	586	586	586	321	265	607	607
Adj. R-sq	0.508	0.720	0.735	0.626	0.610	0.047	0.574

Standard errors in parentheses

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

Table 6.6. NAICS 443 Electronics and Appliances Stores – Mediation Analysis

	Indirect Effects				*Direct Effect*	*Total Effect*
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_elec)		Quality of Service (wage_elec)		Household Expenditure (ln_elec)	Household Expenditure (ln_elec)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.036** (0.014)	(0.008, 0.063)	0.001 (0.005)	(-0.008, 0.011)	0.097*** (0.017)	0.135*** (0.022)
Consumer2	0.041*** (0.012)	(0.018, 0.064)	-0.008 (0.005)	(-0.018, 0.002)	0.024 (0.016)	0.057** (0.021)
Consumer3	-0.003 (0.012)	(-0.027, 0.021)	0.002 (0.005)	(-0.008, 0.012)	0.002 (0.014)	-0.000 (0.018)
Consumer4	-0.001 (0.011)	(-0.022, 0.019)	-0.002 (0.005)	(-0.012, 0.009)	-0.004 (0.015)	-0.007 (0.019)
Consumer5	-0.032** (0.012)	(-0.056, -0.009)	-0.023*** (0.007)	(-0.038, -0.012)	-0.044** (0.015)	-0.099*** (0.019)
stores_super	-1.024* (0.518)	(-2.066, 0.008)	-0.203 (0.224)	(-0.643, 0.227)	0.539 (0.636)	-0.688 (0.823)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

6.4 NAICS 44511 Supermarkets and Other Grocery (except Convenience) Stores

Results of five OLS regressions are presented in Table 6.7. Since consumers normally shop for groceries around their home and workplace (Chen et al. 2020), it is very unlikely for consumers to go out-of-town and shop only for groceries, but it is possible that they will shop for groceries while traveling out-of-town for other purposes. Thus, Model 6 and Model 7 are not applicable for this industry and therefore they are not included in the analysis.

Results of Model 1 and Model 2 indicate that Supermarkets and Other Grocery (except Convenience) Stores is in intertype competition with Warehouse Clubs and Supercenters. The existence of warehouse clubs or supercenters in the focal area significantly draw sales away from supermarkets and other grocery stores. However, the quantity and quality of service provided by supermarkets and other grocery stores slightly mitigate the negative effect of intertype competition.

Results of Model 3, Model 4, and Model 5 indicate that there is no significant difference in the expenditure at supermarkets and other grocery stores between metropolitan and micropolitan areas. Furthermore, supermarkets and other grocery stores located in metropolitan areas face much more intense competition with warehouse clubs and supercenters than those in micropolitan areas. On average, consumer group 1 – rich young professionals spend significantly more amount of money at supermarkets and other grocery stores than baseline consumers; consumer group 2 – large full-nest households spend less, and consumer group 5 – large families with low education spend more than baseline consumers.

Table 6.8 presents the results of a mediation analysis on Model 3. The results show that the quantity of service is important and influential for this industry. The marketing mix variable – quantity of service significantly mediates the effects of all five consumer groups on household expenditure, but for consumer group 2 and consumer group 5 which present the two consumer groups of large households, the quantity of service may not work well for them. Moreover, high quality service provided at supermarkets and other grocery stores can mitigate the impact of competition from warehouse clubs and supercenters.

Table 6.7. NAICS 44511 Supermarkets and Other Grocery (except Convenience) Stores – OLS Results

Variables	Model 1	Model 2	Model 3	Model 4 (Metro)	Model 5 (Micro)
Consumer1			0.092*** (0.014)	0.085*** (0.014)	0.204*** (0.047)
Consumer2			-0.029* (0.012)	-0.028* (0.014)	-0.023 (0.023)
Consumer3			0.002 (0.013)	0.024 (0.015)	-0.030 (0.021)
Consumer4			-0.018 (0.013)	0.004 (0.016)	-0.078** (0.027)
Consumer5			0.044*** (0.013)	0.047** (0.015)	0.053* (0.026)
labor_food		0.009*** (0.001)	0.008*** (0.001)	0.008*** (0.001)	0.007*** (0.002)
wage_food		0.999*** (0.106)	0.810*** (0.103)	0.758*** (0.118)	0.769*** (0.181)
stores_super	-2.858*** (0.549)	-2.803*** (0.498)	-1.216* (0.512)	-1.993** (0.690)	-0.171 (0.787)
msa_type	0.109*** (0.030)	0.016 (0.028)	-0.028 (0.030)		
sales_total	0.016*** (0.001)	0.016*** (0.001)	0.013*** (0.001)	0.010*** (0.002)	0.014*** (0.002)
_cons	7.815*** (0.051)	6.636*** (0.112)	6.818*** (0.109)	7.005*** (0.142)	6.878*** (0.190)
N	489	489	489	292	197
Adj. R-sq	0.294	0.458	0.526	0.504	0.516

Standard errors in parentheses

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

Model 6 and Model 7 are not applicable for this industry, as consumers barely go out-of-town to shop for groceries.

Table 6.8. NAICS 44511 Supermarkets and Other Grocery (except Convenience) Stores – Mediation Analysis

	Indirect Effects				***Direct Effect***	***Total Effect***
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_food)		Quality of Service (wage_food)		Household Expenditure (ln_food)	Household Expenditure (ln_food)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.023*** (0.006)	(0.012, 0.036)	0.008 (0.005)	(-0.001, 0.018)	0.092*** (0.014)	0.123*** (0.015)
Consumer2	-0.010* (0.005)	(-0.021, 0.000)	-0.006 (0.004)	(-0.014, 0.003)	-0.029* (0.012)	-0.045*** (0.014)
Consumer3	0.018*** (0.006)	(0.006, 0.030)	0.005 (0.005)	(-0.005, 0.016)	0.002 (0.013)	0.025 (0.014)
Consumer4	0.024*** (0.006)	(0.014, 0.036)	0.001 (0.005)	(-0.009, 0.010)	-0.018 (0.013)	0.007 (0.014)
Consumer5	-0.032*** (0.008)	(-0.049, -0.018)	0.026*** (0.006)	(-0.080, -0.038)	0.044*** (0.013)	0.038** (0.014)
stores_super	-0.842*** (0.236)	(-1.317, -0.405)	0.954*** (0.192)	(0.582, 1.334)	-1.216* (0.512)	-1.103* (0.545)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

6.5 NAICS 447 Gasoline Stations

Results of three OLS regressions are presented in Table 6.9. Since Gasoline Stations does not compete with any of the retail subsectors/industries included in this dissertation, Model 1 and Model 2 are not applicable. Furthermore, consumers rarely go out-of-town to shop only for gas, Model 6 and Model 7 are not performed for this industry, either. Only results of Model 3, Model 4, and Model 5 are available.

Results of the three regressions indicate that gasoline expenditure is significantly lower in metropolitan areas than micropolitan areas. This is probably due to the widely available public transportation system in metropolitan areas. Consumers and households in micropolitan areas rely more on private vehicles and therefore they spend more on gas. On average, consumer group 1 – rich young professionals and consumer group 3 – college students spend less at gasoline stations compared with baseline consumers. Consumer group 1 is primarily located in metropolitan areas and therefore they spend less on gasoline. For consumer group 3, it is possible that majority of college students live in university dorms and they commute primarily by walking and carpooling with roommates or friends, and therefor they spend very little on gas on average. Unsurprisingly, consumer group 2 – large full-nest households spend more on gasoline than baseline consumers. Given the demographic characteristics of this group of consumers – large families with kids, it may be economically beneficial to commute with private vehicles compared with taking public transportation. Furthermore, this group of consumers usually shop at large superstores for value and convenience (Chen et al. 2020), and thus private vehicles facilitate their shopping needs at large but distant stores.

Table 6.10 presents the results of a mediation analysis on Model 3. The results show that consumer group 3 – college students and consumer group 4 – middle-class families are sensitive to the quantity of service provided at gasoline stations. Although quantity of service may not work well for consumer group 5 – large families with low education, quality of service is important for this group of consumers.

Table 6.9. NAICS 447 Gasoline Stations – OLS Results

Variables	Model 3	Model 4 (Metro)	Model 5 (Micro)
Consumer1	-0.145*** (0.012)	-0.148*** (0.012)	-0.143*** (0.032)
Consumer2	0.042*** (0.011)	0.032* (0.013)	0.049** (0.017)
Consumer3	-0.032** (0.010)	-0.042** (0.014)	-0.026 (0.015)
Consumer4	-0.013 (0.011)	-0.010 (0.015)	-0.015 (0.020)
Consumer5	0.010 (0.011)	0.000 (0.013)	0.024 (0.019)
labor_gas	0.037*** (0.005)	0.026*** (0.006)	0.047*** (0.007)
wage_gas	1.047*** (0.116)	0.826*** (0.172)	1.162*** (0.163)
msa_type	-0.126*** (0.025)		
sales_total	0.017*** (0.001)	0.018*** (0.001)	0.017*** (0.001)
_cons	6.983*** (0.097)	7.086*** (0.144)	6.834*** (0.142)
N	731	353	378
Adj. R-sq	0.501	0.521	0.463

Standard errors in parentheses

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

Model 1 and Model 2 are not applicable since this industry does not compete with other subsectors/industries included in this dissertation; Model 6 and Model 7 are not applicable since consumers barely go out-of-town to shop for gas.

Table 6.10. NAICS 447 Gasoline Stations – Mediation Analysis

	Indirect Effects				*Direct Effect*	*Total Effect*
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_gas)		Quality of Service (wage_gas)		Household Expenditure (ln_gas)	Household Expenditure (ln_gas)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.001 (0.004)	(-0.007, 0.007)	0.004 (0.004)	(-0.004, 0.013)	-0.145*** (0.012)	-0.140*** (0.013)
Consumer2	0.001 (0.004)	(-0.006, 0.008)	-0.001 (0.004)	(-0.008, 0.006)	0.042*** (0.011)	0.042*** (0.012)
Consumer3	0.013*** (0.004)	(0.006, 0.022)	0.006 (0.004)	(-0.002, 0.014)	-0.032** (0.010)	-0.014 (0.011)
Consumer4	0.033*** (0.006)	(0.022, 0.045)	-0.004 (0.004)	(-0.012, 0.003)	-0.013 (0.011)	0.017 (0.011)
Consumer5	-0.015*** (0.005)	(-0.025, -0.007)	0.014*** (0.004)	(0.006, 0.024)	0.010 (0.010)	0.009 (0.011)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

6.6 NAICS 448 Clothing and Clothing Accessories Stores

Results of seven OLS regressions are presented in Table 6.11. Results of Model 1 and Model 2 indicate that Clothing and Clothing Accessories Stores is in intertype competition with Discount Department Stores and Warehouse Clubs and Supercenters. The use of service by clothing stores can mitigate the negative impact from intertype competition.

Results of Model 3, Model 4, and Model 5 indicate that household expenditure at clothing stores is significantly higher in metropolitan areas than micropolitan areas. It is also evident that on average large households – consumer group 2 and consumer group 5 tend to spend more than baseline consumers, at clothing stores. Consumer group 4 – middle-class families tend to spend less at clothing stores, especially for those live in metropolitan areas.

Results of Model 6 and Model 7 indicate that consumers do go out-of-town to shop for clothing and accessories. Both the quantity and quality of service provided by local clothing stores can prevent some consumers from shopping out-of-town and thus reduce the harm to their businesses from competitors located in nearby large areas.

Table 6.12 presents the results of a mediation analysis on Model 3. This table shows that quantity of service is extremely important for this industry, as consumers need immediate access to employees while shopping in the stores for styling advice and sizing information. This service variable significantly mediates the effects of all five consumer groups on household expenditure at clothing stores. Furthermore, in order to compete with traditional department stores, clothing stores can do so by providing adequate and quality service to consumers.

Table 6.11. NAICS 448 Clothing and Clothing Accessories Stores – OLS Results

Variables	Model 1	Model 2	Model 3	Model 4 (Metro)	Model 5 (Micro)	Model 6	Model 7
Consumer1			0.043 (0.023)	0.021 (0.020)	0.123 (0.105)		
Consumer2			0.054** (0.020)	0.011 (0.018)	0.125* (0.063)		
Consumer3			-0.016 (0.020)	-0.000 (0.018)	-0.065 (0.056)		
Consumer4			-0.087*** (0.021)	-0.081*** (0.021)	0.007 (0.074)		
Consumer5			0.057** (0.019)	0.070*** (0.016)	-0.297*** (0.071)		
labor_cloth		0.150*** (0.010)	0.132*** (0.011)	0.131*** (0.010)	0.067** (0.025)		0.209*** (0.008)
wage_cloth		1.019*** (0.244)	0.797** (0.244)	1.718*** (0.283)	-0.145 (0.457)		0.608* (0.244)
stores_dept	-1.381 (1.044)	1.671 (0.876)	2.167* (0.891)	1.960 (1.042)	2.503 (1.581)		
stores_ddept	-4.081*** (0.916)	-3.322*** (0.751)	-2.177** (0.769)	-1.341 (0.920)	2.011 (1.618)		
stores_super	-3.314*** (0.993)	-1.612* (0.818)	-0.736 (0.850)	1.000 (0.858)	-0.048 (1.810)		
msa_type	0.481*** (0.054)	0.236*** (0.049)	0.204*** (0.050)				
sales_total	0.031*** (0.002)	0.024*** (0.002)	0.022*** (0.002)	0.023*** (0.002)	0.011** (0.004)		
pop_large100						-0.037*** (0.011)	-0.023** (0.008)
pop_large200						-0.015*** (0.004)	-0.009** (0.003)
_cons	6.126*** (0.096)	4.147*** (0.215)	4.416*** (0.221)	3.916*** (0.229)	4.919*** (0.447)	7.082*** (0.040)	4.738*** (0.195)
N	454	454	454	328	106	663	663
Adj. R-sq	0.424	0.616	0.638	0.657	0.461	0.045	0.514

Standard errors in parentheses

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

Table 6.12. NAICS 448 Clothing and Clothing Accessories Stores – Mediation Analysis

	Indirect Effects				***Direct Effect***	***Total Effect***
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_cloth)		Quality of Service (wage_cloth)		Household Expenditure (ln_cloth)	Household Expenditure (ln_cloth)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.069*** (0.016)	(0.039, 0.102)	0.011* (0.005)	(0.003, 0.024)	0.043 (0.023)	0.123*** (0.025)
Consumer2	0.031* (0.013)	(0.005, 0.058)	0.004 (0.004)	(-0.003, 0.014)	0.054** (0.020)	0.088*** (0.023)
Consumer3	0.026* (0.013)	(0.000, 0.052)	-0.004 (0.005)	(-0.015, 0.003)	-0.016 (0.020)	0.006 (0.023)
Consumer4	-0.053*** (0.014)	(-0.082, -0.028)	-0.005 (0.004)	(-0.014, 0.001)	-0.087*** (0.021)	-0.145*** (0.024)
Consumer5	0.032** (0.014)	(0.006, 0.060)	-0.002 (0.003)	(-0.009, 0.004)	0.057** (0.019)	0.087*** (0.022)
stores_dept	-1.958*** (0.556)	(-3.205, -0.990)	0.186 (0.176)	(-0.079, 0.619)	2.167* (0.891)	0.395 (1.021)
stores_ddept	-0.113 (0.505)	(-1.085, 0.902)	0.186 (0.161)	(-0.139, 0.513)	-2.177** (0.769)	-2.104* (0.892)
stores_super	-0.027 (0.579)	(-1.114, 1.218)	0.033 (0.186)	(-0.374, 0.382)	-0.736 (0.850)	-0.730 (0.988)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

6.7 NAICS 452111 Department Stores (except Discount Department Stores)

Results of five OLS regressions are presented in Table 6.13. Since there are only 13 micropolitan areas in the sample for this industry, I exclude these observations on Model 1, Model 2, and Model 3, and I examine metropolitan areas only for the first three models. Since

Model 4 is the same as Model 3, it is not listed in the table. Model 5 is not applicable due to the small sample size of micropolitan areas.

Results of Model 1 and Model 2 indicate that traditional department stores are in intratype competition with discount department stores in metropolitan areas. However, interestingly, the existence of clothing stores has a positive impact on expenditure at traditional department stores. There are two possible explanations. First, traditional department stores are in intertype competition with clothing stores, but consumers prefer to shop at traditional department stores for clothing and clothing accessories, therefore, traditional department stores take sales away from clothing stores located in the same area. This is evident in Table 6.11 that the sign of *stores_dept* on *ln_cloth* (Model 1) is negative, although it is not significant. Second, this is probably due to the "agglomeration effect" (Miller et al. 1999). Since traditional department stores and clothing stores are usually located in clusters or in the same shopping area (*i.e.*, shopping center), traditional department stores benefit from the consumer flow going to the same shopping area, which is especially true in high density metropolitan areas. Furthermore, quality of service provided at traditional department stores can mitigate the negative impact from intratype competition with discount department stores.

Results of Model 3 indicate that consumer group 1 – rich young professionals spend less than baseline consumers at traditional department stores. Results of Model 6 and Model 7 indicate that consumers would go 100 miles to shop at department stores in large areas. Service provided by local department stores especially high-quality service can mitigate the negative effect of consumer out-of-town shopping.

Table 6.14 presents the results of a mediation analysis on Model 3. The results indicate that for traditional department stores quality of service is vital, as it significantly mediates the effects of consumer group 2 – large full-nest households and consumer group 5 – low-income large families on their household expenditures. Furthermore, quality of service is an important indicator when consumers decide whether to shop at clothing stores, traditional department stores, or value-driven discount department stores for clothing related products.

Table 6.13. NAICS 452111 Department Stores (except Discount Department Stores) – OLS Results

Variables	Model 1 (Metro)	Model 2 (Metro)	Model 3 (Metro)	Model 6	Model 7
Consumer1			-0.055** (0.019)		
Consumer2			0.014 (0.020)		
Consumer3			-0.012 (0.023)		
Consumer4			-0.037 (0.022)		
Consumer5			0.024 (0.020)		
labor_dept		0.000 (0.000)	0.000 (0.001)		0.001 (0.001)
wage_dept		1.178*** (0.272)	0.980*** (0.281)		1.095*** (0.266)
stores_cloth	0.118** (0.041)	0.057 (0.041)	0.043 (0.046)		
stores_ddept	-2.117* (1.026)	-0.098 (1.097)	0.083 (1.130)		
stores_super	-0.945 (0.827)	-0.265 (0.888)	-0.838 (1.038)		
sales_total	0.012*** (0.003)	0.013*** (0.003)	0.015*** (0.003)		
pop_large100				-0.015* (0.007)	-0.014* (0.007)
pop_large200				0.004 (0.003)	0.005 (0.003)
_cons	5.945*** (0.087)	4.931*** (0.253)	5.047*** (0.266)	6.396*** (0.026)	5.501*** (0.210)
N	200	200	200	213	213
Adj. R-sq	0.297	0.353	0.379	0.019	0.090

Standard errors in parentheses.

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

Model 4 is the same as Model 3; Model 5 is not applicable for this industry since there are only 13 micropolitan areas.

Table 6.14. NAICS 452111 Department Stores (except Discount Department Stores) – Mediation Analysis

	Indirect Effects				*Direct Effect*	*Total Effect*
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_dept)		Quality of Service (wage_dept)		Household Expenditure (ln_dept)	Household Expenditure (ln_dept)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.006 (0.007)	(-0.006, 0.021)	-0.009 (0.005)	(-0.020, 0.000)	-0.055** (0.019)	-0.058** (0.018)
Consumer2	0.004 (0.005)	(-0.005, 0.014)	0.009* (0.006)	(-0.001, 0.025)	0.014 (0.020)	0.030 (0.019)
Consumer3	-0.002 (0.003)	(-0.009, 0.003)	0.010 (0.007)	(-0.003, 0.026)	-0.012 (0.023)	-0.004 (0.024)
Consumer4	-0.003 (0.004)	(-0.012, 0.003)	-0.003 (0.005)	(-0.014, -0.006)	-0.037 (0.022)	-0.044* (0.022)
Consumer5	0.004 (0.005)	(-0.004, 0.014)	0.012* (0.005)	(0.001, 0.022)	0.024 (0.020)	0.040* (0.019)
stores_cloth	0.001 (0.004)	(-0.007, 0.009)	0.046** (0.018)	(0.013, 0.081)	0.043 (0.046)	0.090* (0.045)
stores_ddept	-0.066 (0.122)	(-0.392, 0.085)	-1.558** (0.613)	(-2.694, -0.331)	0.083 (1.130)	-1.541 (1.061)
stores_super	-0.096 (0.162)	(-0.520, 0.130)	-0.492 (0.276)	(-0.998, 0.030)	-0.838 (1.038)	-1.426 (1.048)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

6.8 NAICS 452112 Discount Department Stores

Results of five OLS regressions are presented in Table 6.15. For the same reasons mentioned in section 6.7, there are only five models available for Discount Department Stores. Results of Model 1 and Model 2 indicate that discount department stores are in intratype competition with warehouse clubs and supercenters. However, discount department stores benefit from the existence of traditional department stores in the focal area, which indicates that discount department stores takes sales away from traditional department stores. This is also evident in Table 6.13 (Model 1) that the effect of *stores_ddept* on *ln_dept* is negative and significant. Since discount department stores primarily sell past-season and clearance products which are no longer carried by traditional department stores, shopping at discount department stores saves consumers' economic cost (*i.e.*, dollars) spent on the same brand-name products they would have bought from traditional department stores. Furthermore, quantity and quality of service mitigate the negative impact from competition with warehouse clubs and supercenters and strengthen the benefit from the existence of traditional department stores, but also make the competition with clothing stores more severe.

Results of Model 3 indicate that on average consumer group 4 – middle-class families spend more at discount department stores than baseline consumers. Results of Model 6 indicate that occasionally consumers from large areas are willing to travel to nearby small areas to shop at their discount department stores. This behavior is very likely driven by 1) utilities from buying brand-name but reduced-price items (*i.e.*, discounted); and 2) psychological satisfaction of getting a good deal. It is also possible that discount department stores are located in relatively less populated areas, where traditional department stores are very rare.

Table 6.16 presents the results of a mediation analysis on Model 3. This table shows that consumer group 1 – rich young professionals, consumer group 3 – college students, and consumer group 5 – large families with low education are very sensitive to quantity of service provided by discount department stores. On the other hand, consumer group 2 – large full-nest households and consumer group 4 – middle-class families do not care too much about service provided at discount department stores.

Table 6.15. NAICS 452112 Discount Department Stores– OLS Results

	Model 1 (Metro)	Model 2 (Metro)	Model 3 (Metro)	Model 6	Model 7
Consumer1			0.029 (0.023)		
Consumer2			0.029 (0.023)		
Consumer3			0.029 (0.027)		
Consumer4			0.145*** (0.022)		
Consumer5			-0.024 (0.023)		
labor_ddept		0.009*** (0.001)	0.010*** (0.001)		0.005*** (0.001)
wage_ddept		1.715*** (0.306)	1.267*** (0.300)		1.809*** (0.345)
stores_cloth	0.043 (0.060)	-0.214*** (0.050)	-0.076 (0.050)		
stores_dept	3.772* (1.671)	8.089*** (1.336)	7.729*** (1.337)		
stores_super	-7.398*** (1.415)	-1.142 (1.235)	-2.619* (1.200)		
sales_total	0.012* (0.005)	0.018*** (0.004)	0.010** (0.004)		
pop_large100				0.019 (0.010)	0.011 (0.009)
pop_large200				0.012** (0.004)	0.006 (0.003)
_cons	6.713*** (0.130)	4.019*** (0.259)	4.487*** (0.253)	6.948*** (0.035)	5.152*** (0.240)
N	205	205	205	225	225
Adj. R-sq	0.131	0.514	0.608	0.077	0.306

Standard errors in parentheses.

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

Model 4 is the same as Model 3; Model 5 is not applicable for this industry since there are only 20 micropolitan areas.

Table 6.16. NAICS 452112 Discount Department Stores – Mediation Analysis

	Indirect Effects				***Direct Effect***	***Total Effect***
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_ddept)		Quality of Service (wage_ddept)		Household Expenditure (ln_ddept)	Household Expenditure (ln_ddept)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.056** (0.018)	(0.023, 0.092)	-0.007 (0.007)	(-0.022, 0.006)	0.029 (0.023)	0.078** (0.030)
Consumer2	0.000 (0.024)	(-0.050, 0.044)	-0.002 (0.008)	(-0.018, 0.014)	0.029 (0.023)	0.028 (0.031)
Consumer3	0.064** (0.023)	(0.023, 0.111)	0.041** (0.013)	(0.017, 0.068)	0.029 (0.027)	0.134*** (0.034)
Consumer4	-0.009 (0.017)	(-0.040, 0.027)	0.014 (0.007)	(0.002, 0.030)	0.145*** (0.022)	0.151*** (0.030)
Consumer5	0.068*** (0.022)	(0.031, 0.117)	0.002 (0.007)	(-0.012, 0.017)	-0.024 (0.023)	0.045 (0.029)
stores_cloth	0.088* (0.034)	(0.018, 0.151)	0.072** (0.024)	(0.033, 0.127)	-0.076 (0.050)	0.085 (0.064)
stores_dept	-2.882** (1.061)	(-4.920, -0.772)	-0.647 (0.434)	(-1.524, 0.226)	7.729*** (1.337)	4.201* (1.760)
stores_super	-3.876*** (1.351)	(-6.954, -1.716)	-0.122 (0.505)	(-1.260, 0.732)	-2.619* (1.200)	-6.617*** (1.534)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

6.9 NAICS 45291 Warehouse Clubs and Supercenters

Results of five OLS regressions are presented in Table 6.17. For the same reasons mentioned in previous sections, only Model 1, Model 2, Model 3, Model 6, and Model 7 are available for this industry. However, due to the small sample size of only 53 metropolitan areas, I’m afraid that this section is only for informational purposes. Results obtained from the analysis may not be generalized to the broader metropolitan area sample.

Results of Model 1 and Model 2 indicate that when there exist all types of businesses in the focal area, Warehouse Clubs and Supercenters is in intertype competition with Supermarkets and Other Grocery (except Convenience) Stores, Clothing and Clothing Accessories Stores, and Department Stores (except Discount Department Stores). Warehouse clubs and supercenters take sales away from traditional department stores, but the existence of supermarkets and clothing stores reduces household expenditures at warehouse clubs and supercenters. Quantity and quality of service can mitigate the negative impact of competing with supermarkets, and interestingly attract more consumers to shop at warehouse clubs and supercenters for electronics and appliances. Results of Model 3 indicate that on average consumer group 4 – middle-class families favor warehouse clubs and supercenters and they spend more than baseline consumers. Results of Model 6 and Model 7 show that consumers are willing to travel out-of-town to shop at warehouse clubs and supercenters.

Table 6.18 presents the results of a mediation analysis on Model 3. This table shows that neither quantity or quality of service mediate the effects of five consumer groups on their household expenditures at warehouse clubs and supercenters.

Table 6.17. NAICS 45291 Warehouse Clubs and Supercenters – OLS Results

Variables	Model 1 (Metro)	Model 2 (Metro)	Model 3 (Metro)	Model 6	Model 7
Consumer1			-0.066 (0.051)		
Consumer2			0.082 (0.088)		
Consumer3			-0.151 (0.103)		
Consumer4			0.187* (0.089)		
Consumer5			-0.051 (0.082)		
labor_super		0.003* (0.002)	0.001 (0.001)		0.001 (0.001)
wage_super		1.574* (0.614)	0.301 (0.732)		1.445* (0.651)
stores_fur	1.017 (0.673)	1.033 (0.663)	0.394 (0.592)		
stores_elec	1.791 (1.180)	2.735* (1.167)	1.720 (0.999)		
stores_food	-0.717* (0.328)	-0.691* (0.319)	0.326 (0.433)		
stores_cloth	-0.970** (0.339)	-1.175** (0.334)	-0.584 (0.319)		
stores_dept	16.573** (5.901)	14.718* (5.587)	9.377 (5.158)		
stores_ddept	-6.751 (4.354)	-7.679 (4.128)	-15.003*** (3.856)		
sales_total	0.020 (0.010)	0.024* (0.010)	0.013 (0.009)		
pop_large100				-0.009 (0.019)	-0.018 (0.018)
pop_large200				-0.031*** (0.009)	-0.030*** (0.009)
_cons	7.651*** (0.521)	5.025*** (1.018)	7.386*** (1.167)	8.467*** (0.065)	6.679*** (0.880)
N	53	53	53	57	57
Adj. R-sq	0.330	0.416	0.617	0.217	0.258

Standard errors in parentheses

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

Model 4 is the same as Model 3; Model 5 is not applicable for this industry since there are only 4 micropolitan areas.

Table 6.18. NAICS 45291 Warehouse Clubs and Supercenters – Mediation Analysis

	Indirect Effects				*Direct Effect*	*Total Effect*
	Marketing mix Variables				**Dependent Variable**	**Dependent Variable**
	Quantity of Service (labor_super)		Quality of Service (wage_super)		Household Expenditure (ln_super)	Household Expenditure (ln_super)
Variables	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (Bootstrap S.E.)	Bootstrap C.I.	Coeff. (S.E.)	Coeff. (S.E.)
Consumer1	0.002 (0.015)	(-0.034, 0.031)	0.001 (0.014)	(-0.032, 0.027)	-0.066 (0.051)	-0.063 (0.050)
Consumer2	0.012 (0.042)	(-0.055, 0.120)	-0.021 (0.056)	(-0.144, 0.082)	0.082 (0.088)	0.073 (0.073)
Consumer3	-0.008 (0.035)	(-0.093, 0.052)	-0.000 (0.024)	(-0.044, 0.059)	-0.151 (0.103)	-0.160 (0.100)
Consumer4	-0.002 (0.021)	(-0.057, 0.036)	0.026 (0.072)	(-0.108, 0.177)	0.187* (0.089)	0.211** (0.063)
Consumer5	0.003 (0.026)	(-0.055, 0.053)	0.004 (0.023)	(-0.035, 0.059)	-0.051 (0.082)	-0.044 (0.079)
stores_fur	-0.107 (0.306)	(-0.887, 0.382)	0.003 (0.122)	(-0.234, 0.2706	0.394 (0.592)	0.290 (0.558)
stores_elec	-0.221 (0.581)	(-1.288, 1.124)	-0.003 (0.198)	(-0.486, 0.366)	1.720 (0.999)	1.496 (0.918)
stores_food	-0.044 (0.173)	(-0.403, 0.328)	0.054 (0.202)	(-0.376, 0.470)	0.326 (0.433)	0.336 (0.407)
stores_cloth	0.060 (0.151)	(-0.264, 0.372)	0.021 (0.113)	(-0.134, 0.326)	-0.584 (0.319)	-0.504 (0.288)
stores_dept	0.725 (3.020)	(-3.164, 9.381)	-0.518 (1.849)	(-5.172, 2.390)	9.377 (5.158)	9.584 (4.890)
stores_ddept	-0.287 (1.499)	(-4.172, 2.252)	-0.295 (1.232)	(-2.738, 2.484)	-15.003*** (3.856)	-15.585*** (3.655)

Bootstrap C.I.: 95% bootstrap confidence intervals using 5,000 bootstrap samples.

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

Chapter 7 Discussion

In this chapter, I discuss the empirical results presented in Chapter 6, and then I summarize the theoretical contributions and managerial implications of this dissertation.

7.1 Discussion of Results

In this section of the chapter, I organize the discussion of the empirical results by variables.

7.1.1 Consumer Socioeconomic and Demographic Variables

Recall the five consumer groups and the baseline consumers:

- Consumer Group 1 – “rich young professionals” characterized by highly educated, high income (at least $75,000 annual household income), middle age (35 – 44), living in high density populated areas.
- Consumer Group 2 – “large full-nest households” characterized by large household size with children and their parents under age 44.
- Consumer Group 3 – “college students” characterized by highly educated, young (age between 15 and 24), living primarily in university housing.
- Consumer Group 4 – “middle-class families” characterized by annual household income between $75,000 to $99,999, living in less populated areas where autos (*i.e.*, private vehicles) are needed to commute.
- Consumer Group 5 – “large families with low education” characterized by large household size and low level of education.
- Baseline consumers – “female single-head” characterized by middle income ($50,000 - $74,999), age between 45 and 59, single-head of the household.

Since these consumer groups are highly characterized by their income, household size, family lifecycle, and auto ownership, I will discuss the results based on the abovementioned consumer and household characteristics. It has been shown in the literature that household income positively influences retail expenditure across lines of retail trade (e.g., Ingene and Lusch 1981, Chen et al. 2020). Intuitively, households with more income will have more disposable income to spend on buying products and services. Therefore, in general, consumer groups which a higher than average (($50,000 - $74,999) income will purchase more and incur more expenditure than the baseline consumers. However, it varies across lines of trade. Furthermore, household size is found to have a positively relationship with household expenditures, since more people per household will lead to higher demand and needs for products and services. Therefore, larger households will spend more than baseline consumers, on average. For households which are characterized by full nest, they are in special needs of necessities and clothing products, and therefore will spend more on food and clothing than the baseline consumers. Automobile ownership has been shown to positively influence sales in the literature (e.g., Ingene 1984). For consumers living in less populated and low-density areas, autos become necessary as means of transportation, which also facilitate their shopping at large but distant stores such as supermarkets and warehouse clubs and supercenters. Therefore, households with more autos tend to spend more.

Table 7.1 summarizes the results on the set of consumer group variables.

For consumer group 1 – rich young professionals, compared with baseline consumers, they spend about the same amount of money at new car dealers, clothing stores, discount department stores, and warehouse clubs and supercenters; they spend more at furniture stores, electronics stores, and supermarkets; but they spend less at gasoline stations and traditional department stores. Since this consumer group is characterized by high income and located primarily in high density metropolitan areas where public transportation is widely available, it is very likely that this group of consumers have less need for private vehicles and therefore incur less expenditure at gasoline stations. Thus, on average, they spend about the same amount on new cars but less expenditure at gasoline stations. It is not surprising that they spend more at furniture stores, electronics and appliances stores, and supermarkets, because they can afford to buy more quantity and more expensive items. Furthermore, this consumer group consists primarily of young professionals who are single, therefore, their per household expenditure is equivalent to their per capita expenditure. If they tend to shop more at clothing stores and discount department stores but less at traditional department stores, it makes sense that on average when compared with the baseline consumers, they spend about the same as a single-head household which may consist of multiple people per household at clothing stores and discount department stores, and less at traditional department stores. Lastly, warehouse clubs and supercenters are usually favored by large families, thus, it is not surprising that this consumer group spend about the same as a single-head household at warehouse clubs and supercenters.

For consumer group 2 – large full-nest households, results indicate that compared with baseline consumers, they spend about the same amount for new cars, electronics, traditional department stores, discount department stores, and warehouse clubs and supercenters; they spend less at furniture stores and supermarkets; but more at gasoline stations and clothing stores. Since this group of consumers is characterized by its large household size with children in all ages, it would be more economically and temporally beneficial for this group of consumers to shop at large stores where they can do a one-stop shopping for the entire family. Furthermore, private vehicles may be economically and temporally more efficient for this group of consumers to commute, thus, it is not surprising that they spend more than baseline consumers at gasoline stations. Moreover, large full-nest households are in special needs of groceries and clothing products, it is possible that an average income large full-nest household would favor warehouse clubs and supercenters, compared to its competitors – furniture stores, supermarket, etc. Thus, they would reduce expenditures at furniture stores and supermarkets, and instead shop at warehouse clubs and supercenters. The sign of *Consumer2* on *ln_super* is positive but not significant, probably due to the small sample size at warehouse clubs and supercenters. In summary, consumer group 2 is in greater needs of gas to commute in private vehicles and they spend more on clothing products. It is possible that this group favors warehouse clubs and supercenters but given the small sample size the effect is not significant.

For consumer group 3 – college students, results indicate they incur about the same expenditures as baseline consumers across the nine lines of trade except for gasoline stations. This consumer group is characterized by low income, living primarily in university housing. When averaging this group of consumers on a per household basis, they can be treated as a regular household. Therefore, it is not surprising that they exhibit similar expenditure patterns as a single-head household. Since they live primarily in university dorms, they may use other means of transportation more frequently than private vehicles, such as walking, carpooling, bicycling, or riding a bus. Thus, they spend less at gasoline stations.

As for consumer group 4 – middle-class families, they spend about the same as baseline consumers on new cars, electronics, food and beverage at supermarkets, gas, and at traditional department stores; they spend less at furniture stores and clothing stores, but more at discount department stores and warehouse clubs and supercenters. It is likely that middle-class families favor discount department stores and warehouse clubs and supercenters over their competitors. It is evident that their expenditures at furniture stores and clothing stores are lower, which implies that they may shop for these needs at discount department stores and warehouse clubs and supercenters.

Lastly, for consumer group 5 – large families with low education, it shows that this consumer group tends to spend about the same amount of money as baseline consumers at new car dealers, gasoline stations, traditional department stores, discount department stores, and warehouse clubs and supercenters; and they spend less at furniture stores and electronics stores, but more at supermarkets and clothing stores. Since education is positively related with income, it is likely that this group of consumers are primarily large families with low education and low income. For low-income families, whom usually have a tight budget on expenditures, they would try to lower their economic costs associated with their shopping needs. Therefore, they would allocate their budget primarily on necessities – food and clothing for those households with children. Thus, this group spends more at supermarkets and clothing stores.

Table 7.1. Results – Consumer Groups

	Consumer1	Consumer2	Consumer3	Consumer4	Consumer5
ln_car	0	0	0	0	0
ln_fur	+	-	0	-	-
ln_elec	+	0	0	0	-
ln_food	+	-	0	0	+
ln_gas	-	+	-	0	0
ln_cloth	0	+	0	-	+
ln_dept	-	0	0	0	0
ln_ddept	0	0	0	+	0
ln_super	0	0	0	+	0

7.1.2 Intertype and Intratype Competition Variables

The logic is straightforward for intertype and intratype competitions – competitors take sales away and therefore retail expenditures will decrease in the focal subsector/industry. Table 7.2 summarizes the results regarding the competition variables. Below I discuss some of the results which are worth more investigation.

For clothing and clothing accessories stores, it has shown that when there exists traditional department store, discount department stores, and warehouse clubs and supercenters in the trading area, the threat from traditional department stores is very minor and not significant. Instead, discount department stores and warehouse clubs and supercenters do take sales away from clothing stores.

For traditional department stores, surprisingly, the existence of clothing stores has a positive impact on their sales. There are two possible explanations. One, traditional department stores takes sales away from clothing stores. Second, when stores are located in clusters, traditional department stores benefit from the consumer flow going to the surrounding area – agglomeration effect (Miller et al. 1999). This is especially true in metropolitan areas where population and store density are relatively high. Furthermore, discount department stores located in the trading area would take sales away from traditional department stores, but the negative impact from competition with warehouse clubs and supercenters is not significant.

For discount department stores, they face direct competition with warehouse clubs and supercenters. Moreover, the existence of traditional department stores benefits discount department stores, which indicates that discount department stores take sales away from traditional department stores.

For warehouse clubs and supercenters, due to its small sample size, results may not be representative. It has shown that with all the intertype competitors and intratype competitors in the area, supermarkets and clothing stores are strong competitors of warehouse clubs and supercenters. Furthermore, warehouse clubs and supercenters benefit from existence of traditional department stores in the area, which indicates that they take sales away from traditional department stores.

Table 7.2. Results – Intertype and Intratype Competition Variables

Retail Subsector/Industry	Intertype Competition with	Results	Intratype Competition with	Results
NAICS 442 Furniture Stores	NAICS 45291	✓		
NAICS 443 Electronics	NAICS 45291	✓		
NAICS 44511 Supermarkets and Other Grocery (except Convenience) Stores	NAICS 45291	✓		
NAICS 448 Clothing & Accessories Stores	NAICS 452111	0		
	NAICS 452112	✓		
	NAICS 45291	✓		
NAICS 452111 Department Stores (except Discount Department Stores)	NAICS 448	+	NAICS 452112	✓
			NAICS 45291	0
NAICS 452112 Discount Department Stores	NAICS 448	0	NAICS 452111	+
			NAICS 45291	✓
NAICS 45291 Warehouse Clubs & Supercenters	NAICS 442	0	NAICS 452111	+
	NAICS 443	0	NAICS 452112	0
	NAICS 44511	✓		
	NAICS 448	✓		

"✓" indicates there exists competition.
"0" indicates there is no competition from the focal industry's perspective.
"+" indicates the focal industry benefits from the existence of competitors.

7.1.3 Marketing mix Variables

Marketing mix variables are known to have a strong positive impact on retail sales across various lines of retail trade. The two marketing mix variables examined in this dissertation are quantity and quality of service. As discussed in Chapter 3, service provided by retailers will reduce consumers' temporal cost of shopping inside the store, and thus increase their overall shopping utilities, *ceteris paribus*. More employees per store (*i.e.*, labor) and higher quality of service (*i.e.*, wage) will provide consumers with immediate access to assistance when needed. Consumers will be able to experience faster check-outs, longer store operating hours, personal shopping advice, etc. All service effort made by retailers will transform to customer satisfaction and thereby raising household expenditures. Therefore, quantity and quality of service should positively influence household expenditures across all lines of retail trade.

The results presented in Chapter 6 have indicated that the two marketing mix variables – quantity and quality of service significantly increase retail expenditures across nine retail subsectors/industries, except for quantity of service at NAICS 452111 Department Stores (except Discount Department Stores). It has shown that quantity of service does not influence retail expenditure at traditional department stores, but quality of service does. The rational is that traditional department stores usually target at middle-class or upper-class consumers who look for a personalized shopping experience. Therefore, these consumers may weigh more on the quality dimension of service rather than quantity. Another possible explanation is that this zero-effect may be due to the small sample size.

7.1.4 Consumer Out-shopping Variables

Theoretically, Reilly (1931) demonstrated that consumers would go out-of-town to shop at large cities for a wider variety of products. Recall Figure 2.2 in Chapter 2, population and distance are two determinants for consumers' decision on out-of-town shopping. In general, if there exists a nearby city or area which has larger population than the focal area, then the focal area would lose sales to the nearby large areas.

Results of this set of consumer out-shopping variables are summarized in Table 7.3. In general, consumers do out-of-town shopping, with a range of up to 200 miles, across various lines of retail trade. However, the results of the three industries within NAICS 452 may be biased since only a small number of observations are available at micropolitan areas. Furthermore, when metropolitan areas are in close proximity to a few other metropolitan areas, the effects of these population variables may be confounding and inaccurate.

Table 7.3. Results – Consumer Out-shopping Variables

	pop_large100	**pop_large200**
ln_car	✓	✓
ln_fur	✓	✓
ln_elec	✓	✓
ln_cloth	✓	✓
ln_dept	✓	0
ln_ddept	0	✓
ln_super	0	✓

"✓" indicates that there exists consumer out-shopping.
"0" indicates that there is no consumer out-shopping.

7.2 Theoretical Contributions

This section presents the theoretical contributions of this dissertation to the marketing literature. Recall that this dissertation aims to address four research questions:

1. how intertype and intratype competition influence retail expenditures per household.
2. what key variables drive retail expenditures per household across different retail lines.
3. whether those key variables differ in metropolitan areas vs. micropolitan areas.
4. whether large areas draw retail sales (*i.e.*, expenditures) from nearby small areas.

This dissertation contributes to the literature by empirically addressing the abovementioned research questions.

First, only a few studies in the literature have examined intertype and intratype competition of different retail lines at the aggregate level (e.g., Ingene 1983, Miller et al. 1999). In this dissertation, I show that several lines of retail trade are in competition with each other. Specifically, warehouse clubs and supercenters are competing with all other retail lines including furniture stores, electronics stores, supermarkets, clothing stores, traditional department stores, and discount department stores. Moreover, clothing stores also face competition with traditional department stores and discount department stores, and traditional department store suffer from the existence of discount department stores.

Second, I show that key variables driving retail expenditures vary across different retail lines, which has both theoretical contributions and managerial implications for businesses. I show that the five consumer groups exhibit different expenditure patterns across nine retail lines; the two marketing mix variables are significant across nine retail lines except for quantity of service provided at traditional department stores; intertype and intratype competition exist in seven retail lines, except for New Car Dealers and Gasoline Stations; and consumer out-shopping variables significantly influence local sales.

Third, I show that key variables of retail expenditures are slightly different between metropolitan areas and micropolitan areas. Micropolitan areas have not been examined in the literature, to the best of my knowledge, since information regarding these areas became available only since 2012.

Fourth, I construct and examine a set of variables measuring consumers' out-of-town shopping behavior. Based on Reilly's *Law of Retail Gravitation* (1931), I construct two variables capturing the population from nearby large areas within 200 miles of the focal area. The results are consistent with Reilly's work, in that population from large areas significantly draw sales away but the effect diminishes as distance increases from 100 miles to 200 miles.

Furthermore, I test the mediation effects of the two marketing mix variables, respectively, as proposed in the theoretical framework. Results indicate that there exist mediation effects of consumer group variables (*i.e.*, socioeconomic and demographic variables) and competitive variables on household expenditures, which implies that marketing mix variables are indeed endogenously determined in light of consumers' socioeconomic and demographic characteristics, and the competitive forces in the local market.

In summary, this dissertation has built upon and extended our current understanding of the marketing (*i.e.*, retailing) literature by empirically examining the theoretical framework developed in this dissertation.

7.3 Managerial Implications

In this section, I discuss the managerial implications from this dissertation.

First, understanding competition is vital for businesses. Competition exists from intertype and intratype competitors, and from competitors outside the focal area if it is located within close proximity (*i.e.*, 200 miles) to a few highly populated areas. The results indicate that warehouse clubs and supercenters are in competition with almost all retail subsectors/industries examined in this dissertation. Because warehouse clubs and supercenters provide a wide variety of products encompassing all departments, at a competitive price level, consumers can save both their economic and temporal costs with this type of "value-driven one-stop shopping". Therefore, for businesses that are in direct competition with warehouse clubs and supercenters, such as furniture stores, electronics and appliances stores, supermarkets, clothing and accessories stores, traditional department stores, and discount department stores, decisions on store location and store size are very important for their success. Since every store has a maximum range that consumers are willing to travel to, if located too close and in the travel range of a competitor, it is possible that this store will lose sales due to competition. Store size also matters because large stores can carry more product assortment while small stores can only carry a limited assortment, but one should note that shopping at large stores would incur a greater temporal cost inside the store, compared with a relatively low temporal cost of shopping at small stores. Therefore, store location and store size influence consumers' shopping benefits and shopping costs, and thus impact store profitability.

Second, I have shown that marketing mix efforts, specifically the quantity and quality of service significantly influence and increase household expenditures at nine retail subsectors/industries except for quantity of service at traditional department stores. This is consistent with the findings in the extant literature. Furthermore, I have demonstrated that retailers' service efforts are managerially determined in response to the socioeconomic and demographic characteristics and competitive forces they face in the focal area. When there exists competition in the focal area, adequate and quality service will help the business mitigate the negative impact of competition. Service will also prevent consumers from shopping out-of-town, to some extent. In summary, both quantity and quality of service are important factors for increasing sales and retaining customers.

Third, I have shown that the five consumer groups have different expenditure patterns across the nine retail subsectors/industries. This is important to retailers and marketers for market segmentation purposes. Moreover, I demonstrate in this dissertation that the five consumer groups exhibit slightly different expenditure patterns in metropolitan areas versus micropolitan areas. This will facilitate retailers' effort in market segmentation one step further.

Fourth, I have compared the key variables driving retail expenditures in both metropolitan areas and micropolitan areas. It is important for retailers to understand that even the same type of consumers may exhibit different expenditure patterns in different geographic areas. This will be useful for retailers as they look to open new stores in micropolitan areas.

Lastly, I have empirically examined the phenomenon of consumer out-of-town shopping. Although the results are intuitive that consumers would go to large cities/areas for shopping purposes, it is unclear in the literature that how far consumers would travel, what they would shop out-of-town, and whether the negative impact to local businesses is significant. I have shown that consumers are willing to travel up to 200 miles (*i.e.*, approximately 3-hour driving at average speed of 65 miles/h) for various shopping needs including new cars, furniture, electronics and appliances, clothing products, department stores, discount department stores, and warehouse clubs and supercenters. It is possible that a consumer's out-of-town shopping involves multiple lines of retail trade, rather than one single retail line. Recall consumer's shopping benefits and shopping costs discussed in Chapter 3 that consumers would minimize their economic and temporal costs while shopping. Thus, it is both economically and temporally beneficial for consumers to plan a comprehensive out-of-town shopping trip that maximizes their shopping utilities. For retailers who are looking to open new stores, it is recommended to research not only the consumer and competitive environments inside the focal area, but also the external environment in the surrounding areas. This applies to retailers looking to open new stores in both metropolitan areas and micropolitan areas.

In summary, this dissertation provides insights to retailers as they make decisions on store location, store size, strategies against competition, market segmentation, and the use of marketing mix variables.

Chapter 8 Conclusions

This chapter concludes this dissertation and provides avenues for future research.

8.1 Conclusions

In this dissertation, I have presented an examination of the U.S. retail market structure across seven retail subsectors (*i.e.*, nine retail lines), using the 2012 Economic Census and 2012 American Community Survey (5-year estimates). These seven retail subsectors together generate 73.68% of the total retail sales revenue in 2012. This dissertation covers 735 metropolitan and micropolitan areas, among which 354 are metropolitan areas and 381 are micropolitan areas. I have addressed four research questions in this study.

First, I analyze the impact of intertype and intratype competition on household expenditures. I show that both intertype and intratype competition take sales away from the focal industry. Warehouse clubs and supercenters are strong competitors for many lines of retail trade including furniture stores, electronics and appliance stores, supermarkets, clothing stores, traditional department stores, and discount department stores. Furthermore, there may exist agglomeration effect for intertype competitors of clothing stores and department stores in metropolitan areas.

Second, I show that a set of socioeconomic and demographic variables (*i.e.*, consumer groups), marketing mix variables (*i.e.*, quantity and quality of service), competitive variables (*i.e.*, intertype and intratype competition), and consumer out-shopping variables (*i.e.*, population from nearby areas) influence household expenditures across various retail lines. Specifically, the two marketing mix variables significantly increase household expenditures at all nine retail lines except for quantity of service at traditional department stores. Furthermore, with adequate and quality service, local retailers are able to lessen the negative impact of competition from local rivals and prevent consumers from out-of-town shopping, to some extent.

Third, for each retail subsector/industry, I compare the results of metropolitan areas versus those of micropolitan areas. I show that for some retail subsectors/industries, the key determinants are slightly different in metropolitan areas versus micropolitan areas.

Fourth, I examine the phenomenon of consumer out-of-town shopping. In this dissertation, I show that consumers do out-of-town shopping for seven retail lines. They do so by traveling to nearby large areas within 200 miles of their driving range.

Lastly, I show that the two marketing mix variables – quantity and quality of service indeed mediate the effects of consumer groups on their household expenditures, across several retail lines. In other words, the five consumer groups (*i.e.*, consumer segments) all respond to quantity and or quality of service, but differently, which indicates that in general consumers are willing to pay extra for services provided by local retailers.

8.2 Future Research Directions

This section presents the limitations of this dissertation and provides avenues for future research.

First, since this dissertation uses only cross-sectional data from the 2012 Economic Census, one can build on this dissertation and examine expenditure patterns over time using Economic Censuses prior to 2012. However, one should note that Economic Census is conducted every five years, and information regarding micropolitan areas were not available for census years prior to 2012. Therefore, using past censuses will only ensure researchers to study expenditure patterns in metropolitan areas. When the 2017 Economic Census becomes available, one can compare the data in 2012 with 2017, and examine the changes in expenditure patterns occurred during those years.

Second, each retail subsector or industry is examined independently in this dissertation. One can build on this research to simultaneously examine all the competing subsectors/industries. In this dissertation, I treat intertype competition and intratype competition as exogenous, due to data limitation. However, one should note that the number of competitor stores is endogenously determined since their existence in a geographic area should be managerially optimized in response to the local socioeconomic and demographic characteristics and competitive forces.

Third, this dissertation only covers brick-and-mortar retail businesses. It will be interesting to examine how e-shopping has impacted brick-and-mortar businesses, if data is available.

Fourth, as discussed in Chen et al. (2020), an important factor that is influential to retail businesses is the rate of population change. Since every brick-and-mortar store requires a minimum population in the geographic area to cover its fixed cost, if a decline of population is expected in the area, retailers will have to either shrink their store size to lower fixed cost, or seek other ways to create demand, perhaps establishing an online presence. Many retailers are adopting a hybrid business format of selling both locally and online. This will be an interesting research topic to be explored in the future.

Fifth, I only examine seven of the twelve retail subsectors in this dissertation, due to their importance in the retail sector and relevance to consumer out-shopping. One can examine other lines of retail trade which are not covered in this dissertation.

Sixth, the consumer out-shopping variables are constructed using population from nearby areas. This may not be an optimal measure for the phenomenon of consumer out-of-town shopping, but it is the optimal solution at hand. One can develop and examine how to measure this phenomenon more accurately, as this phenomenon significantly influences local retail sales.

Lastly, I use the U.S. census data to examine its retail market structure. One can build upon the theoretical framework developed in this dissertation and extend it to other economic and cultural contexts.

In summary, there are many opportunities for researchers to extend this study, and I have provided seven possible avenues for future research: investigating expenditure patterns across a time horizon, simultaneously examining the competitive industries, considering the impact of online shopping, addressing the concern of population decline, examining other lines of retail trade, refining the measure for consumer out-shopping, and extending this study to other contexts.

References

Becker, G. (1965). A Theory of the Allocation of Time. *The Economic Journal*, 75(299), 493-517.

Bruce, G. D. (1969). "The ecological structure of retail institutions." *Journal of Marketing Research*, 48-53.

Bucklin, L. P. (1967). "The concept of mass in intra-urban shopping." *Journal of Marketing*, 31(4), 37-42.

Bucklin, L. P. (1972). *Competition and evolution in the distributive trades*. Englewood Cliffs, NJ: Prentice-Hall.

Chen, Si, C. A. Ingene, & I. Takahashi. (2020). "Household Expenditures at Japanese Food and Beverage Retailers." *Journal of Business Research*, 107, 245-255. https://doi.org/10.1016/j.jbusres.2018.08.022.

Christaller, W. (1933). *Central Places in Southern Germany*. *Translated by C. W. Baskins*. Englewood Cliffs, NJ, Prentice-Hall, 1966.

Converse, P. D. (1949). "New laws of retail gravitation." *Journal of marketing*, *14*(3), 379-384.

Cronin Jr, J Joseph, M. K. Brady, & G Tomas M Hult. (2000). "Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments." *Journal of Retailing*, *76*(2), 193-218.

Ferber, R. (1958). "Variations in retail sales between cities." *Journal of Marketing*, 295-303.

Forbes, J. D. (1972). "Central place theory: An analytical framework for retail structure." *Land Economics*, *48*(1), 15-22.

Ford, P. (1935). "Excessive competition in the retail trades. Changes in the numbers of shops, 1901-1931." *The Economic Journal*, *45*(179), 501-508.

Gardner, J., & J. R. Hendrickson. (2018). "If I Leave Here Tomorrow: An Option View of Migration When Labor Market Quality Declines." *Southern Economic Journal*, *84*(3), 786-814.

Google Map Developers. https://www.mapdevelopers.com/batch_geocode_tool.php/, Accessed October 20, 2018

Hall, M., J. Knapp, & C. Winsten. (1961). *Distribution in Great Britain and North America: A study in structure and productivity*. London: Oxford University Press.

Herrmann, R. O., & L. L. Beik. (1968). "Shoppers' movements outside their local retail area." *Journal of Marketing*, 45-51.

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* Guilford publications.

Holian, M. J., & M. E. Kahn. (2015). "Household demand for low carbon policies: Evidence from California." *Journal of the Association of Environmental and Resource Economists*, *2*(2), 205-234.

Huff, D. L. (1964). "Defining and estimating a trading area." *Journal of Marketing*, 34-38.

Ingene, C. A. (1983). "Intertype competition: restaurants versus grocery stores." *Journal of Retailing*, *59*(3), 49-75.

Ingene, C. A. (1984). "Structural determinants of market potential." *Journal of Retailing*, *60*(1), 37-64.

Ingene, C. A., & A. Ghosh. (1990). "Consumer and producer behavior in a multipurpose shopping environment." *Geographical Analysis*, 22(1), 70-93.

Ingene, C. A., & Robert F. Lusch. (1980). "Market selection decisions for department stores." *Journal of Retailing*, *56*(3), 21-40.

Ingene, C. A., & Robert F. Lusch. (1981). "A model of retail structure." *Research in marketing*, *5*(2), 101-164.

Ingene, C. A., & E. Yu. (1981). "Determinants of retail sales in SMSAs." *Regional Science and Urban Economics*, *11* (4), 529-547.

Ingene, C. A., & James R. Brown. (1987). "The structure of gasoline retailing." *Journal of Retailing*, *63*(4), 365.

Kaiser, H. F. (1974). "An index of factorial simplicity." *Psychometrika*, *39*(1), 31-36.

Liu, B. C. (1970). "Determinants of retail sales in large metropolitan areas, 1954 and 1963." *Journal of the American Statistical Association*, *65*(332), 1460-1473.

Miller, C. E., J. Reardon, & D. E. McCorkle. (1999). "The effects of competition on retail structure: An examination of intratype, intertype, and intercategory competition." *Journal of Marketing*, 107-120.

North American Industry Classification System: NAICS. https://www.naics.com/six-digit-naics/?code=44-45/, Accessed February 15, 2019.

Office of Management and Budget. OMB Bulletin No. 18-04. https://www.whitehouse.gov/wp-content/uploads/2018/09/Bulletin-18-04.pdf/, Accessed January 02, 2019.

Reidenbach, R. E., M. B. Cooper, & M. C. Harrison. (1984). “A factor analytic comparison of outshopping behavior in larger retail trade areas.” *Journal of the Academy of Marketing Science*, *12*(1-2), 145-158.

Reilly, W. (1931). *The law of retail gravitation*. New York: Knickerbocker Press.

Reynolds, F. D., & W. R. Darden. (1972). “Intermarket patronage: A psychographic study of consumer outshoppers.” *Journal of Marketing*, 50-54.

Reynolds, F. D., & W. S. Martin. (1974). “A multivariate analysis of intermarket patronage: some empirical findings.” *Journal of Business Research*, *2*(2), 193-200.

Reynolds, R. (1953). “A Test of the Law of Retail Gravitation.” *Journal of Marketing*, *17*(3), 273-277.

Russell, V. K. (1957). “The relationship between income and retail sales in local areas.” *Journal of Marketing*, *21*(3), 329-332.

Schwartzman, D. (1971). *The Decline of Service in Retail Trade: An Analysis of the Growth of Sales Per Manhour, 1929-1963*. Bureau of Economic and Business Research, Washington State University.

Shah, A. (2018). *ASDOC: Stata module to create high-quality tables in MS Word from Stata output. Statistical Software Components S458466*, Boston College Department of Economics.

Takeuchi, H., & L. P. Bucklin. (1977). “Productivity in Retailing: Retail Structure and Public Policy.” *Journal of Retailing*, *53*(1), 35.

Tauber, E. M. (1972). “Marketing notes and communications: why do people shop?.” *Journal of Marketing*, 36(4), 46-49.

Teas, R Kenneth. (1993). “Expectations, performance evaluation, and consumers’ perceptions of quality”. *Journal of Marketing*, 18-34.

Thompson, J. R. (1971). “Characteristics and behavior of out-shopping consumers.” *Journal of Retailing*, *47*(1), 70-80.

U.S. Bureau of Labor Statistics, All Employees: Retail Trade [CEU4200000001], retrieved from FRED, Federal Reserve Bank of St. Louis. https://fred.stlouisfed.org/series/CEU4200000001/, Accessed April 02, 2019.

U.S. Bureau of Labor Statistics, Consumer Expenditure Survey 2017, https://www.bls.gov/news.release/cesan.nr0.htm/, Accessed January 23, 2019.

U.S. Bureau of Labor Statistics, Current Population Survey 2012, https://www.bls.gov/cps/aa2012/cpsaat17.pdf/, Accessed May 2, 2019.

U.S. Census Bureau, American Community Survey, https://www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html/, Accessed March 12, 2019

U.S. Census Bureau. Annual Retail Trade Survey 2017. https://www.census.gov/data/tables/2017/econ/arts/annual-report.html/, Accessed January 02, 2019.

U.S. Census Bureau, Historical Delineation Files, Counties with Metropolitan and Micropolitan Statistical Area Codes, Dec. 2009, https://www.census.gov/geographies/reference-files/time-series/demo/metro-micro/historical-delineation-files.html/, Accessed October 16, 2018

U.S. Census Bureau, Historical Delineation Files, Principal Cities of Metropolitan and Micropolitan Statistical Areas, Dec. 2009, https://www2.census.gov/programs-surveys/metro-micro/geographies/reference-files/2009/historical-delineation-files/list2.txt/, Accessed October 16, 2018

U.S. Census Bureau, Gazetteer Files for Counties, 2012, https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.2012.html/, Accessed October 16, 2018

U.S. Census Bureau. Monthly Retail Trade Report. https://www.census.gov/retail/index.html/, Accessed April 02, 2019.

U.S. Census Bureau. 2012 Economic Census. https://www.census.gov/programs-surveys/economic-census/year.html/, Accessed January 02, 2019.

Van Tassel, C. (1965). *Analysis of factors influencing retail sales.* East Lansing, MI: Bureau of Business and Economic Research, Michigan State University.

Weber, S., & M. Péclat. (2017). “A simple command to calculate travel distance and travel time.” *The Stata Journal*, *17*(4), 962-971.

The World Bank, U.S. GDP Growth. https://data.worldbank.org/indicator/NY.GDP.MKTP.KD.ZG?end=2017&locations=US&start=1992/, Accessed April 15,2019.

Appendix A: 2012 Economic Census – 18 NAICS Industrial Sectors

The 2012 Economic Census contains information on the following 18 industrial sectors classified using NAICS.

NAICS 21 – Mining, quarrying, and oil and gas extraction

NAICS 22 – Utilities

NAICS 23 – Construction

NAICS 31-33 – Manufacturing

NAICS 42 – Wholesale trade

NAICS 44-45 – Retail trade

NAICS 48-49 – Transportation and warehousing

NAICS 51 – Information

NAICS 52 – Finance and insurance

NAICS 53 – Real estate and rental and leasing

NAICS 54 – Professional, scientific, and technical services

NAICS 55 – Management of companies and enterprises

NAICS 56 – Administrative and support and waste management and remediation services

NAICS 61 – Educational services

NAICS 62 – Health care and social assistance

NAICS 71 – Arts, entertainment, and recreation

NAICS 72 – Accommodation and food services

NAICS 81 – Other services (except public administration)

Appendix B: Estimated Number of Employees, Estimated Annual Salary, and Relative Salary to the Retail Trade Sector

Estimated # of Employees in retail subsector/industry i (i = 1, 2, …, 9) is calculated as follows.

Step 1:

Average quarterly payroll$_i$ = $\frac{Annual\ payroll_i}{4}$

Step 2:

Estimated # of employees$_i$ = $\frac{Average\ quarterly\ payroll_i}{First\text{−}quarter\ payroll_i}$ * First-quarter # of employees$_i$

I do so for all the nine retail subsector/industry examined in this dissertation and for the entire retail trade sector.

Estimated annual salary is calculated as follows.

Estimated annual salary$_i$ = $\frac{Annual\ payroll_i}{Estimated\ \#\ of\ Employees\ _i}$

I do so for all the seven retail subsectors examined in this dissertation and for the entire retail trade sector.

Relative salary is calculated as follows.

Relative salary$_i$ = $\frac{Estimated\ annual\ salary_i}{Estimtaed\ annual\ salary\ of\ the\ retail\ trade\ sector}$

Appendix C: A List of 182 Metropolitan & Micropolitan Areas Not Included in This Dissertation

Geographic Area	**# of Retail Stores**	**Total Retail Sales ($M)**	**1 Quarter # of Retail Employees**
Altus, OK Micro Area	99	358	1302
Americus, GA Micro Area	151	330	1490
Andrews, TX Micro Area	30	170	392
Arkadelphia, AR Micro Area	96	274	1206
Arkansas City-Winfield, KS Micro Area	121	337	1472
Atchison, KS Micro Area	49	127	669
Bainbridge, GA Micro Area	136	325	1344
Bardstown, KY Micro Area	159	535	1860
Bastrop, LA Micro Area	85	237	1007
Beatrice, NE Micro Area	112	262	1063
Beeville, TX Micro Area	78	331	1113
Bennettsville, SC Micro Area	98	185	768
Big Spring, TX Micro Area	109	404	1353
Big Stone Gap, VA Micro Area	260	823	3411
Bloomington, IL Metro Area	632	2662	10134
Boone, IA Micro Area	69	217	1034
Borger, TX Micro Area	79	205	915
Breckenridge, CO Micro Area	346	627	3003
Brookings, SD Micro Area	121	376	1777
California-Lexington Park, MD Metro Area	294	1227	4950
Campbellsville, KY Micro Area	126	342	1443
Cedar City, UT Micro Area	172	568	1940
Clarksdale, MS Micro Area	117	295	1040
Cleveland, MS Micro Area	161	371	1567
Clewiston, FL Micro Area	101	258	1000
Columbus, NE Micro Area	161	541	2156
Connersville, IN Micro Area	75	234	1091
Cordele, GA Micro Area	119	304	1299
Craig, CO Micro Area	72	202	758
Crescent City, CA Micro Area	60	193	998
Cullowhee, NC Micro Area	157	404	1650
Dayton, TN Micro Area	107	281	1162

Geographic Area	# of Retail Stores	Total Retail Sales ($M)	1 Quarter # of Retail Employees
Decatur, IN Micro Area	133	377	1518
Deming, NM Micro Area	74	210	985
DeRidder, LA Micro Area	104	384	1375
Dickinson, ND Micro Area	160	883	2372
Dodge City, KS Micro Area	126	499	1881
Dumas, TX Micro Area	75	276	991
Edwards, CO Micro Area	455	806	3545
Effingham, IL Micro Area	203	935	3006
Elk City, OK Micro Area	141	715	1730
Elkins, WV Micro Area	142	353	1491
Emporia, KS Micro Area	162	494	1817
Evanston, WY Micro Area	92	390	1160
Fairfield, IA Micro Area	86	172	997
Fallon, NV Micro Area	69	235	986
Fitzgerald, GA Micro Area	78	200	734
Forrest City, AR Micro Area	113	316	1173
Fort Leonard Wood, MO Micro Area	138	455	1765
Fort Morgan, CO Micro Area	92	257	984
Frankfort, IN Micro Area	102	239	1022
Garden City, KS Micro Area	183	589	2503
Gillette, WY Micro Area	184	896	2566
Glenwood Springs, CO Micro Area	520	1294	4527
Grants, NM Micro Area	70	257	864
Great Bend, KS Micro Area	128	420	1659
Greenfield Town, MA Micro Area	262	704	2960
Greensburg, IN Micro Area	110	357	1257
Grenada, MS Micro Area	120	415	1382
Guymon, OK Micro Area	76	226	849
Hailey, ID Micro Area	198	338	1549
Hays, KS Micro Area	183	621	2188
Heber, UT Micro Area	88	267	1119
Helena-West Helena, AR Micro Area	85	223	965
Hereford, TX Micro Area	66	276	783
Hermiston-Pendleton, OR Micro Area	237	822	3113
Hillsdale, MI Micro Area	139	370	1348
Holland, MI Micro Area	335	950	3587
Hood River, OR Micro Area	157	315	1367

Geographic Area	# of Retail Stores	Total Retail Sales ($M)	1 Quarter # of Retail Employees
Huron, SD Micro Area	79	258	1094
Indianola, MS Micro Area	100	244	951
Ionia, MI Micro Area	143	430	1601
Jackson, OH Micro Area	117	340	1386
Jackson, WY-ID Micro Area	279	579	2394
Jamestown, ND Micro Area	101	379	1394
Jefferson, GA Micro Area	232	1483	2788
Jesup, GA Micro Area	120	293	1211
Junction City, KS Micro Area	96	346	1262
Juneau, AK Micro Area	142	491	1996
Ketchikan, AK Micro Area	121	236	974
Kingsville, TX Micro Area	110	524	1488
Kirksville, MO Micro Area	145	385	1749
Lafayette-West Lafayette, IN Metro Area	611	2379	9428
La Grande, OR Micro Area	103	318	1408
Lamesa, TX Micro Area	43	317	683
Laramie, WY Micro Area	141	480	1839
Las Vegas, NM Micro Area	81	221	961
Levelland, TX Micro Area	61	221	900
Lewisburg, TN Micro Area	102	289	1024
Lexington, NE Micro Area	112	427	1532
Liberal, KS Micro Area	91	317	1222
Lincoln, IL Micro Area	88	292	1087
Logan, WV Micro Area	118	551	1808
Los Alamos, NM Micro Area	28	103	466
Los Angeles-Long Beach-Anaheim, CA Metro Area	37817	166583	542118
Ludington, MI Micro Area	113	370	1547
Macomb, IL Micro Area	121	341	1670
Madison, IN Micro Area	128	391	1566
Magnolia, AR Micro Area	103	197	1028
Malvern, AR Micro Area	89	254	946
Marietta, OH Micro Area	227	756	2861
Marion, NC Micro Area	125	458	1706
Marshall, MN Micro Area	134	459	2108
Marshall, MO Micro Area	97	253	1046
Maryville, MO Micro Area	72	260	1173

Geographic Area	# of Retail Stores	Total Retail Sales ($M)	1 Quarter # of Retail Employees
Maysville, KY Micro Area	93	364	1339
Mexico, MO Micro Area	102	254	1144
Middlesborough, KY Micro Area	128	365	1447
Mineral Wells, TX Micro Area	118	282	1113
Mitchell, SD Micro Area	134	537	2082
Moberly, MO Micro Area	97	297	1171
Moscow, ID Micro Area	139	356	1940
Mountain Home, ID Micro Area	76	256	1000
Mount Pleasant, TX Micro Area	133	478	1796
Newport, OR Micro Area	307	581	2917
New Ulm, MN Micro Area	109	347	1692
North Vernon, IN Micro Area	66	197	777
Oskaloosa, IA Micro Area	103	243	1251
Othello, WA Micro Area	51	158	568
Ottawa, KS Micro Area	90	271	1069
Oxford, MS Micro Area	211	549	2591
Oxford, NC Micro Area	132	350	1372
Ozark, AL Micro Area	138	323	1216
Pampa, TX Micro Area	100	315	1136
Parsons, KS Micro Area	89	199	971
Pecos, TX Micro Area	29	162	448
Pierre, SD Micro Area	131	443	1690
Pinehurst-Southern Pines, NC Micro Area	352	1095	4558
Plainview, TX Micro Area	114	347	1489
Portales, NM Micro Area	52	157	687
Port Clinton, OH Micro Area	144	428	1581
Port Lavaca, TX Micro Area	63	435	864
Price, UT Micro Area	86	311	1117
Prineville, OR Micro Area	70	184	630
Pullman, WA Micro Area	100	334	1329
Raymondville, TX Micro Area	32	118	422
Rexburg, ID Micro Area	151	444	1972
Rock Springs, WY Micro Area	199	921	2622
Salem, OH Micro Area	338	1126	4123
Sandpoint, ID Micro Area	199	438	2138

Geographic Area	# of Retail Stores	Total Retail Sales ($M)	1 Quarter # of Retail Employees
Santa Maria-Santa Barbara, CA Metro Area	1509	4854	19295
Shawano, WI Micro Area	127	404	1576
Sheridan, WY Micro Area	151	476	1606
Snyder, TX Micro Area	63	242	760
Sonora, CA Micro Area	173	573	2264
Spearfish, SD Micro Area	142	434	1456
Spencer, IA Micro Area	111	313	1496
Spirit Lake, IA Micro Area	111	293	1225
Starkville, MS Micro Area	163	472	1927
Steamboat Springs, CO Micro Area	215	361	1585
Stephenville, TX Micro Area	172	535	1918
Sterling, CO Micro Area	100	320	1173
Storm Lake, IA Micro Area	94	281	1147
Summerville, GA Micro Area	74	158	676
Summit Park, UT Micro Area	295	1014	4168
Susanville, CA Micro Area	79	216	810
Sweetwater, TX Micro Area	53	210	707
The Dalles, OR Micro Area	123	389	1425
Thomaston, GA Micro Area	93	208	974
Toccoa, GA Micro Area	99	263	1152
Troy, AL Micro Area	143	374	1496
Urban Honolulu, HI Metro Area	2889	13036	46865
Uvalde, TX Micro Area	105	424	1283
Van Wert, OH Micro Area	87	283	1297
Vermillion, SD Micro Area	45	123	717
Vernal, UT Micro Area	135	503	1706
Vernon, TX Micro Area	48	255	703
Vineyard Haven, MA Micro Area	205	374	1496
Wahpeton, ND-MN Micro Area	101	346	1230
Washington, IN Micro Area	122	486	1505
Washington Court House, OH Micro Area	183	675	2444
Watertown, SD Micro Area	184	649	2748
Wauchula, FL Micro Area	76	172	653
Weatherford, OK Micro Area	149	531	1731

Geographic Area	# of Retail Stores	Total Retail Sales ($M)	1 Quarter # of Retail Employees
Weirton-Steubenville, WV-OH Metro Area	372	1195	5221
Williston, ND Micro Area	114	798	2260
Winnemucca, NV Micro Area	78	325	1017
Wisconsin Rapids-Marshfield, WI Micro Area	279	1102	5073
Woodward, OK Micro Area	114	435	1324
Worthington, MN Micro Area	111	328	1533
Yankton, SD Micro Area	121	369	1744
Zapata, TX Micro Area	32	76	283