A Comprehensive Analysis of Travel Mode Choices for the Urban Grocery Shopping Trip: A Case Study of Salt Lake County, Utah

A Thesis

Presented in Partial Fulfilment of the Requirements for the

Degree of Master of Science

with a

Major in Geography

in the

College of Graduate Studies

University of Idaho

by

Ensheng Dong

Major Professor: Haifeng (Felix) Liao, Ph.D.

Committee Members: Karen S. Humes, Ph.D.; Hejun Kang, Ph.D.

Department Administrator: Karen S. Humes, Ph.D.

May, 2015

Authorization to Submit Thesis

This thesis of Ensheng Dong, submitted for the degree of Master of Science with a major in Geography and titled "A Comprehensive Analysis of Travel Mode Choices for the Urban Grocery Shopping Trip: A Case Study of Salt Lake County, Utah," has been reviewed in final form. Permission, as indicated by the signatures and dates given below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor ____________________________ Date ________________

Haifeng (Felix) Liao, Ph.D.

Committee

Members ____________________________ Date ________________

Karen S. Humes, Ph.D.

____________________________ Date ________________

Hejun Kang, Ph.D.

Department

Administrator ____________________________ Date ________________

Karen S. Humes, Ph.D.

Abstract

This research comprehensively investigates influential factors which govern people's choices concerning which travel mode they use for shopping trips. The data about 1,294 home-based shopping trips was collected from the 2012 Utah Travel Survey. 61.75% of the respondents drove alone to their grocery stores, whereas 32.46% selected carpooling and the rest chose other modes. A descriptive analysis and multinomial logit models (MNL) were conducted to estimate the impact of individual socioeconomic characteristics, accessibility variables, and household and store built-environments on mode choices during shopping trips. The characteristics of built-environments in the neighborhoods around households and stores were measured while taking spatial scales into account. Results show that strong predictors of walking or riding public transit to the grocery store were age, household composition (the number of household members), vehicle ownership, household annual income, land use mix, street density, and distance to the central business district. Results also suggest that different factors have different optimal geographical scales for studying shopping trip travel modes. The straight-line buffer is good for the sales amount, the network buffer is suitable for the household built-environment factors, and the census block and block group scales apt for the store built-environment factors. Future studies of the travel mode choice for urban grocery shopping trips should consider the effect of spatial scales on built-environment factors and also additional store characteristics, such as the number of employees, the number of parking lots, and the variety of shops.

Acknowledgements

I would never have been able to finish my thesis without the guidance of my committee members, help from friends, and support from my family.

First and foremost, I would like to express my deepest and most sincere gratitude to my major advisor, Dr. Felix Liao, who came up with the topic of this thesis, negotiated with his previous colleagues, and managed to help me complete this research. His correction of my writing with patience and formative ideas did help me a lot. I would also like to thank Dr. Karen S. Humes, the Department Chair, for her providing me with an excellent atmosphere and necessary funding for doing research. She was always willing to help and give her best suggestions.

I would also like to thank Dr. Hejun Kang, for her guidance and caring. Dr. Kang kindly provided me with good ideas and gave me the chance to enhance my research ability with financial support. She also warmly encouraged me to take advantage of the opportunity to present in the 2013 AAG conference in Los Angeles, CA.

Special thanks goes to Dr. Tim G. Frazier and his professional group. Dr. Frazier was the first to cheer me on to continue doing my research with his generous and warm heart. He and his students also offered me many opportunities: figuring out the topic, seeking another degree, attending his group meetings, etc.

Many thanks to Dr. Raymond J. Dezzani, for his research attitude and classes set examples in my academic life. Also many thanks to Michele Vachon, Kevin Henry, Brendan Gordon, Laura Courtox, Li Huang, and other friends in the Department of Geography for helping and encouraging me in different stages of my thesis. My research would not have been possible without their help.

Finally, I would like to thank my family. They are always supporting me and encouraging me with their best wishes.

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Background and Research Questions

The questions about "why," "when," "where" and "how" to make daily trips reflect people's travel demands, and questions also interest geographers, transportation engineers and urban planners (Chen, Gong, & Paaswell, 2008; Guo, Bhat, & Copperman, 2007). According to the 2009 National Household Travel Survey (NHTS), there are several basic categories of people's daily trip purposes: to/from work, work-related business, shopping, other family/personal errands, school/church, social and recreational trips and other purposes. Among them, there is a continuously increasing tendency of shopping trips. For instance, the average annual personal miles of travel (PMT) of shopping in the U.S. increased from 2,567 miles in 1983 to 4,620 miles in 2009; the average number of annual personal trips per household for shopping increased from 474 in 1983 to 725 in 2009; and the length of average personal shopping trip increased from 5.4 miles in 1983 to 6.5 miles in 2009 (Santos, McGuckin, Nakamoto, Gray, & Liss, 2011). Obviously, the travel behavior study is related to not only a specific research field but also people's daily life. Furthermore, the increasing amount of non-work trips, or the shopping trips particularly, also contributes to travel safety, traffic congestion, greenhouse gas emission, photochemical smog, etc. (Finkelstein, Fiebelkorn, & Wang, 2004; Künzli et al., 2000).

Traditionally trip mode choice studies have primarily focused on the work trip (Bhat, 1998b; Horowitz, 1993). Non-work trip has recently become another major fraction in this field, such as customers' shopping trips (McDonald, 2008; Müller, Tscharaktschiew, & Haase, 2008; Zhu & Lee, 2008) and children's school trips (Lee, Zhu, Yoon, & Varni, 2013; McDonald, Deakin, & Aalborg, 2010). Grocery shopping trip have attracted interdisciplinary attention especially in the fields of geography, transportation, urban planning, economics, and marketing. Most of these fields mainly investigate variables related to individual and household characteristics, shopping features, and built-environment (Cervero, 1996). Our research tries to delve more into understanding these factors with a particular focus on how demographic and built-environment factors affect travel mode choices.

Driving a private vehicle, walking, riding a bike and taking public transit are major surface transportation modes for people's daily life. As for the literature on the shopping trip travel behavior, the travel mode is usually categorized as auto and non-auto (Jiao, Moudon, & Drewnowski, 2012). Some scholars also focus on only one travel mode. For instance, Boivin (2008) analyzed the bus travel mode; Rodrigues, Popovich, & Handy (2014) studied bicycles, and Cervero (1996) and Hess & Moudon (1999) talked about walking. On the other hand, the geographical scale is seldom studied comprehensively. Geographical scale or unit can influence the relationship between neighborhood context and people's travel behavior (Handy, Boarnet, Ewing, & Killingsworth, 2002), as a consequence, choosing a proper scale is always challenging, because of uncertain factors, such as the location of place, neighborhood factors, the study purpose, as well as the limitation of data source (Fan et al., 2014; Yamada et al., 2012).

In this thesis, we are trying to answer questions: Why do people prefer a certain travel mode among driving alone, carpool and others? To what extent do built-environment and store characteristics play a role in shaping mode choices? Does the geographical scale influence people's decision?

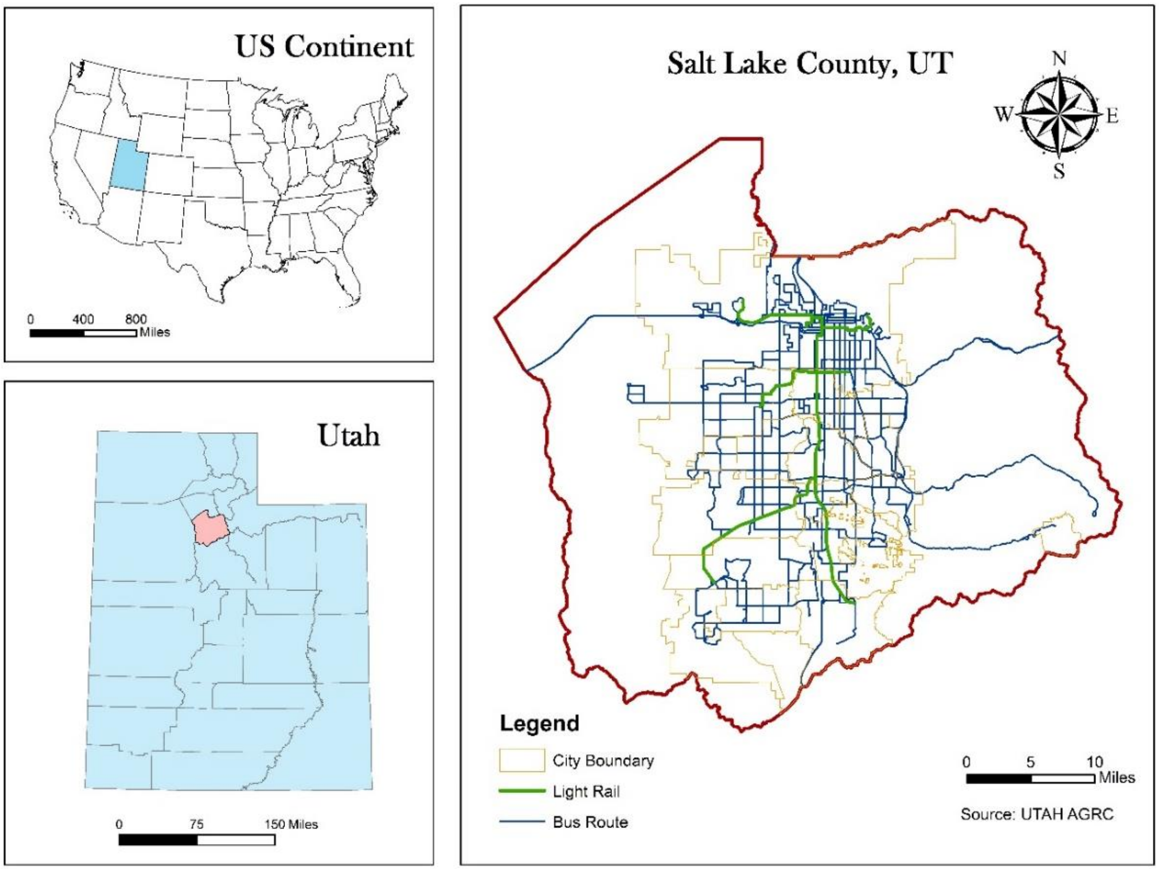


Figure 1.1: Study area

1.2 Study Area

Like most Americans, Utahans rely heavily on automobiles for transportation (NCHRP, 1988). But there has been an emerging transformation in Utah's transportation system, especially after the successful 2002 Winter Olympic Games, facilitating a more public-transit oriented urban area (Farber, Bartholomew, Li, Páez, & Nurul Habib, 2014). The extensive light rail (TRAX) and commuter rail system (FrontRunner) in the Salt Lake metropolitan area also result in the fact that more and more residents in Wasatch Front become supporters of public transit (Figure1.1). Due to this variety of transportation modes, Salt Lake County, Utah becomes the study area of this research.

Relevant stakeholders, including the Utah Transit Authority (UTA), the Wasatch Front Regional Council (WFRC), the Utah Department of Transportation (UDOT) as well as other organizations, have been cooperating to prepare a long-range transportation plan. For example, the Wasatch Choice 2040 was initiated in 2004 to find a more effective approach to smart growth, transportation, and land use planning coordination in Weber, Davis, Salt Lake, and Utah counties (Scheer, 2012).

As the population and the vehicle miles traveled (VMT) increased, providing multimodal travel options has been an important policy to meet people's future travel demands. For example, policy makers have been putting more effort on the expansion of the road/bus/light rail system, the construction of the commuter rail, the establishment of the bus rapid system and the bicycle and pedestrian network, the policy guidance on the carpool travel, etc. (Utah 2011 - 2040 Long Range Transportation Plan (LRP), and the Preferred Scenario for the 2015-2040 Wasatch Regional Transportation Plan (RTP)). Nevertheless, a comprehensive understanding of the most influential factor affecting residents' trip mode choices in this area is still lacking. Therefore, results of this study will also provide most recent evidence to support a smart transportation solution from the perspective of shopping trips. This will help local policy makers fulfill the transportation strategic goals in providing interconnected transportation choices, enhancing the regional economy, strengthening the senses of community, and protecting the environment, etc.

1.3 Research Objectives and Thesis Organization

The objectives of this thesis are threefold. First, it integrates the individual demographic and household socioeconomic background, built-environment, and characteristics of grocery stores in the models of travel mode choices for urban grocery shopping trips. A descriptive analysis was conducted and a multinomial logit model (MNL) was developed. Second, it applies a multi-scale approach to more accurately gauge the impacts of built-environment factors on shopping travel mode choices. Third, some policy recommendations are made for designing healthy and environmental-friendly communities.

The basic elements in this research are shown in the flow-chart in Figure 1.2: three (3) shopping trip travel modes, and four (4) levels of influential factors with seven (7) different geographical scale. A descriptive analysis and multinomial logit modal were developed to find factors affecting people's shopping trip travel mode choices.

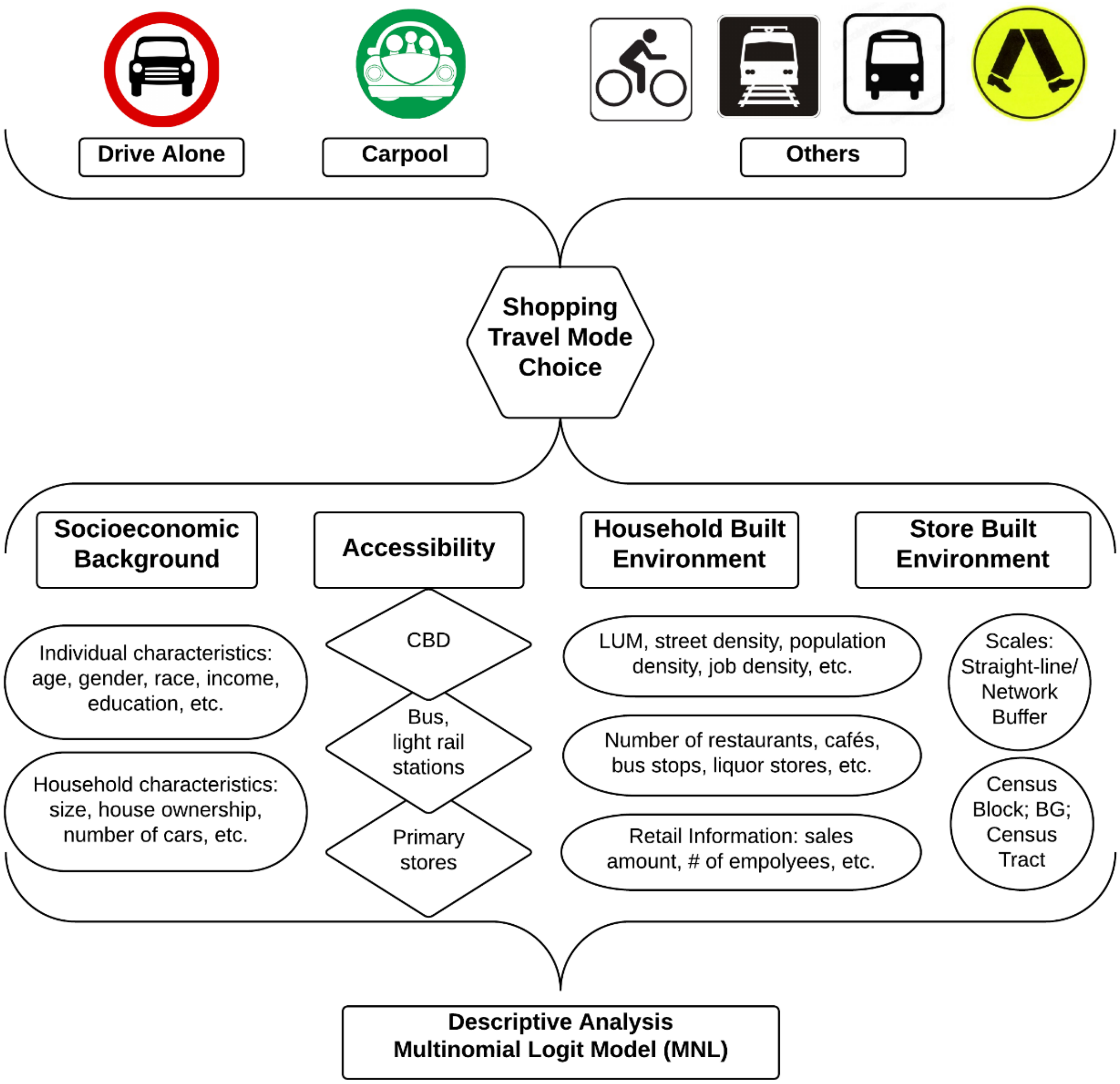


Figure 1.2: The flow-chart of analysis

This thesis is organized in five sections. After this introduction, the second chapter will briefly summarize the literature on factors affecting travel mode choice. This is followed by the explanation of the methodology and data acquisition. Then the model results are discussed based on a case in Salt Lake County, UT. The major findings are summarized and suggestions for future studies are also discussed in the conclusion chapter.

Chapter 2 Literature Review

There have been a large body of literature on travel behavior and travel mode choices. Some of them focus on work trip travel behavior (Elmi, Badoe, & Miller, 2015; Horowitz, 1993; Liu, 2007), some address shopping trip travel frequency (Blaylock, 1989; Cao & Mokhtarian, 2005; Smith & Carsky, 1996; Wilde & Ranney, 2000), but the number of studies about grocery shopping trip travel mode choices and its research method is limited (Jiao et al., 2012). This chapter first reviews literatures on factors affecting mode choices of shopping trips, which include individual and household background factors, shopping characteristics and built-environment factors, and then discusses methods that have been used in analyzing travel mode choices such as the multinomial logit model.

2.1 Factors Affecting Mode Choices of Shopping Trips

2.1.1 Individual Demographic Characteristics

Individual backgrounds, including age, gender, race, education, personal income, employment status, etc., were found as having an impact on people's shopping trip travel behavior. Su & Bell (2009) mentioned that older people would like to choose a cheaper travel mode, and they care less about travel times. Schmöcker, Quddus, Noland, & Bell (2008) found that disabled older people have a preference for taxis. Based on 1,000 surveys in late May 1995 in Austin, Texas, Handy (1996) found that younger women were more likely to walk, which showed that the age as well as the gender might be an important correlated factor. In contrast, Guo et al. (2007) reported that people older than 65 years are more likely to choose an automobile for shopping. In addition, the older the shopper, the more likely their shopping frequency is more than once a week and they may use different travel modes, since the elderly may have a greater desire for fresh food than the younger people (Blaylock, 1989).

Gender is always a major issue in analyzing shopping behavior, and researchers have come to different conclusions about the influence of gender in shopping trips. The topic of gender and transportation was first introduced around 1980 (Law, 1999; Rosenbloom, 1978; Scheiner & Holz-Rau, 2012). If a female is the household head, then she is likely to shop once a week or more with different kinds of travel modes, which may be due to women having to take care of other household responsibilities, such as child care (Blaylock, 1989). A more recent record from the 2009 National Household Travel Survey (NHTS) shows that women, overall, took more trips than men for family errands, including shopping. Other researchers also found that being a female plays an important role in deciding whether to use auto or non-auto travel modes (Bhat, 1998b; Handy, 1996; Matthies, Kuhn, & Klockner, 2002).

Race and habit may also have influence on determining the shopping behavior. Generally, in previous studies, the black families have a lower household income than the non-black families, which in turn may influence the number of vehicles per household adult, shopping frequency and travel mode (Blaylock, 1989; Farber et al., 2014). On the contrary, many studies found that the awareness or habit of travel mode choice did not strength or weaken the car use or people's travel behavior (Garvill, Marell, & Nordlund, 2003; Klöckner & Matthies, 2004).

As for personal income and employment status, Carrasco (2008) found that individual income was more decisive than age and gender. High-income individuals assign a high value to their time and hence are less willing to spend time on shopping trips. Bhat (1998) found that it is less likely for the unemployed respondents than the employed to go shopping by driving alone.

2.1.2 Household Socioeconomic Characteristics

The household is a congregation of its members, hence household socioeconomic characteristics could provide a more comprehensive way to explain the travel behavior. The number of household members (family size), house ownership, household income, car ownership and their number, employment status for adult members, etc. are main factors found to be important in literature about travel behavior (Carrasco, 2008). In addition, the household head's age and the family's participation in the Food Stamp Program are some other detailed considerations of household characteristics. Larger families need to buy larger quantities of products due to their higher level of consumption, which, given fixed storage capacity in their homes, necessitates a larger number of driving trips. Clifton (2004) found that households with a lower income preferred driving, either alone or carpool in Austin, Texas. Handy's research showed people who owned a car preferred to go shopping by driving, because of the convenience and the time duration (Handy 1996). In Europe, people preferred to walk for shopping when cars were not available (Coveney & O'Dwyer, 2009). The number of vehicles for each adult in one household is also strongly related to the car usage. If there are more cars per adult in one household, the family members will be more likely to drive to the store (Scheiner & Holz-Rau, 2012).

2.1.3 Shopping Characteristics

Other factors affecting customers' behavior are shopping characteristics and preferences, such as shopping spending (include travel cost and inventory cost) (Bawa & Ghosh, 1999; K. Clifton, Morrissey, & Ritter, 2012), customer loyalty (Bell, Ho, & Tang, 1998), store open hours (Geiger, 2007), public transportation fees (Litman, 2009), shopping time, shopping composition (or the percentage of food) and so on. However, most literatures studying these factors aim to discuss shopping frequency, departure time choice or store choice instead of shopping travel mode choice (Bhat, 1998a; Jiao et al., 2012).

Shopping distance and time duration are also important aspects in studying shopping behavior. Holz-Rau found that driving is the primary mode in Berlin when this distance was more than 670 m (about 0.4 mi) from the household to the reported store (Jiao, Moudon, & Drewnowski, 2011).

2.1.4 Built-environment Factors

Built-environment, or the physical neighborhood structure, can significantly shape the mode choice (Cervero, 2002; Ewing & Cervero, 2010; Schwanen & Mokhtarian, 2005). In the field of urban planning, transportation researchers coined and defined “three Ds,” density, diversity and design, to describe the built-environment (Cervero & Kockelman, 1997). Recent studies extended this description to “five Ds”, density, diversity, design, destination accessibility and distance to transit (Ewing & Cervero, 2010). The location of residence, land use mix, employment rate, job-to-population ratio, street density, population and job density, points of interest, distance to a store, and distance to the closest transit stop or grocery store are the most frequently used built-environment factors in travel mode studies.

In general, suburban residents tend to use private vehicles rather than other travel modes, such as public transit. On the contrary, urban residents tend to use public transit, biking and walking more, but still prefer private vehicles (Schwanen & Mokhtarian, 2005). Frank & Pivo (1994) also found that people living in neighborhood with a higher land use mix were less likely to choose driving alone for shopping; and Scheiner & Holz-Rau (2007) concluded that higher land use mix usually meant a good transit system, which may lower the possibility that people will drive their own cars. A number of studies also proved that the higher intensity of land use meant the surrounding neighborhood was more friendly to the pedestrian and the public transit service was also better (Cervero & Kockelman, 1997; Cervero & Radisch, 1996; Cervero, 1996, 2002; Frank & Pivo, 1994; Schwanen, 2004).

Similar to the land use mix, the street density and the number of intersections reflect the network connectivity and the accessibility to stores. Guo et al. (2007) mentioned that the higher the density of bikeway, the more likely people chose to go shopping by bicycle or on foot. Other factors, such as the employment rate, the population density rate, and the number of convenient stores, quick-service restaurants, coffee shops, bus stations and the number of traffic signals within a certain distance are all detailed considerations about built-environment (Jiao et al., 2012).

This literature review suggests a number of topics that warrant further research. First, the impact of built-environments centered on the grocery stores deserves more research efforts. Second, the characteristics of stores, such as sale volumes, tend to escape scholarly attention. In fact, store characteristics, such as the store size, the sales amount, the number of employees, and the shopping distance are important attracting factors behind shopping behaviors and mode choice (Jiao et al., 2012). Third, the notion of built-environment is fuzzy (Handy et al., 2002). Previous studies tested either household built-environment or the store built-environment, and little has been done to combine the two perspectives together while taking into account the impact of spatial scales.

2.2 Geographical Scale

Geographical scale could help LUM to measure the variety of a neighborhood, and the relationship between neighborhood context and people's travel behavior (Duncan et al., 2010; Handy et al., 2002). However, there is no fixed definition about the size of a neighborhood (Brownson, Hoehner, Day, Forsyth, & Sallis, 2009; Gauvin et al.), and this open question is also known as modifiable areal unit problem (MAUP), which is associated with the use of geographically aggregated data, and the unit or scale of those data may not be the same (Coombes, Green, & Openshaw, 1986; Openshaw, 1977). Different spatial partitions were adopted in different researches, varying from census block, census block group, census tract and ZIP code to buffers with different sizes (Cervero & Radisch, 1996; Coveney & O'Dwyer, 2009; Hess & Moudon, 1999; Jiao et al., 2012). No matter the geographical scale is defined by the census administration boundaries or the buffer distance, they are used to measure the walkability or driving possibility. In one of the latest researches about walkability and health in Salt Lake County, Utah, Fan et al. (2014) suggest that different focuses should adopt different scales. For instance, when analyzing the large grocery store, ZIP code zone or a buffer with more than a 1,000-meter radius is a better choice; when studying the limited-service, census tract level or larger would be effective; while for the full-service restaurants, the census tract or block group level is considered better; if more factors are investigated together, census tract would also be a better choice than census block group or ZIP code.

In this study, there will be seven geographical scale surrounding the household and the store: half-mile buffer circle (0.5 mi circle), one-mile buffer circle (1 mi circle), half-mile network buffer (0.5 mi network) and one-mile network buffer (1 mi network), census block (Block), census block group (BG), census tract (CT). The demographic and socioeconomic predictors, as well as distance factors, are controlled when different scales are applied to the household or store built-environment factors. We assume that the combination of different geographical scale may provide a more nuanced understanding of the significance of factors affecting mode choices in urban shopping trips.

2.3 Modeling Shopping Behavior

The most frequently used travel mode choice models are binary or binomial model (BL), multinomial logit model (MNL), and nested logit model (NL), since the travel mode is usually regarded as categorical data. The dependent of the binomial logistic regression model is the travel mode, which is dichotomous. For instance, driving or not driving, auto or non-auto. Jiao et al. (2012) built the binomial model level by level: Level 1 was socioeconomic characteristics, Level 2 was the distance model, Level 3 was the home neighborhood model, and Level 4 was the primary store neighborhood model. Rodrigues et al. (2014) used the binary logistic regression model to examine the socio-economic and attitudinal differences between bicyclists and drivers. The Akaike criterion and likelihood-ratio test are frequently used to test the goodness of fit of the binomial model.

If the dependent variable has three or more unordered levels, multinomial logit model (MNL) will be a better choice. Unlike the BL model, MNL relies on the assumption of independence of irrelevant alternatives (IIA), which will be discussed in more detail in Chapter 3. An alternative is the nested logit model (NL), which assumes a hierarchical decision process instead of the independence assumption (Hoffman & Duncan, 1988). When studying the intercity business travel mode choice in the Ontario-Québec corridor of Canada, Forinash & Koppelman (1993) found that NL with bus-train or car train nests was superior to the MNL. The original models mentioned above have been improved by many researchers to fit their own study conditions. The Scobit-based model took the BL model as a special case (Zhang & Timmermans, 2015). Combined with an Ordered Generalized Extreme Value (OGEV), Bhat (1998) found that MNL-OGEV model was a better method than the simple MNL or the nested logit model for a 1990 San Francisco Bay Area travel survey. Based on the collected data, MNL became an optimal choice to identify correlations between socioeconomic characteristics, neighborhood factors and shopping trip travel mode choice.

Chapter 3 Methodology

Based on the previous literature and our research objectives, we identified the following questions to address with the survey data: Do individual demographic, socio-economic and household built-environment factors affect shopping travel mode, no matter whether positively or negatively? Does a store's built-environment as well as its sale amount have strong influences on shopping trips? Does spatial scale matter for the built-environment of both household and store? Chapter 3 demonstrates the methodology used to address these questions.

3.1 Data

The data used in this study is from the 2012 Utah Travel Study, conducted by five agencies: the Wasatch Front Regional Council (WFRC), the Mountainland Association of Governments (MAG), the Cache Metropolitan Planning Organization (CMPO), the Dixie Metropolitan Planning Organization (DMPO), the Utah Department of Transportation (UDOT), and the Utah Transit Authority (UTA). This study is comprised of two stages: a statewide household travel diary survey, and seven randomly-distributed complementary surveys, including the long distance survey, the college travel diary, the bike/pedestrian debrief survey, the bike/pedestrian barriers survey, the attitude debrief survey, the Dixie (SunTran) onboard survey, and the residential choice stated preference survey. Basic information about residents' travel behavior for the Utah statewide transportation planning is provided by this comprehensive travel study. Several fruitful studies referring to this dataset have concentrated on different topics. For instance, Farber et al. (2014) developed an innovative way to assess social equity impacts of distance-based public transit fees; using a latent class analysis, Liao, Farber, & Ewing (2014) studied the significant heterogeneity in residential location preferences in the Wasatch Front region.

This thesis analyzes the shopping travel mode choice based on the household travel diary survey. The original purpose of this survey was for the Wasatch Choice 2040 long range development and transportation plan, and the original data set is categorized into four parts: household information, adult travel diary, child travel diary and debrief questionnaires. This thesis mainly concentrates on the household information and the adult travel diary, with a focus on Salt Lake County, Utah, given that the public transportation system is more developed in this county compared with other counties.

In the survey, the household information is supposed to be completed by one adult in the household. This adult was asked to provide household and personal demographic information as well as the vehicle data. The survey collected the number of adults and children in each participating household (household size), household location, years living at current residence, housing type, the average household annual income, etc. The personal data includes gender, age, race, education, employment status, number of jobs of household members and so forth.

Each person aged 18 or above also completed the adult travel diary, which listed all locations (longitude and latitude by Google Map) the respondents visited on their assigned travel data, as well as the start and end times, trip purpose and travel mode used. Since most children's travel destinations were school or sport sites, this shopping travel mode research does not consider the child travel diary and other non-applicable supplementary surveys.

In addition, the shopping trips could also be categorized into different types according to a variety of criteria. Based on the origin location, there are home-based shopping trips and non-home-based shopping trips; based on trips' origin (O) and destination (D), we have internal-internal (region resident, O and D in the same region), internal-external (region resident, O is in the region, D is outside of that region), external-internal (outside resident, O is outside of the region, D is within that region) and external-external trips (outside resident, O and D are outside of the region). This thesis focuses on the home-based internal-internal shopping trips within Salt Lake County, Utah.

The 2012 Utah Travel Study questioned 9,155 households in the entire state of Utah. The following analyses are based on 1,294 shopping trips from 1,091 households derived from the Utah Travel Survey. Among them, more than 90% of shopping trips were completed by either driving alone (799 trips) or carpooling (420 trips). Only 5.8% of trips (75 trips) used other transportation modes: public transit (16 trips), bicycle (15 trips), walking (41 trips), and others (3 trips). Driving alone, carpooling, and others (bicycling, walking, taking public transit, etc.) are dependent variables in this research.

In addition to the built-environment factors around the household, a store's built-environment factors are also important for the mode choice. We collect data, such as sale amount, number of employees and size of the store, to better characterize a store. These data were collected from the infoUSA website (http://www.infousa.com/) about grocery stores, supermarkets, shopping centers and markets. Grocery stores are the smaller contenders, with annual revenues of $2 million or less, whereas supermarkets generate more than $2 million in revenue within an average space ranging anywhere from 20,000 to 65,000 square feet. Supercenters, known also as hypermarkets, are in essence department stores combined with supermarkets together under one roof, occupying an average of 170,000 square feet. There are 293 reported shopping destinations that were used in this case study of Salt Lake City. Points of interests, such as restaurants, cafés, and pharmacies, surrounding both household and stores, are collected also from the infoUSA website. The data regarding employment rate is from the American FactFinder website (http://factfinder.census.gov/), and the Utah Automated Geographic Reference Center (AGRC) provides the number of network intersections.

3.2 Variable Selection

Individual's background, habit and household environment have an influence on their choice of travel mode (Bamberg & Schmidt, 2010), and this study analyzes factors in four main classes: socioeconomic characteristics, distance variables, household built-environment factors and store built-environment factors.

3.2.1 Socioeconomic Characteristics

Summarizing from data sources mentioned above, we gathered individual and household socioeconomic data: gender, age, employment status, education, race, household income, house ownership, the number of vehicles, bicycles, children, and adults in one household, number of years lived at current residence, etc. Selected questionnaire questions and screenshots of the survey webpage can be found in the Appendix A.

In order to reduce the bias, we rearranged the categories of several factors, such as age and employment status. According to the 2014 Utah Driver Handbook, a person younger than 17 years of age should be accompanied by a licensed driver who is at least 21 years ago and occupying a seat next to the driver. Also limited by the category of the 2012 Utah Travel Survey, this study, therefore, eliminated trips taken by people under 18 years to avoid counting too many carpooling. As for the employment status, we also combined full-time employment, part-time employment and self-employment to one group: employment.

3.2.2 Distance Variables

Using ArcGIS Geocoding and Network Analysis tools, all respondents' home locations and the reported stores in Salt Lake County can be displayed on the transportation network map. Under the Network Analyst Toolbar, employing the OD Cost Matrix, we estimated the distance to the downtown area (CBD distance) and the home-store network distance (OD distance). The distance from the nearest public transit station to each respondent's household (Transit distance) was also calculated by taking advantage of the OD Closest Facility function.

3.2.3 Land Use Mix Index

The land use mix index (LUM) is an entropy index, which measures the balance and basic characteristics of land zones (Cervero & Kockelman, 1997), such as residential, agricultural, commercial, educational, etc. Previous works found that different computations or classifications of the land use mix index reflect the availability of destinations where people choose to walk or ride bicycles (Cervero & Gorham, 1995; Christian et al., 2011; Crane & Crepeau, 1998). The calculation of land-use mix by Frank, Andresen, & Schmid (2004) is

$$LUM = -\sum_{i=1}^{n} p_i \ln p_i \,/\, \ln n \tag{1}$$

where p_i = the proportion of the area covered by land use i against the summed area for land use classes of interest (including i), and n is the number of land use classes of interest. If the value of LUM is close to 1, then there are more different types of land, or maximally heterogeneous land use; when the LUM value is close to 0, there are not too many various land use types, or homogeneous land use. There are two alternatives to get statistical references of mixed land use: Shannon's index (Shannon & Weaver, 1949) and Simpson's index (Simpson, 1949). Unlike Frank's entropy score, Shannon's and Simpson's indices present higher diversity when there are more types of land use categories (Nagendra, 2002; Yamada et al., 2012). In our study, Frank's index is adopted to calculate the land use mix index (LUM), reflecting the diversity in different geographical scales.

Controversy remains regarding the number of land use types in calculating LUM. When analyzing the relationship between obesity and community design, Frank et al. (2004) first introduced the above formula and suggested four land use measures: residential, commercial, office, and institutional. Duncan et al. (2010) adopted five categories: residential, commercial, industrial/institutional, recreational, and "others." In another obesity research, Yamada et al. (2012) used six categories: single-family residential, multi-family residential, retail, office, education and entertainment, which could be traced back to Frank et al. (2006) and Brown et al. (2009). After comparing different calculation methods, Christian et al. (2011) did not suggest that "public open space" should be one of the categories, and neither did Yamada et al. (2012).

Based on the property type codes provided by Salt Lake County assessor's office, this thesis adopted six categories: single-family residential, multi-family residential, commercial, office, educational/institutional and recreational. Industrial and open public area were not included in our categories. The "single-family residential" means the general single-family residence as well as the mobile home land. The "multi-family residential" land use category indicates duplex, three or more unit apartment, mixed apartment, and condo, as well as non-private facilities such as hotel, motel, fraternity/sorority, etc. The "commercial" land use includes shopping center, drug store, food market, restaurant as well as post office, bank, auto service and other business services. "Office" refers to general business office. The “educational/institutional” land use type contains university, college and other schools as well as government building, hospital, church, etc. “Recreational” land includes park, theater, health club, etc.

The geographical scale also varies in different research on the land use mix factor. Frank et al. (2004) used a 1-kilometer distance from each respondent’s household, while Yamada et al. (2012) focused on three different spatial scales: census tract, block group and 1-kilometer street-network buffer. In our research, we address the built-environment factors at different scales when exploring their impacts on travel mode choices. The LUM is examined at seven different scales: half-mile buffer, one-mile buffer, half-mile network buffer, one-mile network buffer, census block, census block group, and census tract.

The calculation of Frank’s entropy is tedious, due to the large database and the research purpose. There is no ready-made tool in ArcMap, but other tools provided by the third party are neither outdated nor expensive, for example, Hawth’s Analysis Tools for ArcGIS, XTools Pro, Patch Analyst 5. To facilitate the calculation, we developed an Arcpy tool in ArcMap 10.2, which allows the user to select target point (household or store), buffer distance or geographical scale polygons, and the output location of LUM results. Codes are attached in the Appendix B.

3.2.4 Density Factors

The land use mix index (LUM) is not the only measurement to quantify the neighborhood characteristics. Some research results showed that the link between non-work travel and land-use patterns at neighborhood level were not that significant or were inconclusive compared with other factors, therefore the transportation policy based on those factors were unreliable (Boarnet & Sarmiento, 1998; Kitamura, Mokhtarian, & Laidet, 1997). Two studies about Seattle found that a shopper would take non-auto shopping trip travel mode if the store was located in a higher density neighborhood (Chen et al., 2008; Jiao et al., 2011). Density factors, therefore, are also introduced in this research, and are supposed to be significant in explaining the correlation between surrounding environments and shopping trip travel mode. We studied the street density, residential density, job density, population density, as well as densities of points of interest, such as the number of traditional restaurants, bus stations, cafés, quick-service restaurants, convenience stores, liquor stores, and road intersections within different geographical scales.

Street density or road density, which provides us an indicator of the connectivity and accessibility of an area, is the total length of the entire road network per square mile or square kilometer. We intersected the selected geographic polygon with the Salt Lake County network and calculated street density by the field calculator in the ArcMap.

The residential density is calculated by counting the total number of single-family and multi-family residences within a certain geographical scale. Population and employment data from the Census Bureau website and American FactFinder were used to calculate the population density and job density.

It is relatively straightforward to calculate the population density and the job density within the census block group and census tract level, since the data from the US Census Bureau is based on census block. However, it is difficult to gather specific information within the irregular boundaries of the buffer, either the straight-radius buffer or the network buffer. Hence, we assume that the population and job distribution are proportional to the percentage of the area within its census block. This estimation was also completed in ArcMap with a compiled Arcpy snippet.

As for the point of interest, we collected information from the Utah Automated Geographic Reference Center (AGRC). We counted the number of each type of point of interest within different geographical scales, and spatially joined household or store with these values.

3.2.5 Sales Amount

The store characteristics, such as the sales amount, the food pricing, the number of employees, the size, and the number of parking lots, tend to determine people's shopping involvement, like the shopping frequency, expenditure as well as the travel mode choice (Bawa & Ghosh, 1999; Putrevu & Ratchford, 1997; Smith & Carsky, 1996). In our study, the infoUSA provides the location of stores and their annual sales amount and the number of employees.

In order to test the aggregation effect of stores, this study concentrated on not only the reported stores but also the overall sales amount in the neighborhood. For instance, if a respondent said his trip was a home-based shopping trip but the location he pointed on the map was a gas station close to a large shopping center. Thus, using the total sales amount within a certain geographic area would objectively reflect his primary purpose or the attraction of the shopping mall leading to his visit to that gas station.

3.3 Geographical Scale Selection

As mentioned in Chapter 2, the correlation between neighborhood environment and travel mode choices is also sensitive to geographical scale. In general, the census block is the smallest unit followed by half-mile buffer, census block group, one-mile buffer, and the census tract. The areas of network buffer varies according to the surrounding number and types of streets.

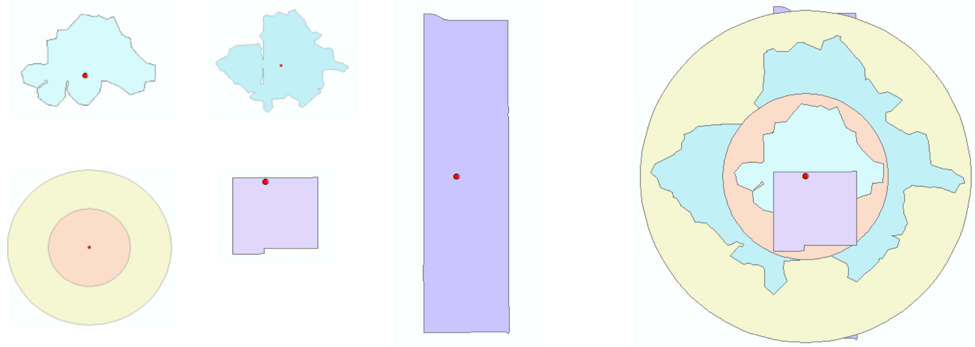


Figure 3.1: An example of different geographical scales:

(First line, network buffer half-mile and one-mile; Second line, straight-line buffer half-mile and one-mile, block group; Middle, census tract; Right, comparison of different geographical scale areas)

The half-mile and one-mile buffers were chosen because previous researchers have demonstrated the effectiveness and usefulness, no matter for fixed radius buffer or the network buffer. The fixed buffer is a circle centered in a household or a store based on a straight-line distance, half or one mile in our study; the street-network buffer reflects the actual driving or walking paths along the surrounding streets. The network buffer could also be defined by a fixed time period. Due to the difference of the speed of driving, walking, biking and riding a bus, the fixed distance network buffer was taken into effect (Figure 3.1).

The US Census Bureau defines census blocks as statistical areas bounded by visible features (such as streets, roads, streams, and railroad tracks), and by nonvisible boundaries (such as selected property lines and city, township, school district, and county limits and short line-of-sight extensions of streets and roads). Generally, census blocks are small in area; for example, a block in a city bounded on all sides by streets. Census blocks (Block) in suburban and rural areas may be large, irregular, and bounded by a variety of features, such as roads, streams, and transmission lines. In remote areas, census blocks may encompass hundreds of square miles.

Block groups (BG) are statistical divisions of census tracts, which are generally defined to contain between 600 and 3,000 people, and are used to present data and control block numbering. Most BGs were delineated by local participants in the Census Bureau's Participant Statistical Areas Program. The Census Bureau delineated block groups only where a local or tribal government declined to participate, and a regional organization or State Data Center was not available to participate. A block group usually covers a contiguous area (Figure 3.1). Each census tract contains at least one block group, and block groups are uniquely numbered within the census tract. Within the standard census geographic hierarchy, block groups never cross state, county, or census tract boundaries but they may cross the boundaries of any other geographic entity.

Census tracts (CT) are small, relatively permanent statistical subdivisions of a county, and it is uniquely numbered in each county with a numeric code. Census tracts average about 4,000 inhabitants. The minimum population is about 1,200, while the maximum population is around 8,000. Census tracts are designed to be relatively permanent over time. Any changes are documented so data can be compared from decade to decade.

3.4 Model Selection

The generalized linear modelling technique of multinomial logistic regression can be used to model unordered categorical response variables. This model can be understood as a simple extension of logistic regression that allows each category of an unordered response variable to be compared to a reference category providing a number of logit regression models. A binary logistic regression model compares one dichotomy (for example, passed-failed, died-survived, etc.) whereas the multinomial logistic regression model compares a number of dichotomies. This procedure outputs a number of logistic regression models that make specific comparisons of the response categories. When there are j categories of the response variable, the model consists of $j-1$ logit equations which are fit simultaneously. Multinomial logistic regression is a technique that basically fits multiple logistic regressions on a multi-category unordered response variable that has been dummy coded. The general multinomial logistic regression model is shown below:

$$log \frac{\Pr(Y=j)}{\Pr(Y=j')} = \alpha + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \epsilon \tag{2}$$

where j is the identified travel mode (carpool and walking, for instance), and j' is the reference travel mode (driving alone, for instance). In order to make sure the multinomial logit model can be safely used, the IIA (independent of irrelevant alternative) hypothesis has been applied.

Two main prediction models, multinomial logit model (MNL) and generalized nested logit model (GNL), have been widely employed in grocery shopping behavior analyses. For example, Recker & Kostyniuk (1978) estimated the destination choice for the urban grocery shopping trip, using multinomial logit model and found accessibility had a nonlinear influence on customer's choice. Wen & Koppelman (2001) examined high-speed rail in Toronto-Montreal corridor, by applying a generalized nested logit model.

Some other statistical analysis for grocery shopping benefited from binomial model (Frisbie Jr., 1980) or binary logit model (Innes, Ircha, & Badoe, 1990), stochastic model (Davies & Pickles, 1987), Cox regression model (Schwanen, 2004), etc. With these comprehensive considerations, we believe that MNL will provide us a more scientific way to build a model containing both continuous and categorical variables (McFadden, 1977).

Chapter 4 Descriptive Analysis

The descriptive analysis provides a preliminary overview of various factors in different geographic scales. This initial understanding is that travel mode choices are related to socioeconomic characteristics (gender, age, race, employment status, household annual income, and household composition), household distribution, and built-environment factors (sales amount, LUM, street density, residential density, and the number of restaurants, convenience stores, bus stops, etc.).

4.1 Individual and Household Socioeconomic Background

Gender differences have a strong effect on the travel mode choice (Scheiner & Holz-Rau, 2012). Among the 1,294 respondents, 61% were female and 39% were male (Table 4.1). Figure 4.1 shows that male respondents preferred driving alone and using other transportation modes than female respondents: there were 65.6% and 7.3% male respondents chose to drive alone and use other modes respectively, while the percentages of those for female were 59.2% and 4.8% respectively. Meanwhile, the female inclination of choosing carpooling for shopping was stronger than the male: 35.9% of female respondents used carpooling as opposed to 27.1% of male respondents.

Table 4.1: Individual characteristics

Variables	Frequency	Percentage
Gender		
Male	509	39.33
Female	785	60.66
Age		
[18,24]	78	6.03
[25,34]	280	21.64
[35,44]	245	18.93
[45,54]	195	15.07
[55,64]	256	19.78
Above 64	240	18.55
Employment		
Employed full-time	483	37.33
Employed part-time	148	11.44
Self-employed	83	6.41
Student	67	5.18
Homemaker	211	16.31
Retired	237	18.32
Not currently employed	65	5.02
Education		
Less than high school	14	1.08
High school graduate	131	10.12
Some college	261	20.17
Vocational/technical training	51	3.94
Associates degree	110	8.50
Bachelor's degree	455	35.16
Graduate/post-graduate degree	272	21.02
Race		
Hispanic	55	4.25
African American or Black	8	0.62
American Indian or Alaskan Native	6	0.46
Asian	37	2.86
White or Caucasian	1174	90.73
Other	31	2.40
Prefer not to answer	38	2.94

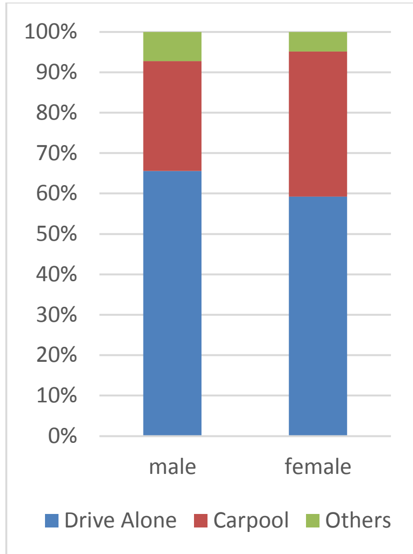


Figure 4.1: The choice of travel mode vs. gender

The average age of all respondents fell in the range between 45 and 54 years old (Table 4.1). Figure 4.2 shows the correlation between age and travel mode. We can see that young people (18-24 and 25-34 years old) were more likely to choose to carpool, and people aged 25-34 and 65-74 were more likely to choose other modes, compared to other age categories. The disadvantaged financial status of the young and the health pursuit of senior citizens may lead to their preference of choosing to drive alone or to carpool.

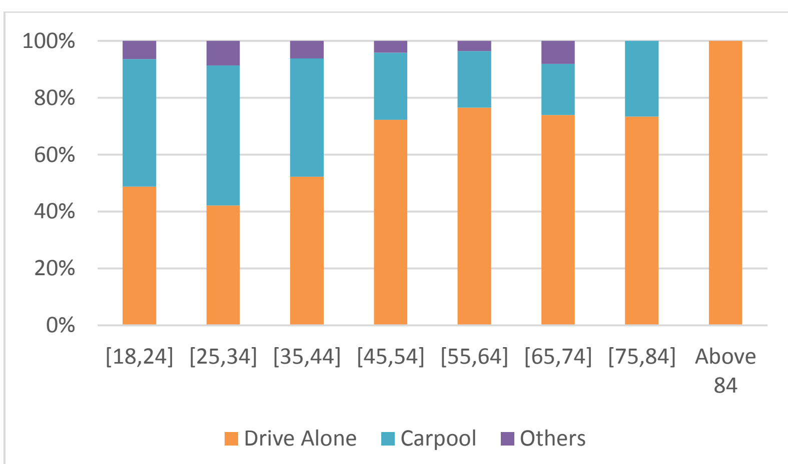


Figure 4.2: Age vs. shopping trip travel mode

Table 4.1 also provides the information about respondents' employment status, education level and race. The unemployment rate in Utah was about 5.8% in January, 2012, which was slightly higher than 5.02% in the studied Salt Lake County. The education level of the database was above the statewide average. Utah ranks 4th in the nation in terms of persons 25-years and older who have attained bachelor degrees (30.8%), whereas the number was 56.18% in this study. The percentage of the white citizens was also slightly higher than that of the state average (89.2%). Hence, based on the individual information, most respondents in this study were employed white young females with a good education background.

As for the household characteristics, household income, family size, family composition and several travel habits were examined (Table 4.2). The average household annual income fell within the $5,000 to $99,999 category. The lower the household annual income was, the higher the possibility that the shopper would walk, bike or take public transit to the store; while the higher the income was, the higher the possibility of the shopper to drive alone for shopping (Figure 4.3). This findings agree with Moudon, Design, & Drewnowski's study in King County, WA (2011), which showed that income per adult was positively related to drive for the grocery shopping.

Table 4.2 Household characteristics

Variables	Frequency	Percent	Variables	Frequency	Percent
Household Income			***Number of adults in the household***		
Under $35,000	217	16.77	1 person	225	17.39
$35,000 - $49,999	158	12.21	2 persons	843	65.15
$50,000 - $99,999	473	36.55	3 persons	140	10.82
$100,000 or more	286	22.10	4 persons	56	4.33
Prefer not to answer	160	12.36	5 persons	20	1.55
Household Life Cycle			6 + persons	10	0.77
No children or retirees	528	40.80	***Number of children in the household***		
With children no retirees	475	36.71	0 child	805	62.21
Household with retirees	291	22.49	1 child	172	13.29
Rent or Own the Home			2 children	143	11.05
Rent	242	18.70	3 children	103	7.96
Own/Buying	1026	79.29	4 children	53	4.10
Other	12	0.93	5 children	14	1.08
Prefer not to answer	14	1.08	6 + children	4	0.31
Number of years lived at residence			***Number of vehicles in the household***		
Less than 1 year	125	9.66	0 vehicle	23	1.78
1-5 years	371	28.67	1 vehicle	317	24.50
6-10 years	259	20.02	2 vehicles	636	49.15
11-15 years	136	10.51	3 vehicles	215	16.62
16-20 years	110	8.50	4 vehicles	72	5.56
More than 20 years	293	22.64	5 or more	31	2.40
Frequency of riding transit			***Number of bikes in the household***		
6-7 days a week	7	0.54	0 (none)	438	33.85
5 days a week	31	2.40	1 bicycle	255	19.71
4 days a week	10	0.77	2 bicycles	337	26.04
3 days a week	21	1.62	3 bicycles	132	10.20
2 days a week	11	0.85	4 bicycles	83	6.41
1 day a week	8	0.62	5 bicycles	30	2.32
A few times per month	77	5.95	6 or more	19	1.47
Less than monthly	495	38.25			
Never	634	49.00			
Typical commute mode to work					
Auto/truck/motorcycle	548	42.35			
Transit	35	2.70			
Walked/wheelchair	11	0.85			
Bicycle	19	1.47			
Other	1	0.08			

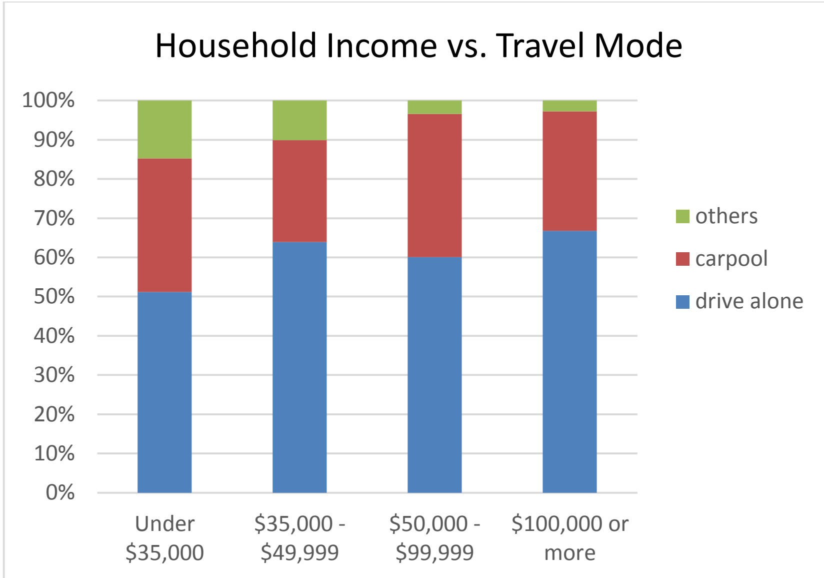


Figure 4.3: Household annual income vs. shopping trip travel mode

As summarized in Table 4.2, the average number of adults was 2.10; the average number of children in per household was 0.83. Compared with the data from the Census Bureau, respondents in Salt Lake County had smaller size and more children per family.

Almost 80% of families had already owned or bought the household, and more than half of participating families had lived in the current residence for at least 10 years; 22.64% had lived there more than 20 years.

As shown in Figure 4.4, people in households whose number of vehicles were more than 1 prefer to drive alone for shopping, people who had no vehicle in the household were more likely to choose walking or biking. The number of bicycles in one household also influenced its members' travel mode choices (Figure 4.5): households with three bicycles were more likely to ride a bike for shopping than other groups.

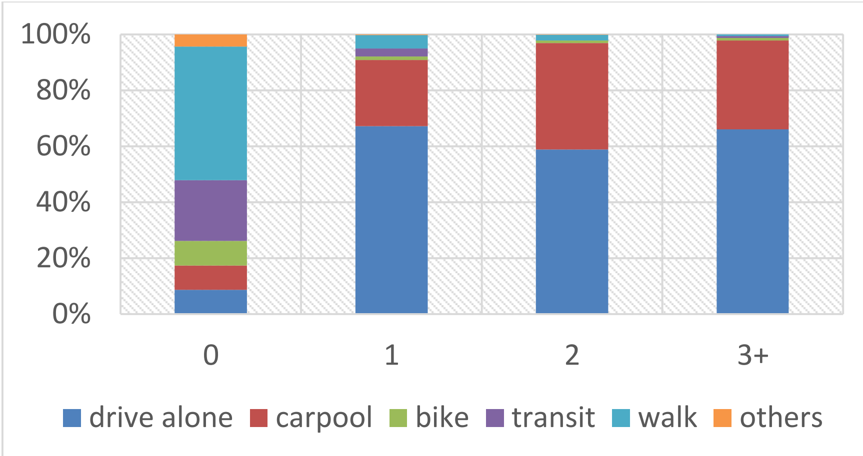


Figure 4.4: Number of vehicles in household vs. mode choices

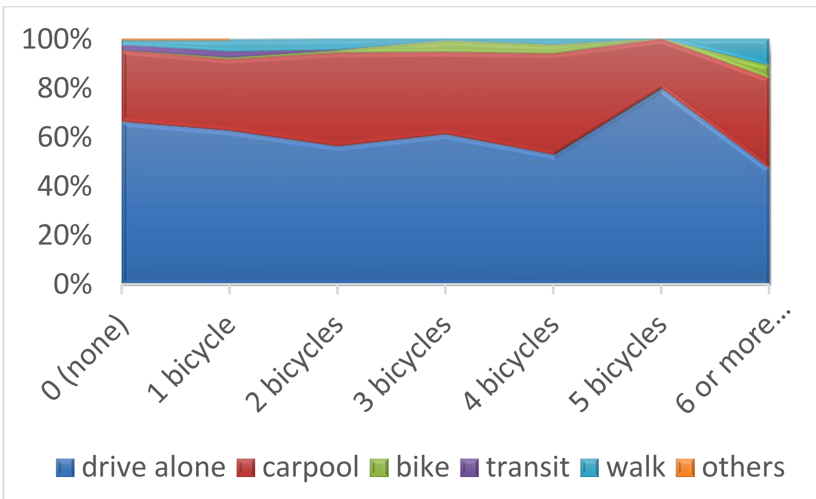


Figure 4.5: Number of bicycles in household vs. shopping trip travel mode

4.2 Built-environment Analysis and Geographical Scales

There were 1,294 trips filtered from the 2009 Utah Travel Survey. All these trips were home-based, internal-internal shopping trips in Salt Lake County, Utah. Among the respondents, 61.75% drove alone, 32.46% carpooled, and 5.80% chose walking, biking, and riding bus or light rail (Table 4.3). This finding is different from other major cities. For instance, Jiao et al. (2012) in their Seattle research found that 88% of 1,885 respondents drove for shopping, and 12% used other travel modes. In Detroit, about

62% of the respondents mentioned they never took bus to the grocery store (Boivin, 2008). Clifton et al. (2012) reported that more than 85% of customers drove their vehicles for shopping in Portland, OR, whereas only 0.9% chose to ride the bus. The high percentage of the respondents who drove compared with the percentage of those using other transportation modes may result from the construction of the auto-oriented infrastructure, the accessibility to the public bus stations and the shopping centers, the commercialization of the surrounding area, and other built-environment factors.

Table 4.3: The percentage of different travel modes

Travel Mode	Freq.	Percent (%)
Drive Alone	799	61.75
Carpool	420	32.46
Bicycle	15	1.16
Transit	16	1.24
Walk	41	3.17
Others	3	0.23
Total	1294	100.00

The kernel density of household locations in Figure 4.6 also reflects people's preference on choosing certain travel modes. This map shows several expected trends, such as: most people who drove alone or carpool for shopping lived close proximity to the traffic network, and customers using other transportation modes lived adjacent to shopping centers, especially the downtown area, which was the most convenient location for them to acquire living supplies.

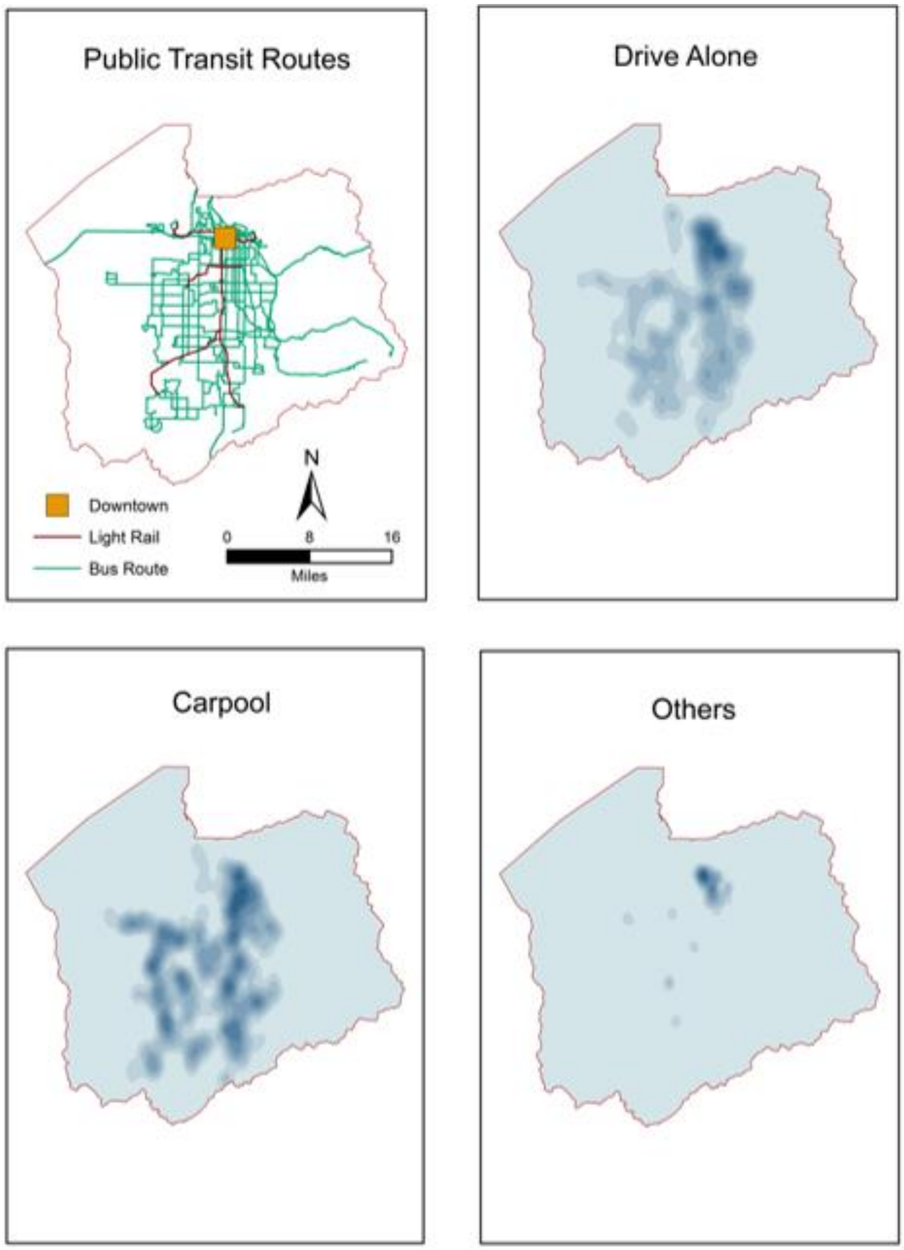


Figure 4.6: Kernel density of household travel mode distribution

Three accessibility factors, OD distance, Public transit distance, and CBD distance, were measured to reflect the spatial distribution of people using different travel modes. As shown in Figure 4.7, on average, people choosing driving alone to the store had the longest distance to the reported store (3.684 mi), to the nearest public transit station (0.434 mi), and to the CBD in downtown (11.851 mi). In contrast, among the three modes, the household locations of people using other transportation modes were the shortest to the reported store (2.671 mi), to the nearest bus stop or light rail station (0.200 mi), and to the CBD area (5.806 mi). We can conclude that people using other transportation modes for shopping lived closer to the store and the public transit station.

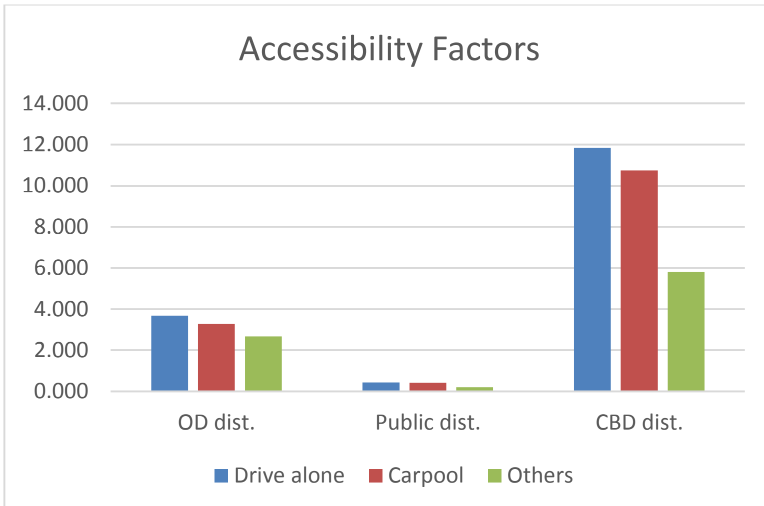


Figure 4.7: Accessibility vs. travel modes

Interestingly, besides the household distribution and distances to the featured destinations, other built-environment factors also showed evident influence on the travel mode choice, especially in different geographical scales. Among studied geographical scales in this research, the one-mile straight-line buffer and the one-mile network buffer contain more points of interest than the half-mile straight-line buffer and the half -mile network buffer respectively. As for the census level, the census block is the smallest scale, followed by the block group and the census tract.

In Figure 4.8, for instance, the number of built-environment factors, such as the number of restaurants, cafés, quick-services, liquor stores and convenience stores, became larger when the geographical scale increased from the half-mile buffer to the one mile buffer, and from the census block to the block group and to the census tract.

Figure 4.8: Numbers of restaurants and convenience stores in different scales

For the number of built-environment factors, there were more points of interest around stores than households in a certain geographical scale. That is to say, the surrounding area around stores was much more commercialized than the surrounding area around households or the residential area. This phenomena could also be reflected by the residential density and the sales amount around households and stores. There were more single and multiple houses around reported household locations than reported stores, and there were more sales amount around reported stores than reported households (Figure 4.9).

Figure 4.9: Residential density and sales amount around households and stores

As mentioned in Section 3.2, the land use mix index (LUM) describes the diversity of land use condition in the neighborhood. In Figure 4.10, most geographical scales demonstrate a relatively diverse neighborhood (LUM > 0.50), and the store's neighborhood was more diverse than the household's neighborhood. The LUM scores were similar among the one-mile straight-line buffer, the network buffers, census block group level, and the census tract level. One exception was the census block level, where the land use tended to be homogenized (LUM < 0.15). Compared with other built-environment factors, therefore, the similarity of LUM values may lead to the insignificance of LUM in the multinomial logit model.

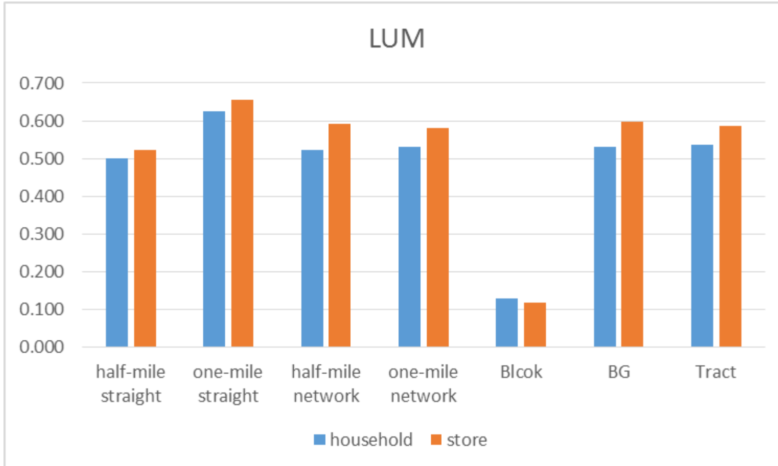


Figure 4.10: Land use mix index in different geographical scales

The street density factor, like LUM, demonstrates the neighborhood diversity from a different perspective. Figure 4.11 shows that the larger the geographical scale was, the lower the street density would be, and the discrepancy was not that significant within the same geographical scale classification. In contrast, there were fewer bus stops around households than around the store neighborhoods. The contradiction between the street density and the number of bus stops may lower people's interest to take public transit for shopping, since more bus stations around households will encourage using public transit as the travel mode. In addition, straight-line buffer circles possessed more bus stops than network buffers, while the network buffers contained a larger street density than straight-line buffer circles. All these differences of built-environment factors in different geographical scales encourage us to study the travel mode choice with a spatial perspective.

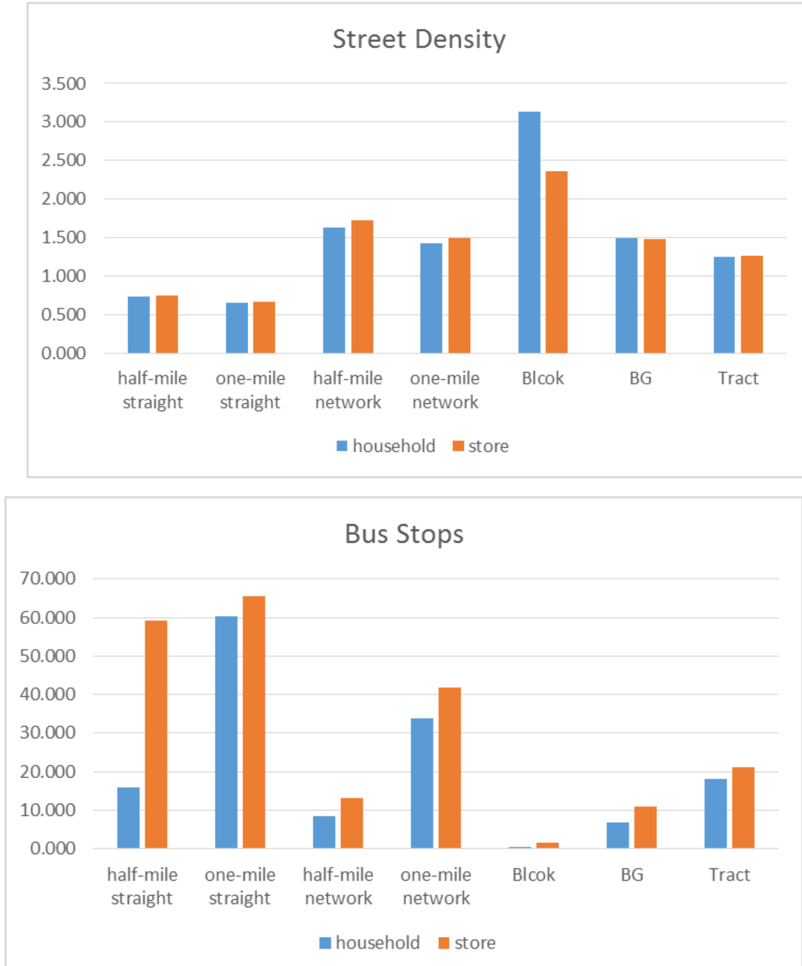


Figure 4.11: Transportation factors

Chapter 5 Multinomial Logit Analysis

5.1 MNL Results

The descriptive results reveal the possible connection between demographic, distance factors, and built-environment, and the shopping trip travel mode. This multinomial logit model (MNL) will test whether those conclusions are still true after taking the interactive influence into account by using driving alone as the alternative model.

The pseudo-R^2 test, one of measurements for the goodness-of-fit, is examined first (Table 5.1), followed by four different models: the base model for demographic factors (Model 1); the accessibility model for OD, the household–public transit stations/CBD distances (Model 2); the household built-environment model (Model 3); and the store built-environment model (Model 4) with sales amount. There are 1,264 trips for each model.

The R^2 statistic measures the variability in the dependent variable that is explained by a linear regression model, and it cannot be computed for multinomial logistic regression models. The pseudo-R^2 statistics are designed to have similar properties to the true R^2 statistic. In this case, most values are greater than 0.15, so the data are consistent with the model assumptions, and the model adequately fits the data. Compared with other travel behavior researches, these pseudo-R^2 values in most of our models are around 0.2, which is relatively low but still acceptable (Jiao et al., 2012; Schwanen & Mokhtarian, 2005).

Table 5.1: R-squared values in different models

R-squared	Base Model	Accessibility Model	Home Neighborhood Model	Store Neighborhood Model
Demographic Factors	0.174			
Distance Factors		0.180		
Half-mile Buffer Circle			0.191	0.223
One-mile Buffer Circle			0.192	0.223
Half-mile Network Buffer			0.193	0.221
One-mile Network Buffer			0.194	0.222
Census Block Level			0.201	0.233
Census Block Group Level			0.196	0.231
Census Tract Level			0.196	0.226

5.1.1 Model 1. Base Model

The base model (Model 1) mainly presents the individual and household background on mode choices, and the pseudo-R^2 of this model is 0.174. Results of Model 1 is in Table 5.2. The study shows how factors such as females aged between 25 and 34, the number of adults, and the number of children in one household positively relate to the likelihood of carpool. These three factors are highly significant ($P < 0.01$), with coefficients of 0.979, 0.525 and 0.463, respectively. Status as a student, the number of bicycles and families who rent enable people to choose other transportation modes. The number of vehicles in one household negatively relate to both carpool and other modes. The respondents' employment status, education level, race, household income, and the number of years living in the current residence are not significant in both the model for carpool and the model for others. Numbers of vehicles and bicycles in the household and families who rent are top predictors for choosing other transportation modes, whose coefficients are -1.294, 0.497 and 1.385, individually. Compared with carpool, respondents having more vehicles in the household are almost six time less likely to choose other modes to go shopping.

Table 5.2: Socioeconomic factors
(Boldface indicates significance ($p < 0.05$) of the test results)

Variables	Model for carpool	Model for others
Female Age [18, 24]	0.601	-0.655
Female Age [25, 34]	**0.979**	0.045
Female Age [35, 44]	0.236	0.064
Female Age [55, 64]	-0.084	-0.316
Female Age [Above 64]	-0.182	0.292
Employed	0.385	-0.109
Student	0.194	**1.023**
Education Level	-0.002	-0.018
Hispanic	0.433	-1.888
Non-White	-0.168	0.619
Income Under $35,000	0.531	0.696
Income [$35,000, $49,999]	-0.213	0.458
Income $100000 or more	-0.305	0.287
# adults	**0.525**	0.239
# children	**0.463**	0.072
# vehicles	**-0.213**	**-1.294**
# bicycles	0.004	**0.497**
Rent	-0.049	**1.385**
Living years	-0.036	-0.204

5.1.2 Model 2. Accessibility Model

The accessibility model (Model 2) consider three network distance factors: the distance between the household and the reported shopping destination (Origin-Destination, OD dist.), the distance between the household and the closest public transit (transit dist.), and the distance between the household and the central business district (CBD dist.). Combined with factors in the base mode, the pseudo-R^2 of this model is 0.180. None of the accessibility factors showed their significance due to the strong affect from the demographic factors. Once these three accessibility factors are respectively tested, their significance began to appear (Table 5.3). The CBD distance is a strong predictor for carpool ($P < 0.01$), and both the transit distance and the CBD distance are also the strong predictors for others. Hence, for an increase of the distance from the household to CBD, people are more likely to carpool and less likely to use other modes; and if the distance to the closest transit station increase, people are also less likely to use other transportation modes.

Table 5.3: Accessibility model
(Note: *Significant at 5%; **Significant at 1%)

Travel Mode	Coef.	P>|z|	[95% Conf. Interval]	
Drive Alone	(base outcome)			
Carpool				
OD distance	0.025	0.158	-0.010	0.059
Transit distance	-0.128	0.321	-0.381	0.125
CBD distance*	**0.034**	**0.003**	**0.012**	**0.057**
Others				
OD distance	-0.012	0.778	-0.099	0.074
Transit distance*	**-1.459**	**0.050**	**-2.925**	**0.007**
CBD distance**	**-0.143**	**0.000**	**-0.202**	**-0.084**

5.1.3 Model 3. Home Built-Environment Model

Model 3 adds thirteen household built-environment factors: land use mix index (LUM), population density, job density, residential density, street density, numbers of traditional restaurants, bus stops, cafés, quick-service restaurants, convenience stores, liquor stores, traffic nodes, and store sales amount. Table 5.4 lists all significant factors in Model 3. Values of pseudo-R^2 become higher and higher in different geographical scales: 0.191 for the half-mile buffer circle, 0.192 for the one-mile buffer circle, 0.193 for the half-mile network buffer, 0.194 for the one-mile network buffer, 0.201 for the census block level, 0.196 for the census block group level, and 0.196 for the census tract level.

In this model, all base model variables remain significant but not for all geographical scales. For instance, status as a student positively relates to use other modes only within the half-mile buffer circle, one-mile network buffer area, the census block and the block group. In comparison with Model 1 and Model 2, households whose annual income are under $35,000, time spent at the current residence, and CBD distance within the census block level become significant as well.

Table 5.4: Results for Model 3
(Boldface indicates significance ($p < 0.05$) of the test results)

Variables	0.5mi circle	1mi circle	0.5mi network	1mi network	Block	BG	CT
Model for carpool							
Female Age [25, 34]	**1.040**	**1.012**	**1.040**	**1.006**	**1.017**	**0.993**	**1.018**
Income Under $35,000	**0.575**	**0.573**	**0.581**	**0.531**	**0.585**	**0.592**	**0.586**
# adults	**0.497**	**0.475**	**0.518**	**0.489**	**0.545**	**0.556**	**0.526**
# children	**0.456**	**0.461**	**0.462**	**0.458**	**0.473**	**0.469**	**0.468**
# vehicles	**-0.225**	**-0.206**	**-0.231**	**-0.228**	**-0.239**	**-0.221**	**-0.225**
Street density	0.945	0.197	0.034	-0.016	**0.101**	0.254	0.329
# restaurants	-0.003	0.000	0.015	-0.024	0.071	**0.018**	0.020
# quick-services	-0.042	0.005	-0.115	-0.035	-0.505	**-0.103**	-0.046
# convenience stores	-0.057	0.056	-0.054	**0.117**	-0.647	0.059	0.027
# traffic nodes	0.001	0.001	0.004	0.001	0.005	**0.004**	0.001
Model for others							
Student	**0.863**	0.840	0.847	**0.951**	**1.061**	**1.123**	0.722
Income Under $35,000	**0.857**	0.806	0.771	**0.866**	**0.858**	**0.797**	0.685
# vehicles	**-1.180**	**-1.223**	**-1.206**	**-1.192**	**-1.207**	**-1.224**	**-1.313**
# bicycles	**0.400**	**0.431**	**0.421**	**0.413**	**0.484**	**0.450**	**0.484**
Rent	**1.092**	**1.012**	**1.231**	**1.129**	**1.019**	**1.001**	**0.939**
Living years	**-0.258**	**-0.258**	**-0.280**	**-0.249**	-0.161	-0.185	-0.199
CBD dist.	0.024	0.002	0.033	0.004	**-0.087**	-0.032	-0.001
Population density	-7.629	-4.372	-0.137	-0.425	0.676	0.855	**7.647**
# restaurants	0.032	**0.024**	0.069	0.039	0.212	**-0.073**	0.018
# convenience stores	-0.054	-0.069	0.133	-0.162	0.336	-0.427	**-0.345**
# liquor stores	-0.272	0.338	-0.074	-0.430	-14.583	**1.133**	0.281
# traffic nodes	0.000	-0.001	0.003	-0.001	**0.021**	0.006	0.002

Model 3 also shows that seven out of thirteen household built-environment factors had impact on people's shopping trip travel mode choices, and their significance should be considered within different geographical scales.

Street density within the block level, the number of traditional restaurants and the number of traffic nodes in the block group level, and the number of convenience stores within the one-mile network buffer around the household location increase the usage of carpooling. The number of quick-service restaurants in the block group level negatively relates to carpool.

Population density in the census tract level, the number of traditional restaurants within the one-mile straight-line buffer, the number of liquor stores in the block group level and the number of traffic nodes in the census block level have a positive impact on choosing other modes. The number of traditional restaurants in the block group level and the number of convenience stores in the census tract level negatively relate to others.

5.1.4 Model 4. Store Built-Environment Model

Similar to Model 3, Model 4 adds thirteen more built-environment factors defining neighborhoods around the reported stores. The values of pseudo-R^2 are obviously higher than the previous corresponding models: 0.223 for the half-mile buffer circle, 0.223 for the one-mile buffer circle, 0.221 for the half-mile network buffer, 0.222 for the one-mile network buffer, 0.233 for the census block level, 0.231 for the census block group level, and 0.226 for the census tract level. All significant factors in Model 4 are listed in Table 5.5. Results show that most socioeconomic factors remained significant in this model, such as females aged between 25 and 34, household income under $35,000, the number of vehicles, and families who rent. A race factor, being a non-white, and the number of adults positively relate to others in Model 4. Nevertheless, no accessibility variables are significant in this model.

Table 5.5: Results for Model 4
(Boldface indicates significance ($p < 0.05$) of the test results)

	Variables	0.5 mi circle	1 mi circle	0.5 mi network	1 mi network	block	BG	CT
	Model for carpool							
Socioeconomic factors	Female Age [25, 34]	**1.051**	**1.002**	**1.078**	**1.010**	**0.933**	**0.994**	**0.989**
	Income < $35,000	**0.561**	**0.579**	**0.578**	**0.544**	**0.711**	**0.626**	**0.609**
	# adults	**0.494**	**0.461**	**0.512**	**0.462**	**0.547**	**0.550**	**0.523**
	# children	**0.462**	**0.472**	**0.461**	**0.461**	**0.504**	**0.472**	**0.478**
	# vehicles	**-0.236**	-0.191	**-0.226**	**-0.219**	**-0.214**	**-0.217**	**-0.228**
Household built-environment	Residential density	0.000	0.000	0.000	0.000	-0.003	**-0.001**	0.000
	Street density	0.659	-0.683	0.025	0.022	**0.112**	0.227	0.139
	# convenience stores	-0.068	0.056	-0.074	**0.114**	-0.667	0.073	0.000
	# traffic nodes	0.001	0.001	0.005	0.002	0.008	**0.004**	0.001
Store built-environment	Job density	-0.042	0.065	-0.037	-0.028	**-0.066**	-0.072	-0.070
	Residential density	0.000	0.000	0.000	0.000	0.002	0.001	0.000
	Street density	**1.205**	1.688	0.157	0.077	**0.138**	**0.544**	0.650
	# quick-services	0.016	**0.033**	-0.002	**0.045**	0.017	0.020	0.053
	# convenience stores	0.026	-0.063	0.017	-0.031	**0.634**	0.013	0.062
	# liquor stores	-0.301	0.004	-0.271	0.055	-0.819	**-0.513**	**-0.319**
	Sales amount	0.027	-0.115	-0.054	-0.224	**0.743**	0.169	0.206
	Model for others							
Socioeconomic factors	Female Age [18, 24]	-1.101	-1.003	-1.058	-0.906	-0.681	**-1.632**	-1.293
	Student	0.632	0.713	0.588	0.632	**1.019**	0.876	0.392
	Non-White	**0.928**	**0.908**	0.819	0.721	**0.902**	0.648	0.374
	Income < $35,000	**1.079**	**1.044**	**1.112**	**1.082**	**0.880**	**1.089**	0.959
	# adults	0.399	**0.576**	0.341	**0.527**	0.274	0.242	0.476
	# vehicles	**-1.202**	**-1.336**	**-1.215**	**-1.347**	**-1.230**	**-1.336**	**-1.334**
	# bicycles	**0.404**	**0.484**	**0.450**	**0.462**	**0.526**	**0.520**	**0.506**
	Rent	**1.431**	**1.485**	**1.546**	**1.384**	**1.048**	**1.205**	**1.190**
	Living years	**-0.295**	-0.213	**-0.276**	-0.209	-0.126	-0.143	-0.141
Household built-environment	LUM	**-3.692**	-0.839	**-2.728**	-1.387	-0.086	-1.579	-0.806
	Population density	-6.766	5.384	-0.043	-0.069	0.609	0.046	**7.068**
	Job density	-0.596	-0.388	**-0.239**	-0.300	-0.021	0.010	-0.201
	# restaurants	0.017	0.022	0.058	0.025	0.138	**-0.085**	0.016
	# liquor stores	-0.156	0.116	0.033	-0.449	-15.248	**1.138**	0.170
	Sales amount	0.473	0.074	0.882	0.036	**4.090**	0.266	0.272

(continued on next page)

	Variables	0.5 mi circle	1 mi circle	0.5 mi network	1 mi network	block	BG	CT
		Model for others						
Store built-environment	LUM	-2.375	-0.156	-3.036	-2.588	**1.223**	1.454	-2.793
	Job density	**0.738**	**0.942**	0.205	**0.566**	-0.010	0.114	0.339
	Residential density	0.001	0.000	**0.003**	0.000	-0.003	0.001	0.000
	Street density	-0.375	-3.824	-0.855	**-2.557**	0.206	0.741	-0.381
	# restaurants	0.020	-0.002	0.028	-0.018	**0.303**	**0.116**	0.049
	# cafés	**0.727**	**0.345**	0.015	**0.450**	0.103	-0.417	0.058
	Sales amount	**-1.170**	**-0.673**	0.063	-0.587	-0.144	**-1.170**	-0.655

LUM, job density and the number of restaurants around the household negatively relate to others, while population density, the number of liquor stores and sales amount around the household positively relate to others. LUM, job density, residential density, street density and the number of traditional restaurants and cafés are positively associated to others in different geographical scales.

5.2 Discussion of MNL Results

The MNL models show that individual and household socioeconomic factors, such as gender, age, status as a student, annual income, household ownership, vehicle ownership and household composition, are the most significant factors in deciding the travel mode, either carpool or others. The transit distance and CBD distance are shown to be significant when not combined with socioeconomic factors. The built-environment factors around households and stores, such as LUM, population density, job density, residential density, street density, the number of points of interest and sales amount, are all associated with the shopping trip travel mode choice, although their significance could not be applied to all geographical scales.

The household annual income under $35,000, and the number of adults and children are consistently positive predictors for carpool, while the number of bicycles and the families who rent are consistently positively predictors for others. This result indicates that the household economic condition does influence people's travel mode choice, especially for those whose annual income is under $35,000 and prefer not driving alone for shopping, and they would share a vehicle or take other travel modes. In addition, the larger the family size (the number of adults and children) is, the higher possibility for this family to share vehicles for shopping. In addition, if there are more vehicles in the household, they are less likely to choose carpool or taking other modes; driving alone would be the best option for them. On the contrary, families having more bicycles, which may suggest they have a healthier life attitude, would choose travel modes other than carpool, which is also consistent with the findings in the descriptive results. If the house is rented, its members are likely to choose other travel modes; however, if living in current locations for a longer period, people prefer not to choose other travel modes for shopping. We may conclude that economic and health considerations are potential factors behind people's preference on shopping trip travel modes.

The MNL results also imply that a longer distance to the CBD could lead people to carpool more and take other travel modes less. This is consistent with the results after adding household built-environment factors at the census block level. In addition, the long distance to the closest public transit stations (both bus stops and light rail stations) prevent people from choosing other travel modes. We may also say that the distances to the aggregated shopping area (CBD) and the nearest transit stations are the most important consideration for shoppers.

The geographical scale is obvious when the built-environment factors are taken into consideration. Table 5.6 showed the correlation between the built-environment and the geographical scales The store built-environment has a greater effect on the travel mode choice than the household built-environment, since there were more significant factors around the store. Most significant household built-environment factors are within the census level, which means the community boundaries (census level) have a larger impact than the distance consideration (buffer level) around the household, whereas both the census level and the buffer level are important for factors around the store.

The census block group is the level whose number of significant factors are the most in both the model for carpooling and the model for others: residential density, street density, sales amount, numbers of restaurants, quick-services, traffic nodes, and liquor stores are all significant factors. Followed by the census block group level is the census block level, whose significant factors include the street density around both the household and the store, as well as the job density, the number of convenience stores, and the sales amount around the store, and these factors are all significant in the model for carpooling, either positively or negatively. The census tract, however, is the only geographical scale which demonstrates the significance of the population density.

Compared with network buffers, straight-line buffers have a better performance in measuring the significant factor in the model for others, which is different from our previous expectation. Around the household, both half-mile straight-line buffer and half-mile network buffer show LUM is negatively related to others. Around the store, job density, the number of cafés, and the sales amount are shown their significance in the model for others in both half-mile and one-mile straight-line buffers; meanwhile, only one-mile network buffer has the similar effect. Nevertheless, the network buffers still show that the residential density, the street density, and the number of convenience stores are significant in the model for others.

Table 5.6: MNL results for built-environments and geographical scales
(Highlighted cells mean the significance of corresponding factors in a certain scale level
C: significant in the model for carpool; O: significant in the model for others.)

	Variables	straight-line		network buffer		census level		
		0.5 mi	1 mi	0.5 mi	1 mi	block	BG	tract
Household built-environment	LUM	O		O				
	Population density							O
	Job density			O				
	Residential density						C	
	Street density					C		
	# restaurants		O				CO	
	# bus stops							
	# cafés							
	# quick-services						C	
	# convenience stores				C			O
	# liquor stores						O	
	# traffic nodes					CO		
	Sales amount					O		
Store built-environment	LUM					O		
	Population density							
	Job density	O	O		O	C		
	Residential density			O				
	Street density	C			O	C	C	
	# restaurants					O	O	
	# bus stops							
	# cafés	O	O		O			
	# quick-services		C		C			
	# convenience stores					C		
	# liquor stores						C	C
	# traffic nodes							
	Sales amount	O	O			C	O	

Several factors exert their effects on the travel mode choice in more than two geographical scales, and we call these factors as strong factors. LUM, and the number of restaurants, and the number of convenience stores around the household are strong factors for either the model of carpool in the model for others. The job density around the store is a continuously positive significant factor on the model of others, especially in buffer levels. The street density around the store is a continuously positive significant factor on the model of carpool. The number of cafés and sales amount are strong factor in the model for others, and the number of quick-services and liquor stores are strong factors in the model for carpool.

Chapter 6 Conclusion

This study investigates travel mode choices in urban shopping trips in Salt Lake County, Utah. As in previous studies of urban shopping trips, individual socio-demographic characteristics and built-environment are key factors behind respondents' mode choices. The most common travel modes include driving alone, ride-sharing, taking public transit, biking, and walking. The various geographical scales raise the question of whether the association between travel mode choices and socio-demographic characteristics and built-environment factors is sensitive to different spatial issues such as scales and boundary effects (Fan et al., 2014; Yamada et al., 2012). Data in this study was derived from the 2012 Utah Travel Survey conducted by the Utah Department of Transportation (UDOT). The empirical methodology employed included both descriptive analysis and MNL regression. Specifically, mixed land use measures and the densities of street, population, job and stores were examined in relationship to travel mode choices in shopping trips at seven different spatial scales. The study results indicate that socioeconomic factors, as well as household and store built-environment factors, do have influence on shoppers' travel mode choices, and geographical scales matters. Furthermore, results show that the individual and household background is the most significant factor driving people's travel mode choices. The distance factors and the surrounding environments of the household or the store also have a strong impact on customers' travel mode choices.

The descriptive analysis and the MNL results demonstrate that there are numerous positive significant factors on choosing other transportation modes, other than driving alone and carpooling. These factors include families who rent, families whose annual income is under $35,000, status as a student, the number of bicycles in one household, the population density around the household in the census tract level, the numbers of traditional restaurants around the store in the one-mile buffer circle, the number of liquor stores around the household in the block group level, and the number of traffic nodes around the household in the census block group level. The factors having a negative impact on choosing other modes, such as walking, biking and riding public transit, are the number of vehicles in one household, the CBD distance in the census block level, LUM around the store, the number of traditional restaurants in the census block group level, the number of convenience stores in the census tract level, etc.

Results also show factors affecting people's choice of carpooling for the shopping trip. Positive determinants are females aged between 25 to 34, families whose annual income is under $35,000, the number of adults and children in one household, street density around the household and the store, the number of convenience stores around the household, the number of quick-services around the store, the sales amount around the store and so on. Negative factors include the number of vehicles in one household, the residential density around the household, the job density around the store, and the number of liquor stores around the store.

Secondly, this study examines how accessibility affects shopping trip travel mode choices. If there is a relatively long distance between household and the selected stores, the nearest public transit stations, and the downtown CBD, customers may not choose walking, biking, or public transit. People appear to share their private vehicles with others, either family members or friends, for shopping if the household-CBD distance is relatively long.

This study further investigates how geographical scale matters when examining the association between people's shopping trip travel mode choice and built-environment, especially for the street density and the sales amount. The results show that two buffer criteria, straight-line buffer and network buffer, exert similar functions for job density. In addition, the smaller the scale the study chooses, the more factors would be significant. For instance, the half-mile buffer circle has more significant factors influencing choices than both the one-mile buffer and the census block group scale, especially when people select other transportation modes.

Although other studies suggest that the diversity factor or the land use mix index is a significant built-environment factor in the travel behavior analysis (Cervero, 1996; Ewing and Cerveco, 2010), this factor in our research only demonstrates its significance in the census block level, and it is insignificant on any other geographical scales. This finding may reflect the controversial debate on whether the land use mix affects the travel mode choices. Some past studies did find the land use mix was not a significant factor in transportation study (Duncan et al., 2010).

Our findings could also provide a good reference for policy-makers in urban transportation planning and healthy communities. The vehicle amount is growing in the US; the urban planner, nevertheless, should consider a way to encourage people to travel by sharing private vehicles with others, walking or utilizing more public transit, such as bus and light rail. In order to achieve that goal, the city authorities may consider re-designing the surrounding environment of residences and shopping stores. For example, they could increase the neighborhood diversity and walking pleasure, improve public transit systems, and allow more small attractions such as restaurants, cafes, quick-service, etc. to be built around a large market, which may lower customers' driving frequency. Findings related to distances from respondents' households to the reported stores, the nearest public transit stations and the CBD also show that future transportation planning in Salt Lake County should focus on the improvement of accessibility to urban business centers through public transportation.

There are some additional questions that are beyond the scope of this research. First, this study has been constrained by the fact that the household travel diary survey is not specifically designed for the shopping trip travel behavior investigation. For example, some important indicators about shopping trips such as shopping frequency, travel expense and shopping cost, were not included in this survey. The research could be improved by using dedicated surveys of shoppers. Second, this research only focuses on the home-based shopping trip and neglects the work-based shopping trip, which is another interesting topic for future investigations.

Lastly, the study points to a number of other interesting topics for future research. An initial study might investigate the influence of personal shopping attitude on travel mode choice for the shopping trip. This thesis comprehensively explores objective factors behind people's travel mode choice, while these choices would also be affected by customers' subjective preferences, such as people's health attitude, body exercise frequency, cognitive distance between household and primary stores, etc.

Second, analyzing the impact of the current and future city development plan is another interesting topic for future research. The transportation system of Salt Lake County is unique but has become an exemplary case of public transit development in the United States (Farber et al., 2014). Therefore, the tendency of people's travel mode choice would have been influenced by this transition.

Third, additional work focusing on the impact of store characteristics is promising. This research successfully reveals that the total sales amount close to households and stores has a strong impact on shoppers' travel mode choices: the higher the total sales amount around a store, the more likely the customers drive alone or carpool, since riding a car would be most convenient if they want to get a large amount of goods in one shopping trip. Other related factors, such as store size, number of employees, price level, number of parking stalls, and mixture of stores in the neighborhood, could also be examined in future studies.

References

Bamberg, S., & Schmidt, P. (2010). Basic and Applied Social Psychology Choice of Travel Mode in the Theory of Planned Behavior : The Roles of Past Behavior , Habit , and Reasoned Action Choice of Travel Mode in the Theory of Planned Behavior : The Roles of Past Behavior , Habit , and Reaso, (October 2014), 37–41. doi:10.1207/S15324834BASP2503

Bawa, K., & Ghosh, A. (1999). A Model of Household Grocery Shopping Behavior. *Marketing Letters*, *10*(2), 149–160. doi:10.1023/A:1008093014534

Bell, D. R., Ho, T., & Tang, C. S. (1998). Determining Where Variable to of Shop : Fixed and Costs Shopping. *Journal of Marketing Research*, *35*(3), 352–369.

Bhat, C. R. (1998a). ANALYSIS OF TRAVEL MODE AND DEPARTURE TIME CHOICE FOR. *Transportation Research Part B: Methodological*, *32*(6), 361–371.

Bhat, C. R. (1998b). Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research Part B: Methodological*, *32*(6), 361–371. doi:10.1016/S0191-2615(98)00004-6

Blattberg, R., Buesing, T., Peacock, P., & Sen, S. (1978). Identifying the Deal Prone Segment. *Journal of Marketing Research*, *15*(3), 369–377.

Blaylock, J. R. (1989). An economic model of grocery shopping frequency. *Applied Economics*, *21*(6), 843–852. doi:10.1080/758520277

Boarnet, M., & Sarmiento, S. (1998). Can Land-use Policy Really Affect Travel Behaviour? A Study of the Link between Non-work Travel and Land-use Characteristics. *Urban Studies*, *35*(7), 1155–1169. doi:10.1080/0042098984538

Boivin, J. (2008). *Bus Transit Use and Grocery Shopping in Detroit.*

Brown, B. B., Yamada, I., Smith, K. R., Zick, C. D., Kowaleski-Jones, L., & Fan, J. X. (2009). Mixed land use and walkability: Variations in land use measures and relationships with BMI, overweight, and obesity. *Health & Place*, *15*(4), 1130–41. doi:10.1016/j.healthplace.2009.06.008

Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., & Sallis, J. F. (2009). Measuring the built environment for physical activity: state of the science. *American Journal of Preventive Medicine*, *36*(4 Suppl), S99–123.e12. doi:10.1016/j.amepre.2009.01.005

Cao, X., & Mokhtarian, P. L. (2005). How do individuals adapt their personal travel? Objective and subjective influences on the consideration of travel-related strategies for San Francisco Bay Area commuters. *Transport Policy*, *12*(4), 291–302. doi:10.1016/j.tranpol.2005.03.003

Carrasco, N. (2008). *Deciding where to shop : disaggregate random utility destination choice modeling of grocery shopping in canton Zurich*. Retrieved from http://www.ivt.ethz.ch/docs/students/sa223.pdf

Cervero, R. (1996). Mixed land-uses and commuting: Evidence from the American Housing Survey. *Transportation Research Part A: Policy and Practice*, *30*(5), 361–377. doi:10.1016/0965-8564(95)00033-X

Cervero, R. (2002). Built environments and mode choice: toward a normative framework. *Transportation Research Part D: Transport and Environment*, *7*(4), 265–284. doi:10.1016/S1361-9209(01)00024-4

Cervero, R., & Gorham, R. (1995). Commuting in Transit Versus Automobile Neighborhoods. *Journal of the American Planning Association*, *61*(2), 210–225. doi:10.1080/01944369508975634

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219. doi:10.1016/S1361-9209(97)00009-6

Cervero, R., & Radisch, C. (1996). Travel choices in pedestrian versus automobile oriented neighborhoods. *Transport Policy*, *3*(3), 127–141. doi:10.1016/0967-070X(96)00016-9

Chen, C., Gong, H., & Paaswell, R. (2008). Role of the built environment on mode choice decisions: additional evidence on the impact of density. *Transportation*, *35*(3), 285–299. doi:10.1007/s11116-007-9153-5

Christian, H. E., Bull, F. C., Middleton, N. J., Knuiman, M. W., Divitini, M. L., Hooper, P., … Giles-Corti, B. (2011). How important is the land use mix measure in understanding walking behaviour? Results from the RESIDE study. *The International Journal of Behavioral Nutrition and Physical Activity*, *8*(1), 55. doi:10.1186/1479-5868-8-55

Clifton, K. J. (2004). Mobility Strategies and Food Shopping for Low-Income Families: A Case Study. *Journal of Planning Education and Research*, *23*(4), 402–413. doi:10.1177/0739456X04264919

Clifton, K., Morrissey, S., & Ritter, C. (2012). Mode Choice and Consumer Spending: An Examination of Grocery Store Trips. *91st Annual Meeting of the Transportation Research …*, 22–26. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Mode+Choice+and+Consumer+Spending:+An+Examination+of+Grocery+Store+Trips#0

Coombes, M. G., Green, A. E., & Openshaw, S. (1986). An Efficient Algorithm to Generate Official Statistical Reporting Areas: The Case of the 1984 Travel-to-Work Areas Revision in Britain. *Journal of the Operational Research Society*, *37*(10), 943–953. doi:10.1057/jors.1986.163

Coveney, J., & O'Dwyer, L. a. (2009). Effects of mobility and location on food access. *Health & Place*, *15*(1), 45–55. doi:10.1016/j.healthplace.2008.01.010

Crane, R., & Crepeau, R. (1998). Does neighborhood design influence travel?: A behavioral analysis of travel diary and GIS data. *Transportation Research Part D: Transport and Environment*, *3*(4), 225–238. doi:10.1016/S1361-9209(98)00001-7

Davies, R. B., & Pickles, A. R. (1987). A joint trip timing store-type choice model for grocery shopping, including inventory effects and nonparametric control for omitted variables. *Transportation Research Part A: General*, *21*(4-5), 345–361. doi:10.1016/0191-2607(87)90054-9

Duncan, M. J., Winkler, E., Sugiyama, T., Cerin, E., DuToit, L., Leslie, E., & Owen, N. (2010). Relationships of land use mix with walking for transport: do land uses and geographical scale matter? *Journal of Urban Health : Bulletin of the New York Academy of Medicine*, *87*(5), 782–95. doi:10.1007/s11524-010-9488-7

Elmi, A., Badoe, D., & Miller, E. (2015). Transferability Analysis of Work-Trip-Distribution Models. *Transportation Research Record: Journal of the Transportation Research Board*. Retrieved from http://trrjournalonline.trb.org/doi/abs/10.3141/1676-21

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association*, *76*(3), 265–294. doi:10.1080/01944361003766766

Fan, J. X., Hanson, H. A., Zick, C. D., Brown, B. B., Kowaleski-Jones, L., & Smith, K. R. (2014). Geographic scale matters in detecting the relationship between neighbourhood food environments and obesity risk: an analysis of driver license records in Salt Lake County, Utah. *BMJ Open*, *4*(8), e005458. doi:10.1136/bmjopen-2014-005458

Farber, S., Bartholomew, K., Li, X., Páez, A., & Nurul Habib, K. M. (2014). Assessing social equity in distance based transit fares using a model of travel behavior. *Transportation Research Part A: Policy and Practice*, *67*, 291–303. doi:10.1016/j.tra.2014.07.013

Finkelstein, E. A., Fiebelkorn, I. C., & Wang, G. (2004). State-level estimates of annual medical expenditures attributable to obesity. *Obesity Research*, *12*(1), 18–24. doi:10.1038/oby.2004.4

Forinash, C. V, & Koppelman, F. S. (1993). Application And Interpretation of Nested Logit Models of Intercity Mode Choice. *Transportation Research Record*, (1413). Retrieved from http://trid.trb.org/view.aspx?id=385097

Frank, L. D., Andresen, M. A., & Schmid, T. L. (2004). Obesity relationships with community design, physical activity, and time spent in cars. *American Journal of Preventive Medicine*, *27*(2), 87–96. doi:10.1016/j.amepre.2004.04.011

Frank, L. D., Sallis, J. F., Conway, T. L., Chapman, J. E., Saelens, B. E., & Bachman, W. (2006). Many Pathways from Land Use to Health: Associations between Neighborhood Walkability and Active Transportation, Body Mass Index, and Air Quality. *Journal of the American Planning Association*, *72*(1), 75–87. doi:10.1080/01944360608976725

Frank, L., & Pivo, G. (1994). Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. *Transportation*

Research Record. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Impacts+of+Mixed+Use+and+Density+on+Utilization+of+Three+Modes+of+Travel+:+Single-Occupant+Vehicle+,+Transit+,+and+Walking#0

Frisbie Jr., G. A. (1980). Ehrenberg's Negative Binomial Model Applied to Grocery Store Trips on JSTOR. *Journal of Marketing Research*, *17*(3), 385–390. Retrieved from http://www.jstor.org/stable/3150539?seq=1#page_scan_tab_contents

Garvill, J., Marell, A., & Nordlund, A. (2003). Effects of increased awareness on choice of travel mode. *Transportation*, *30*(1), 63–79. doi:10.1023/A:1021286608889

Gauvin, L., Robitaille, E., Riva, M., McLaren, L., Dassa, C., & Potvin, L. Conceptualizing and operationalizing neighbourhoods: the conundrum of identifying territorial units. *Canadian Journal of Public Health = Revue Canadienne de Santé Publique*, *98 Suppl 1*, S18–26. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18047157

Geiger, S. (2007). Exploring night-time grocery shopping behaviour. *Journal of Retailing and Consumer Services*, *14*(1), 24–34. doi:10.1016/j.jretconser.2006.03.001

Guo, J., Bhat, C., & Copperman, R. (2007). Effect of the Built Environment on Motorized and Nonmotorized Trip Making: Substitutive, Complementary, or Synergistic? *Transportation Research Record*, *2010*(1), 1–11. doi:10.3141/2010-01

Handy, S. (1996). *Non-Work Travel of Women: Patterns, Perceptions and Preferences*. *Women's Travel Issues Second National Conference*. doi:10.1037/e736202011-019

Handy, S., Boarnet, M., Ewing, R., & Killingsworth, R. (2002). How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine*, *23*(2 SUPPL. 1), 64–73. doi:10.1016/S0749-3797(02)00475-0

Hess, P., & Moudon, A. (1999). Site design and pedestrian travel. *Transportation Record Record 1674*, (99). Retrieved from http://trb.metapress.com/index/07463m257t231n13.pdf

Hoffman, S. D., & Duncan, G. J. (1988). A Comparison of Choice-Based Multinomial and Nested Logit Models: The Family Structure and Welfare Use Decisions of Divorced or Separated Women. *The Journal of Human Resources*, *23*(4), 550–562. doi:10.2307/145813

Horowitz, J. L. (1993). Semiparametric estimation of a work-trip mode choice model. *Journal of Econometrics*, *58*(1-2), 49–70. doi:10.1016/0304-4076(93)90113-J

Innes, J., Ircha, M., & Badoe, D. (1990). Factors affecting automobile shopping trip destinations. *Journal of Urban Planning and …*, *116*(3), 126–136. Retrieved from http://ascelibrary.org/doi/abs/10.1061/(ASCE)0733-9488(1990)116:3(126)

Jiao, J., Moudon, A. V., & Drewnowski, A. (2011). Grocery Shopping How Individuals and Built Environments Influence Choice of Travel Mode. *Transportation Research Record*, *2230*, 85–95. doi:10.3141/2230-10

Jiao, J., Moudon, A. V., & Drewnowski, A. (2012). Influence of Individual Characteristics and the Built Environment on Grocery Shopping Travel Frequency. *Transportation Research Board 91st …*. Retrieved from http://trid.trb.org/view.aspx?id=1129932

Kitamura, R., Mokhtarian, P. L., & Laidet, L. (1997). A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. *Transportation*, *24*(2), 125–158. doi:10.1023/A:1017959825565

Klöckner, C. A., & Matthies, E. (2004). How habits interfere with norm-directed behaviour: A normative decision-making model for travel mode choice. *Journal of Environmental Psychology*, *24*(3), 319–327. doi:10.1016/j.jenvp.2004.08.004

Künzli, N., Kaiser, R., Medina, S., Studnicka, M., Chanel, O., Filliger, P., … Sommer, H. (2000). Public-health impact of outdoor and traffic-related air pollution: a European assessment. *Lancet*, *356*(9232), 795–801. doi:10.1016/S0140-6736(00)02653-2

Law, R. (1999). Beyond "women and transport": towards new geographies of gender and daily mobility. *Progress in Human Geography*, *23*(4), 567–588. doi:10.1191/030913299666161864

Lee, C., Zhu, X., Yoon, J., & Varni, J. W. (2013). Beyond distance: children's school travel mode choice. *Annals of Behavioral Medicine : A Publication of the Society of Behavioral Medicine*, *45 Suppl 1*, S55–67. doi:10.1007/s12160-012-9432-z

Liao, F. H., Farber, S., & Ewing, R. (2014). Compact development and preference heterogeneity in residential location choice behaviour: A latent class analysis. *Urban Studies*. doi:10.1177/0042098014527138

Litman, T. (2009). *Transportation cost and benefit analysis*. Retrieved from http://www.researchgate.net/profile/Todd_Litman/publication/235360398_Transportation_Cost_and_Benefit_Analysis_Techniques_Estimates_and_Implications/links/544a94ca0cf2d6347f401152.pdf

Liu, G. (2007). A behavioral model of work-trip mode choice in Shanghai. *China Economic Review*, *18*(4), 456–476. doi:10.1016/j.chieco.2006.09.001

Matthies, E., Kuhn, S., & Klockner, C. A. (2002). Travel Mode Choice of Women: The Result of Limitation, Ecological Norm, or Weak Habit? *Environment and Behavior*, *34*(2), 163–177. doi:10.1177/0013916502034002001

McDonald, N. C. (2008). Household interactions and children's school travel: the effect of parental work patterns on walking and biking to school. *Journal of Transport Geography*, *16*(5), 324–331. doi:10.1016/j.jtrangeo.2008.01.002

McDonald, N. C., Deakin, E., & Aalborg, A. E. (2010). Influence of the social environment on children's school travel. *Preventive Medicine*, *50 Suppl 1*, S65–8. doi:10.1016/j.ypmed.2009.08.016

McFadden, D. (1977). *Quantitative methods for analyzing travel behavior of individuals: some recent developments*. Retrieved from http://core.kmi.open.ac.uk/download/pdf/6448852.pdf

Moudon, A. V., Design, U., & Drewnowski, A. (2011). Grocery shopping : How individuals and built environments influence travel mode choice.

Müller, S., Tscharaktschiew, S., & Haase, K. (2008). Travel-to-school mode choice modelling and patterns of school choice in urban areas. *Journal of Transport Geography*, *16*(5), 342–357. doi:10.1016/j.jtrangeo.2007.12.004

Nagendra, H. (2002). Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Applied Geography*, *22*(2), 175–186. doi:10.1016/S0143-6228(02)00002-4

Openshaw, S. (1977). A Geographical Solution to Scale and Aggregation Problems in Region-Building, Partitioning and Spatial Modelling. *Transactions of the Institute of British Geographers*, *2*(4), 459. doi:10.2307/622300

Putrevu, S., & Ratchford, B. T. (1997). A model of search behavior with an application to grocery shopping. *Journal of Retailing*, *73*(4), 463–486. doi:10.1016/S0022-4359(97)90030-0

Recker, W. W., & Kostyniuk, L. P. (1978). Factors influencing destination choice for the urban grocery shopping trip. *Transportation*, *7*(1), 19–33. doi:10.1007/BF00148369

Rodrigues, C., Popovich, N., & Handy, S. L. (2014). Bicyclists as Consumers: Mode Choice and Spending Behavior in Downtown Davis, CA. *Transportation Research Board 93rd Annual Meeting*, (January), 15 –31. Retrieved from http://trid.trb.org/view.aspx?id=1288626

Rosenbloom, S. (1978). The need for study of women's travel issues. *Transportation*, *7*(4), 347–350. doi:10.1007/BF00168035

Santos, A., McGuckin, N., Nakamoto, H. Y., Gray, D., & Liss, S. (2011). *Summary of Travel Trends: 2009 National Household Travel Survey*. Retrieved from http://trid.trb.org/view.aspx?id=1107370

Scheer, B. (2012). The Utah Model: Lessons for Regional Planning. *Brookings Mountain West Publications*. Retrieved from http://digitalscholarship.unlv.edu/brookings_pubs/21

Scheiner, J., & Holz-Rau, C. (2007). Travel mode choice: affected by objective or subjective determinants? *Transportation*, *34*(4), 487–511. doi:10.1007/s11116-007-9112-1

Scheiner, J., & Holz-Rau, C. (2012). Gendered travel mode choice: a focus on car deficient households. *Journal of Transport Geography*, *24*, 250–261. doi:10.1016/j.jtrangeo.2012.02.011

Schmöcker, J.-D., Quddus, M. A., Noland, R. B., & Bell, M. G. H. (2008). Mode choice of older and disabled people: a case study of shopping trips in London. *Journal of Transport Geography*, *16*(4), 257–267. doi:10.1016/j.jtrangeo.2007.07.002

Schwanen, T. (2004). The determinants of shopping duration on workdays in The Netherlands. *Journal of Transport Geography*, *12*, 35–48. doi:10.1016/S0966-6923(03)00023-1

Schwanen, T., & Mokhtarian, P. L. (2005). What affects commute mode choice: neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography*, *13*(1), 83–99. doi:10.1016/j.jtrangeo.2004.11.001

Shannon, C., & Weaver, W. (1949). The Mathematical Theory of Information.

Simpson, E. H. (1949). Measurement of diversity. *Nature*, *163*(688).

Smith, M. F., & Carsky, M. L. (1996). Grocery shopping behavior consumers, *3*(2), 73–80.

Su, F., & Bell, M. G. H. (2009). Transport for older people: Characteristics and solutions. *Research in Transportation Economics*, *25*(1), 46–55. doi:10.1016/j.retrec.2009.08.006

Wen, C.-H., & Koppelman, F. S. (2001). The generalized nested logit model. *Transportation Research Part B: Methodological*, *35*(7), 627–641. doi:10.1016/S0191-2615(00)00045-X

Wilde, P., & Ranney, C. (2000). The Monthly Food Stamp Cycle: Shooping Frequency and Food Intake Decisions in an Endogenous Switching Regression Framework. *American Journal of Agricultural …*, *82*(February), 200–213. Retrieved from http://ajae.oxfordjournals.org/content/82/1/200.short

Yamada, I., Brown, B. B., Smith, K. R., Zick, C. D., Kowaleski-Jones, L., & Fan, J. X. (2012). Mixed land use and obesity: an empirical comparison of alternative land use measures and geographic scales. *The Professional Geographer : The Journal of the Association of American Geographers*, *64*(2), 157–177. doi:10.1080/00330124.2011.583592

Zhang, J., & Timmermans, H. (2015). A Scobit-Based Travel Mode Choice Model. In D. R. Deeter-Schmelz (Ed.), *Proceedings of the 2010 Academy of Marketing Science (AMS) Annual Conference SE - 88* (pp. 147–151). Springer International Publishing. doi:10.1007/978-3-319-11797-3_88

Zhu, X., & Lee, C. (2008). Walkability and safety around elementary schools economic and ethnic disparities. *American Journal of Preventive Medicine*, *34*(4), 282–90. doi:10.1016/j.amepre.2008.01.024

Appendix A Selected Questionnaire Questions

A.1 Household Data

1. **Household size**
   A. 1 person household
   B. 2-person household
   C. 3-person household
   D. 4-person household
   E. 5-person household
   F. 6+ person household
2. **Number of adults in household**
3. **Number of children in household**
4. **Household life cycle**
   A. Household with no children or retirees
   B. Household with children but no retirees
   C. Household with retirees
5. **Number of adult workers in household**
6. **Household annual income**
   A. Under $35,000
   B. $35,000 - $49,999
   C. $50,000 - $99,999
   D. $100,000 or more
7. **Number of motorized vehicles in household**
   A. Zero vehicle household
   B. 1 vehicle household
   C. 2 vehicle household
   D. 3+ vehicle household
8. **Number of adult bikes**
   A. 0 (none)
   B. 1 bicycle
   C. 2 bicycles
   D. 3 bicycles
   E. 4 bicycles
   F. 5 bicycles
   G. 6 or more bicycles
9. **If children in household: Number of children's bikes**

A. 0 (none)
B. 1 bicycle
C. 2 bicycles
D. 3 bicycles
E. 4 bicycles
F. 5 bicycles
G. 6 or more bicycles

10. **Rent or own home**
    A. Rent
    B. Own/Buying (paying mortgage)
11. **Number of years lived at residence**
    A. Less than 1 year
    B. 1-5 years
    C. 6-10 years
    D. 11-15 years
    E. 16-20 years
    F. More than 20 years

A.2 Personal Data

1. **Age category**
    A. Under 5 years old
    B. 5 - 15
    C. 16 - 17
    D. 18 - 24
    E. 25 - 34
    F. 35 - 44
    G. 45 - 54
    H. 55 - 64
    I. 65 - 74
    J. 75 - 84
    K. 85 or older
2. **Gender**
3. **Employment status**
    A. Employed full-time
    B. Employed part-time
    C. Self-employed (full or part-time)
    D. Student, not employed or employed less than 25 hrs/week

E. Student, employed 25+ hrs/week
F. Homemaker
G. Retired
H. Not currently employed

4. **Education**
   A. Less than high school
   B. High school graduate
   C. Some college
   D. Vocational/technical training
   E. Associates degree
   F. Bachelor's degree
   G. Graduate/post-graduate degree
5. **Hispanic**
   A. Yes
   B. No
   C. Prefer not to answer
6. **Race**
   A. African American or Black
   B. American Indian or Alaskan Native
   C. Asian
   D. White or Caucasian
7. **Frequency: Ride transit**
   A. 6-7 days a week
   B. 5 days a week
   C. 4 days a week
   D. 3 days a week
   E. 2 days a week
   F. 1 day a week
   G. A few times per month
   H. Less than monthly
   I. Never

A.3 Adult Travel Diary

1. **Trip purpose**
   A. Home-based work (HBW)
   B. Home-based school (HBSch)
   C. Home-based shopping (HBShp)

D. Home-based personal business (HBPb)
E. Home-based other (HBO)
F. Non home-based work (NHBW)
G. Non home-based non work (NHBNW)

2. **Trip mode**
A. Auto drive alone
B. Auto occupancy 2
C. Auto occupancy 3+
D. Transit
E. Walk
F. Bicycle
G. Other

A.4 Screenshots of Survey Webpages

Figure A.1: The welcome page

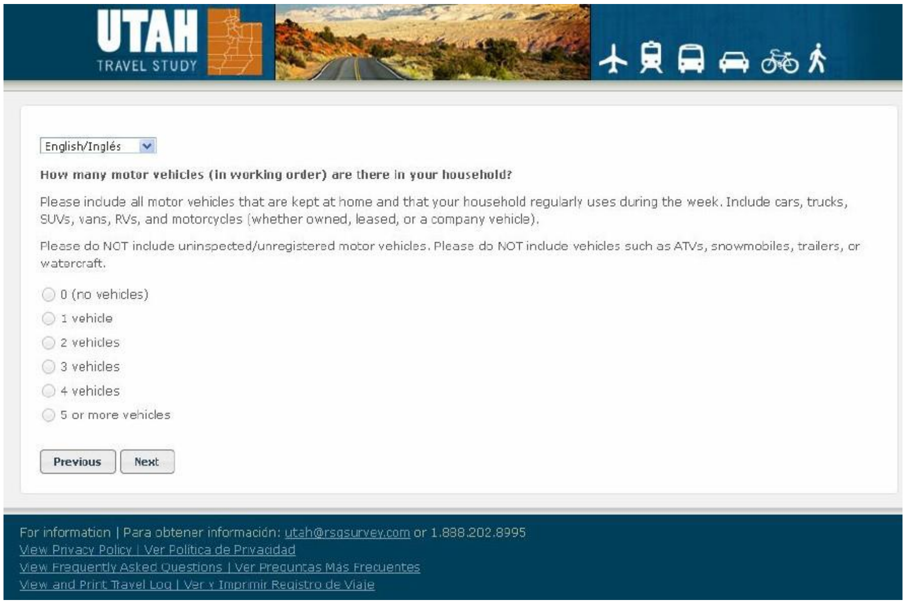
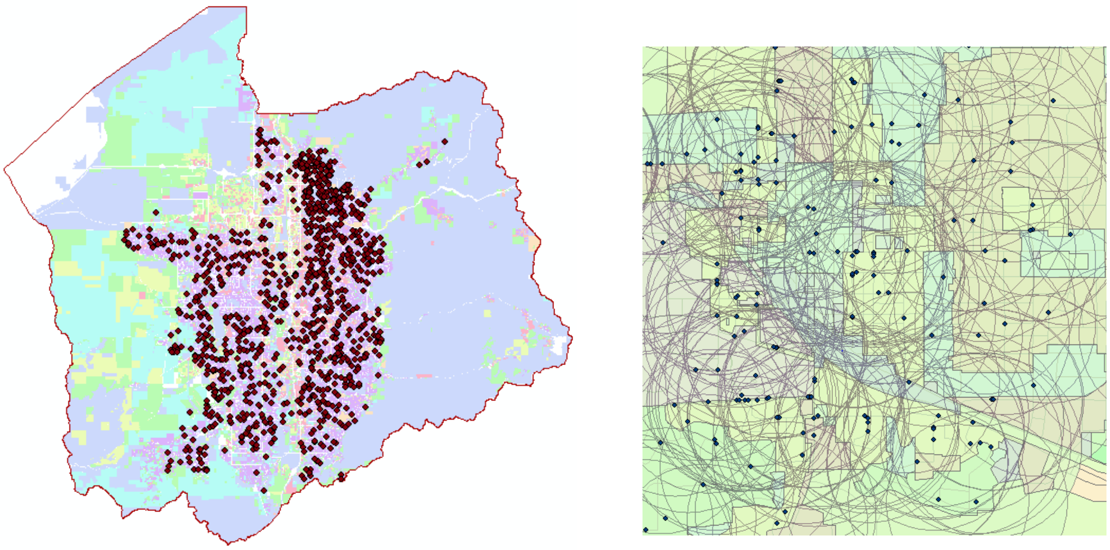


Figure A.2: A questionnaire sample page

Appendix B LUM Calculator

B.1 Background

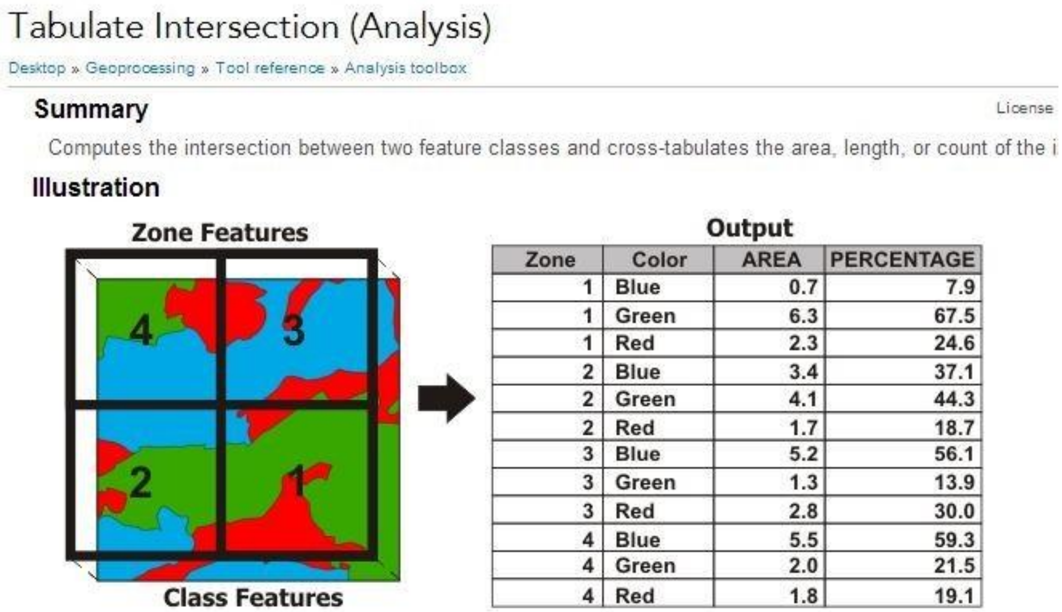


(a) Points (household) and land use types of and land use type polygon Salt Lake County, UT

(b) Intersecting 500 m buffer area

Figure B.1: LUM calculation flow chart based on Arcpy

It is a common subject to find land use condition within two overlapping polygons. Consider the following scenario: a wild animal protection organization wants to determine which land cover class is the largest in 500 m, 1000 m, and 1500 m buffers (or half mile, one mile buffer) around deer vehicle collision points at Latah County, ID. Normally, in this kind of case, what the organization has is the position of collision points, such as the conjunction between Mountain View and B Street even without GPS latitude and longitude. Therefore, in order to get the land use condition within 500 m buffer to a certain accident point, geocoding the position will be the first step, or to find the geographic location from the address. Next, they need to insect local land cover with the buffer circle of collision position. Once they get the area of each land cover within 500 m buffer, they can make a decision.



Tabulate Intersection (Analysis)

Desktop » Geoprocessing » Tool reference » Analysis toolbox

Summary License

Computes the intersection between two feature classes and cross-tabulates the area, length, or count of the i

Illustration

Zone	Color	AREA	PERCENTAGE
1	Blue	0.7	7.9
1	Green	6.3	67.5
1	Red	2.3	24.6
2	Blue	3.4	37.1
2	Green	4.1	44.3
2	Red	1.7	18.7
3	Blue	5.2	56.1
3	Green	1.3	13.9
3	Red	2.8	30.0
4	Blue	5.5	59.3
4	Green	2.0	21.5
4	Red	1.8	19.1

Figure B.2: The screen shot of Tabulate Intersection tool in the ArcGIS Help

Similar scenario will be to calculate the land use mix index (Formula 1). Briefly speaking, we need to find the proportion of the area covered by every land use type i, agriculture for instance, against the summed area for land use classes of interest, and n is the number of land use classes of interest. (Figure B.1)

Unfortunately, however, when intersecting two polygons together, especially one overlaps the other, tens of thousands of tiny areas will be generated. Even though there is a tool - Tabulate Intersection – in ArcGIS 10.2.2, it still cannot deal with the overlapping problem, which will consume lots of running time and low accuracy. (Figure B.2)

B.2 Programming

B.2.1 Land Use Type Calculation Tool

It will take more time to complete the task manually than to write, compile, and debug the program, therefore, I choose to create a new program. In order to fix the overlapping problem, I decide to select points one by one, and intersect its buffer area with the land use type polygon separately, which is shown in a flow-chart (Figure B.3). What is more, the input will be customer-friendly (Figure B.4 and B.5), since the user just provides the land use shapefile and the point address, even the Excel is acceptable, no need to geocode. Figure B.6 is part of the script for reference.

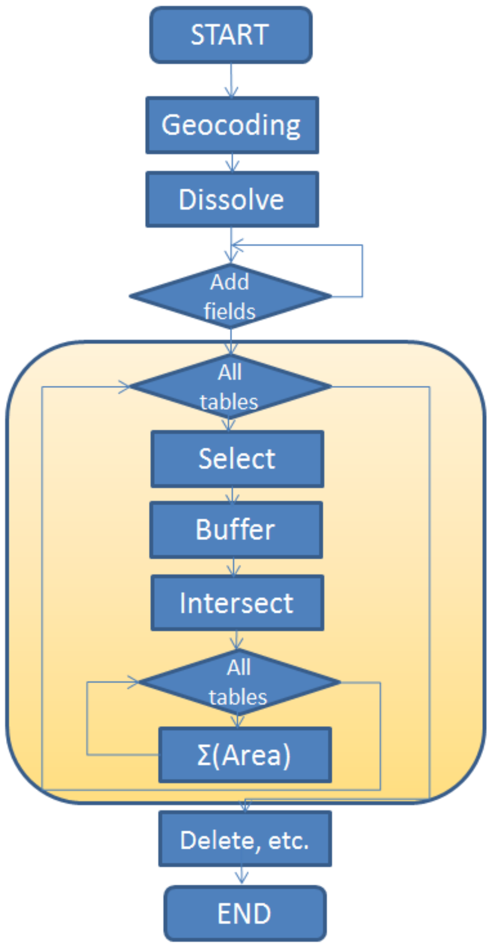


Figure B.3: The flow-chart (the yellow part is the LUM calculator)

Figure B.4: The toolbar for calculating land use mix index and street density

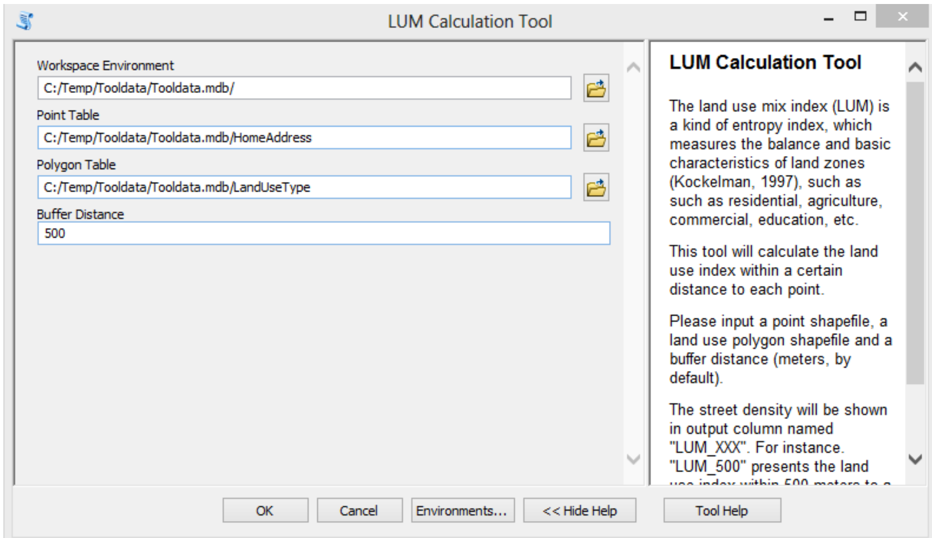


Figure B.5: The dialogue box for LUM calculator

```
# Select one address point
in_features = point
selectName = 'select_' + str(int(row.ID))
# Sample: select_1703182673
out_feature_class = inWS + selectName
firstPart = "[ID]"
secondPart = str(int(row.ID))
where_clause = firstPart + " = " + secondPart
# Sample: where_clause = '[ID] = 1703182673'
arcpy.Select_analysis(in_features, out_feature_class, where_clause)
Printall("Creating " + selectName + ' ...')
#
# Buffer this point
inFC_buffer = inWS + selectName
bufferName = 'buffer_' + str(int(row.ID)) + '_' + bufferDistance
# Sample: buffer_1703182673_100
outFC_buffer = inWS + bufferName
distance = bufferDistance
arcpy.Buffer_analysis(inFC_buffer, outFC_buffer, distance)
Printall("Creating " + bufferName + ' ...')
#
# Intersect this buffer with land use map
bufferFeature = bufferName
inFeatures = [bufferFeature, "LandUseType"]
intersectName = 'intersect_' + str(int(row.ID)) + '_' + bufferDistance
# Sample: intersect_1703182673_100
intersectOutput = intersectName
arcpy.Intersect_analysis(inFeatures, intersectOutput)
Printall("Creating " + intersectName + ' ...')
#
# Calculate land use areas for different address
Printall("Calculating land use area for " + str(int(row.ID)) + '\n')
intersectRows = arcpy.SearchCursor(inWS + intersectName)
for intersectRow in intersectRows:
    fieldName = intersectRow.Type + '_' + bufferDistance
    area = row.getValue(fieldName) + intersectRow.Shape_Area
    row.setValue(fieldName, area)
    rows.updateRow(row)
#
# Delete scratch files
arcpy.Delete_management (out_feature_class)
arcpy.Delete_management (outFC_buffer)
arcpy.Delete_management (intersectOutput)
```

Figure B.6: The snippet for LUM calculator

B.2.2 Geocoding

For the geocoding part, I use the Google Geocoding API, which provides a direct way to access a geocoder via an HTTP request. A Geocoding API request must be in a required form, and the geocoding responses are returned in the format indicates by the output flag within the URL request's path. However, the use of the Google Geocoding API is subject to a query limit of 2,500 geolocation requests per day. Therefore, a geocoded point shapefile is preferable.

For more information about Google Geocoding API, please visit: https://developers.google.com/maps/documentation/geocoding/. To get the python "requests" extension package, please visit: http://www.lfd.uci.edu/~gohlke/pythonlibs/. Figure B.7 is an example.

```
# Request location from Google_API and insert lng and lat to X and Y respectively
n = 0
rows = arcpy.UpdateCursor(tbLct)
for row in rows:
    n += 1
    Printall(str(n))
    str_address = row.Address + ', ' + row.City + ', ' + row.State + ' ' + str(row.Zip)
    # Format sample: "Mountain View and B Street, ID 83843"
    Printall(str_address)
    gapi = "http://maps.googleapis.com/maps/api/geocode/json?address="
    str_tofind = gapi + str_address.replace(' ', '+') + '&sensor=true'
    r = requests.get(str_tofind)
    result_geom = r.json['results'][0]['geometry']
    result_location = result_geom['location']
    if r.status_code == 200:
        Printall("latitude = " + format(result_location.get("lat"), '.15f'))
        Printall("longitude = " + format(result_location.get("lng"), '.15f') + "\n")
        row.latitude = format(result_location.get("lat"), '.15f')
        row.longitude = format(result_location.get("lng"), '.15f')
        rows.updateRow(row)
    else:
        row.X = 0
        row.Y = 0
```

Figure B.7: The snippet for geocoding

B.2.3 Street Density Calculation Tool

With a slight modification of tool 1, or the LUM calculator, we can also calculate the street density with a polygon. Street density or road density is the total length of the entire road network in a certain land area. Here are several home addresses, and we want to know the street density within 100 m buffer area to each address so that we can analyze the convenience of local transportation.

In order to avoid the influence of overlapping, a similar method to calculate the street density is to select each point one by one, and intersect only one point with line shapefile one time. FigureB.10is part of the script for reference.

B.3 Conclusion and Future Improvement

This appendix provides:

I. Calculation of the land use mix index (LUM) around a point in a certain buffer distance;

II. The total length of roads or streets within certain distance to a point;

III. One possible geocoding method with Google Geocoding API.

This programming can be improved in the following ways:

I. Combine SQL, Excel or `numpy` to enable this script to calculate the complete land use mix index (LUM);

II. Use 10.0 North America Geocode Service (ArcGIS Online) to geocode the points.

Appendix C Built-environment Factors

Variables	Description	Min	Max	Mean	Median	SD
Distance Factors						
OD dist.	mile	0.000	22.300	3.375	2.300	3.410
Public transit dist.	mile	0.000	6.832	0.413	0.272	0.512
CBD dist.	mile	0.381	26.114	10.814	10.782	5.889
Home Built-Environment (half-mile straight line buffer)						
LUM	land use mix	0.223	0.897	0.500	0.485	0.108
Population density	people per mile square	0.001	0.345	0.168	0.169	0.055
Job density	job per mile square	0.088	8.113	1.792	1.634	0.978
Residential density	household per mile	16.000	2775.000	1199.692	1177.000	518.726
Street density	mile per mile square	0.128	1.169	0.737	0.746	0.145
Restaurant	number of restaurants	0.000	79.000	5.422	3.000	8.440
Bus stop	number of bus stops	0.000	67.000	15.862	12.000	14.138
Cafe	number of cafes	0.000	6.000	0.257	0.000	0.597
Quick service	number of quick-service stores	0.000	21.000	1.842	1.000	2.997
Convenience store	number of convenience stores	0.000	5.000	0.912	1.000	1.063
Liquor stores	number of liquor stores	0.000	3.000	0.083	0.000	0.306
Traffic node	number of traffic nodes	4.000	337.000	155.599	153.000	43.405
Sales amount	dollar	0.000	229381.000	27419.891	7234.000	43903.374
Store Built-Environment (half-mile straight line buffer)						
LUM	land use mix	0.000	0.782	0.524	0.512	0.101
Population density	people per mile square	0.000	0.318	0.156	0.159	0.063
Job density	job per mile square	0.071	6.129	1.614	1.591	0.538
Residential density	household per mile	0.000	2613.000	1001.539	1030.500	561.371

(continued on next page)

Variables	Description	Min	Max	Mean	Median	SD
Street density	mile per mile square	0.247	1.320	0.755	0.743	0.147
Restaurant	number of restaurants	0.000	131.000	48.046	41.000	27.863
Bus stop	number of bus stops	1.000	146.000	59.213	56.000	31.915
Cafe	number of cafes	0.000	6.000	0.597	0.000	0.916
Quick service	number of quick-service stores	0.000	20.000	6.070	4.000	5.244
Convenience store	number of convenience stores	0.000	6.000	1.297	1.000	1.233
Liquor stores	number of liquor stores	0.000	3.000	0.265	0.000	0.520
Traffic node	number of traffic nodes	3.000	282.000	141.707	143.000	40.439
Sales amount	dollar	0.000	301440.000	94541.618	82328.000	66542.034
Home Built-Environment (one-mile straight line buffer)						
LUM	land use mix	0.413	0.875	0.626	0.630	0.095
Population density	people per mile square	0.004	0.231	0.128	0.128	0.042
Job density	job per mile square	0.055	4.909	1.436	1.387	0.584
Residential density	household per mile	84.000	8501.000	4225.888	4184.000	1670.190
Street density	mile per mile square	0.078	0.884	0.652	0.675	0.121
Restaurant	number of restaurants	1.000	136.000	47.096	40.000	27.046
Bus stop	number of bus stops	0.000	202.000	60.291	49.000	47.611
Cafe	number of cafes	0.000	11.000	1.083	1.000	1.574
Quick service	number of quick-service stores	0.000	38.000	7.811	6.000	7.175
Convenience store	number of convenience stores	0.000	15.000	3.634	3.000	2.861
Liquor stores	number of liquor stores	0.000	6.000	0.374	0.000	0.858
Traffic node	number of traffic nodes	6.000	1027.000	580.753	593.500	143.714
Sales amount	dollar	0.000	590734.000	114067.145	91477.500	103905.045

(continued on next page)

Variables	Description	Min	Max	Mean	Median	SD
Store Built-Environment (one-mile straight line buffer)						
LUM	land use mix	0.112	0.888	0.655	0.664	0.088
Population density	people per mile square	0.001	0.232	0.127	0.130	0.043
Job density	job per mile square	0.042	4.215	1.363	1.339	0.453
Residential density	household per mile	3.000	8694.000	4093.073	3994.000	1849.859
Street density	mile per mile square	0.121	0.924	0.670	0.693	0.100
Restaurant	number of restaurants	1.000	146.000	59.213	56.000	31.915
Bus stop	number of bus stops	0.000	205.000	65.393	51.000	46.191
Cafe	number of cafes	0.000	11.000	1.608	1.000	2.145
Quick service	number of quick-service stores	0.000	41.000	12.300	10.000	8.207
Convenience store	number of convenience stores	0.000	15.000	4.247	4.000	2.939
Liquor stores	number of liquor stores	0.000	6.000	0.660	0.000	1.212
Traffic node	number of traffic nodes	4.000	911.000	569.206	580.000	133.474
Sales amount	dollar	0.000	592964.000	182416.863	135288.000	129573.894
Home Built-Environment (half-mile network buffer)						
LUM	land use mix	0.000	1.046	0.522	0.514	0.132
Population density	people per mile square	0.008	3.526	0.458	0.426	0.240
Job density	job per mile square	0.779	92.290	4.527	3.828	3.637
Residential density	household per mile	0.000	2104.000	598.439	548.500	385.148
Street density	mile per mile square	0.746	6.516	1.631	1.561	0.404
Restaurant	number of restaurants	0.000	48.000	2.486	1.000	4.949
Bus stop	number of bus stops	0.000	50.000	8.380	6.000	9.278
Cafe	number of cafes	0.000	5.000	0.122	0.000	0.396

(continued on next page)

Variables	Description	Min	Max	Mean	Median	SD
Quick service	number of quick-service stores	0.000	16.000	0.804	0.000	1.794
Convenience store	number of convenience stores	0.000	5.000	0.460	0.000	0.775
Liquor stores	number of liquor stores	0.000	3.000	0.043	0.000	0.217
Traffic node	number of traffic nodes	2.000	234.000	71.022	71.000	30.819
Sales amount	dollar	0.000	212832.000	10681.249	1041.500	25256.031
Store Built-Environment half-mile network buffer)						
LUM	land use mix	0.000	0.986	0.592	0.608	0.111
Population density	people per mile square	0.000	3.140	0.378	0.383	0.176
Job density	job per mile square	0.724	38.778	3.950	3.670	2.048
Residential density	household per mile	0.000	1978.000	484.832	449.000	386.490
Street density	mile per mile square	0.814	18.914	1.730	1.533	0.934
Restaurant	number of restaurants	0.000	70.000	9.622	7.000	10.749
Bus stop	number of bus stops	0.000	42.000	13.091	11.000	9.819
Cafe	number of cafes	0.000	4.000	0.397	0.000	0.719
Quick service	number of quick-service stores	0.000	18.000	4.507	3.000	4.307
Convenience store	number of convenience stores	0.000	5.000	0.971	1.000	1.052
Liquor stores	number of liquor stores	0.000	3.000	0.161	0.000	0.404
Traffic node	number of traffic nodes	2.000	161.000	66.576	69.000	30.123
Sales amount	dollar	0.000	224407.000	49493.679	36602.000	48067.440
Home Built-Environment (one-mile network buffer)						
LUM	land use mix	0.000	1.046	0.531	0.529	0.109
Population density	people per mile square	0.007	3.526	0.330	0.306	0.220
Job density	job per mile square	0.707	92.290	3.520	3.056	3.118

(continued on next page)

Variables	Description	Min	Max	Mean	Median	SD
Residential density	household per mile	0.000	5640.000	2176.580	1984.500	1266.639
Street density	mile per mile square	0.746	6.516	1.425	1.393	0.306
Restaurant	number of restaurants	0.000	138.000	12.080	7.000	17.946
Bus stop	number of bus stops	0.000	156.000	33.832	23.000	32.078
Cafe	number of cafes	0.000	8.000	0.581	0.000	1.093
Quick service	number of quick-service stores	0.000	35.000	4.004	3.000	5.112
Convenience store	number of convenience stores	0.000	12.000	2.158	2.000	2.071
Liquor stores	number of liquor stores	0.000	5.000	0.190	0.000	0.579
Traffic node	number of traffic nodes	2.000	593.000	275.151	286.000	104.633
Sales amount	dollar	0.000	482521.000	52099.575	26562.500	72426.193
Store Built-Environment (one-mile network buffer)						
LUM	land use mix	0.000	0.768	0.581	0.604	0.100
Population density	people per mile square	0.000	3.140	0.288	0.290	0.128
Job density	job per mile square	0.305	38.778	3.188	2.950	1.758
Residential density	household per mile	0.000	5637.000	2187.745	2092.000	1336.061
Street density	mile per mile square	0.839	18.914	1.495	1.401	0.812
Restaurant	number of restaurants	0.000	134.000	22.721	18.000	23.453
Bus stop	number of bus stops	0.000	144.000	41.803	33.000	31.933
Cafe	number of cafes	0.000	9.000	1.044	1.000	1.467
Quick service	number of quick-service stores	0.000	34.000	8.997	7.000	6.910
Convenience store	number of convenience stores	0.000	11.000	2.774	3.000	2.002
Liquor stores	number of liquor stores	0.000	5.000	0.471	0.000	0.874
Traffic node	number of traffic nodes	2.000	577.000	296.999	310.000	112.386

(continued on next page)

Variables	Description	Min	Max	Mean	Median	SD
Sales amount	dollar	0.000	460123.000	111342.695	83522.500	93390.882
Home Built-Environment (census block level)						
LUM	land use mix	0.000	1.055	0.129	0.000	0.271
Population density	people per mile square	0.000	1.963	0.295	0.250	0.243
Job density	job per mile square	0.019	69.919	3.208	1.929	4.749
Residential density	household per mile	0.000	539.000	47.506	31.000	50.792
Street density	mile per mile square	0.000	18.573	3.138	2.810	1.808
Restaurant	number of restaurants	0.000	10.000	0.175	0.000	0.681
Bus stop	number of bus stops	0.000	9.000	0.507	0.000	1.075
Cafe	number of cafes	0.000	1.000	0.005	0.000	0.073
Quick service	number of quick-service stores	0.000	5.000	0.073	0.000	0.423
Convenience store	number of convenience stores	0.000	2.000	0.024	0.000	0.158
Liquor stores	number of liquor stores	0.000	1.000	0.003	0.000	0.056
Traffic node	number of traffic nodes	0.000	276.000	7.721	4.000	14.584
Sales amount	dollar	0.000	107924.000	924.129	0.000	6156.647
Store Built-Environment (census block level)						
LUM	land use mix	0.000	1.055	0.118	0.000	0.259
Population density	people per mile square	0.000	1.709	0.100	0.054	0.133
Job density	job per mile square	0.000	136.254	1.764	0.884	4.728
Residential density	household per mile	0.000	359.000	27.594	10.000	43.388
Street density	mile per mile square	0.000	21.652	2.358	1.917	1.728
Restaurant	number of restaurants	0.000	11.000	1.611	1.000	2.393
Bus stop	number of bus stops	0.000	9.000	1.481	1.000	1.363

(continued on next page)

Variables	Description	Min	Max	Mean	Median	SD
Cafe	number of cafes	0.000	1.000	0.100	0.000	0.301
Quick service	number of quick-service stores	0.000	9.000	1.214	1.000	1.819
Convenience store	number of convenience stores	0.000	2.000	0.141	0.000	0.368
Liquor stores	number of liquor stores	0.000	1.000	0.039	0.000	0.193
Traffic node	number of traffic nodes	0.000	276.000	7.964	4.000	15.142
Sales amount	dollar	0.000	141755.000	26090.893	9030.000	34043.144
Home Built-Environment (census block group level)						
LUM	land use mix	0.000	0.909	0.531	0.545	0.126
Population density	people per mile square	0.001	1.305	0.205	0.195	0.120
Job density	job per mile square	0.055	15.996	4.520	4.115	2.359
Residential density	household per mile	1.000	3921.000	614.085	477.000	563.513
Street density	mile per mile square	0.034	3.440	1.496	1.535	0.536
Restaurant	number of restaurants	0.000	74.000	2.145	1.000	4.499
Bus stop	number of bus stops	0.000	97.000	6.784	5.000	7.465
Cafe	number of cafes	0.000	5.000	0.097	0.000	0.362
Quick service	number of quick-service stores	0.000	17.000	0.910	0.000	2.037
Convenience store	number of convenience stores	0.000	5.000	0.400	0.000	0.709
Liquor stores	number of liquor stores	0.000	2.000	0.054	0.000	0.240
Traffic node	number of traffic nodes	3.000	923.000	110.184	67.500	140.448
Sales amount	dollar	0.000	685978.000	14337.246	2068.000	41274.693
Store Built-environment (census block group level)						
LUM	land use mix	0.092	0.911	0.598	0.608	0.104
Population density	people per mile square	0.001	1.305	0.171	0.162	0.104

(continued on next page)

Variables	Description	Min	Max	Mean	Median	SD
Job density	job per mile square	0.071	15.996	4.166	3.801	2.020
Residential density	household per mile	31.000	3921.000	460.074	389.000	322.319
Street density	mile per mile square	0.037	3.199	1.479	1.441	0.426
Restaurant	number of restaurants	0.000	74.000	8.736	5.000	12.789
Bus stop	number of bus stops	0.000	97.000	10.861	9.000	11.140
Cafe	number of cafes	0.000	5.000	0.430	0.000	0.876
Quick service	number of quick-service stores	0.000	17.000	3.986	2.000	4.342
Convenience store	number of convenience stores	0.000	5.000	0.985	1.000	1.216
Liquor stores	number of liquor stores	0.000	2.000	0.196	0.000	0.431
Traffic node	number of traffic nodes	3.000	923.000	101.178	74.000	100.172
Sales amount	dollar	0.000	685978.000	75961.728	45325.000	94616.433
Home Built-Environment (census tract level)						
LUM	land use mix	0.223	0.775	0.535	0.545	0.111
Population density	people per mile square	0.001	0.612	0.190	0.182	0.101
Job density	job per mile square	0.061	9.910	3.249	3.175	1.631
Residential density	household per mile	41.000	5320.000	1567.707	1457.000	861.042
Street density	mile per mile square	0.044	2.210	1.250	1.337	0.421
Restaurant	number of restaurants	0.000	74.000	5.957	4.000	6.760
Bus stop	number of bus stops	0.000	100.000	18.046	16.000	12.442
Cafe	number of cafes	0.000	5.000	0.292	0.000	0.570
Quick service	number of quick-service stores	0.000	17.000	2.554	1.000	3.372
Convenience store	number of convenience stores	0.000	6.000	1.131	1.000	1.216
Liquor stores	number of liquor stores	0.000	3.000	0.141	0.000	0.431

(continued on next page)

Variables	Description	Min	Max	Mean	Median	SD
Traffic node	number of traffic nodes	21.000	1041.000	263.737	208.000	201.178
Sales amount	dollar	0.000	685978.000	36825.632	13388.000	60750.948
Store Built-Environment (census tract level)						
LUM	land use mix	0.238	0.775	0.587	0.594	0.097
Population density	people per mile square	0.001	0.612	0.169	0.160	0.091
Job density	job per mile square	0.061	9.910	3.060	2.853	1.376
Residential density	household per mile	41.000	5320.000	1313.645	1194.000	713.598
Street density	mile per mile square	0.044	2.210	1.264	1.279	0.337
Restaurant	number of restaurants	0.000	74.000	13.866	9.000	13.711
Bus stop	number of bus stops	0.000	100.000	21.029	19.000	12.811
Cafe	number of cafes	0.000	5.000	0.668	0.000	0.979
Quick service	number of quick-service stores	0.000	17.000	5.876	4.000	4.988
Convenience store	number of convenience stores	0.000	6.000	1.810	1.000	1.663
Liquor stores	number of liquor stores	0.000	2.000	0.274	0.000	0.542
Traffic node	number of traffic nodes	21.000	1041.000	240.997	196.000	165.284
Sales amount	dollar	0.000	685978.000	104237.046	71556.500	104109.977