A BEHAVIORAL ANALYSIS OF SHOPPING TRIP CHAINING IN UNITED STATES

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Ting Sun

August 2015

ABSTRACT

This thesis studies travelers' tendency to chain a shopping trip into daily travel tour. A comprehensive analysis of personal shopping trip-chaining probability and its relation to sociodemographic characteristics, household status and land use information is conducted both nationally and area wide in United States. Using the 2009 National Household Travel Survey (NHTS) data, a binary logit model is chosen to analyze the probability of trip chaining, and a negative binomial regression model is used to model individuals daily shopping frequency. Results show that gender, household life cycle, family income, driver status, rural living environment and weekend have significant impact on people's chaining propensity, while influences of age, education, and worker status is insignificant. The same studying method and models are applied to all US census divisions as well. Area behavior, as expected, is consistent with national behavior.

BIOGRAPHICAL SKETCH

Ting Sun was born on November 1, 1990 in Beijing, China. She graduated from Beijing No.101 High School in 2009. She then attended Tsinghua University in Beijing for four years where she graduated with a bachelor's degree in Civil Engineering. After graduation, Ting started her master research in the Field of Transportation System Engineering at Cornell on 2013, and minored in Operations Research. This thesis is the summary of her two-year research work. Following graduation, Ting will start another Master of Science program, Quatitative Finance, in Georgia Institute of Technology.

I dedicated this thesis to my parents, my dearest friend and family: Weixuan and Daoji for their understanding, support and most of all, love. Without them I would not have accomplished as much.

ACKNOWLEDGEMENTS

I am deeply indebted to Professor Linda Nozick whose stimulation motivation and variable ideas help me to complete this work. I would thank Dr. Francis Vanek for his suggestions and advice on the presentation slides. I am also grateful to all my Special Committee members: Professor Oliver Gao, James Renegar for their care and help to me. Special thanks to Professor Peter Jackson, who agreed to be proxy of my minor advisor without hesitation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Travelers often take a series trips when they want to accomplish multiple tasks in a single tour. A series of chained trips form a trip chain, or we call it a tour. Researchers are interested in trip chaining behavior because it is generally considered time saving and fuel reducing. Trip chaining is also a common phenomenon of people's daily travel behavior, which transportation planners care a lot. Studying trip chaining helps transportation planners and policy makers to better understand complex travel behavior and forecast travel demand.

1.1 Background

As early as 1970s, the concept of trip chaining came up when researchers discovered the dependency among trips. The interrelationship of trips plays an important role in travel demand modeling; without it estimations are biased. On the other hand, peoples' time are more valuable nowadays, so travelers often consider more on how to optimize their time schedule. Combining several trips in a tour is an efficient and common way to save time and energy. Researchers found out that the frequency people trip chaining increased significantly compares that of the past. The popularity of private vehicles also significantly increased people's mobility, which makes trip chaining easier and encourages it.

Peoples' shopping travel patterns have been changing. Shopping trip, as one of the main travel purposes in the United States, is increasing in shares of all trips. Shopping trip takes about 20% of all trips in US, according to 2009

National Household Travel Survey [1]. Researchers found out that trip chaining happens a lot when related to shopping trip. More and more shopping stores are built next to each other to attract people, taking advantage of shopping trip chaining property. Understanding shopping trip chaining is crucial for planners and policy makers better model shopping travel demand.

What are the reasons that researchers and scholars are interested in trip chaining? First, researchers are often interested in whether vehicle trips can be done by other environmentally friendly modes of transportation. By looking at short tours and how much respondents of travel day they comprised, researchers can begin to study how much travels can be done by modes other than private vehicles. Second, the more people trip chaining, the more they will save on fuel use, which in turn can help improve air quality. By quantifying trip chaining, researchers can find out how much fuel can be saved by it, or potential of savings if trip chaining increases. Finally, trip chaining can help improve congestion. If a person needs to make several trips, chaining them together can help save time stuck in traffic.

There are tons of papers studying shopping trip and trip chaining separately, but few had linked them together. This thesis provides a link between shopping trips and trip chaining, and makes contribution to quantitatively study the shopping trip chaining behavior. A number of individual's sociodemographics and their household characteristics have been discussed in this thesis, and their influences on shopping trip chaining behavior have been quantitatively analyzed.

This thesis conducted a comprehensive national wide shopping trip chaining behavioral analysis using 2009 National Household Travel Survey (NHTS) data. In addition to analysis in the whole country, an area wide analysis in all divisions in the United States have also been done to confirm and support conclusions in the national study. While the analysis in national study gives researchers a thorough overview of shopping trip chaining behavior in the whole country, the area wide study that considers region difference is more practical in use.

1.2 Definition of Terms

Clear definitions of frequently used terms in this thesis, such as trip, dwell time, trip chain, tour and stop are given below and in the following subsection for convenience of further discussion.

A trip is a direct movement from an origin to a destination. The definition of dwell time is the amount of time that the traveler spent at current location. An anchor means either the start point (origin) or the end (destination) of a trip chain. A series of trips between two anchors is called a trip chain, or a tour. A stop is one of the intermediate destinations in a tour.

1.2.1 Definition of Trip Chaining

There is no formal defination of a trip chain, but I am using the one given by Federal Highway Administration (FHWA), who conducted the national survey and created the dataset used in this thesis.

A trip chain, also referred as a tour, refers to a series of trips that are chained together in time sequence between two anchored destinations. In 2009 NHTS, chains can begin at any of the three types of destinations: home, work or other. Travel to home or work after the beginning of a chain always terminates the trip chain. This means that home or work will never be in the middle of a chain. Other purpose will not automatically end the chain unless the person stays at the middle destination for longer than 30 minutes. According to this stopping-time requirement given by FHWA, to form a trip chain, the amount of time the traveler spend at each of the intervening stops should be less than 30 minutes. We will look at some examples of trip chains in Figure 1.1.

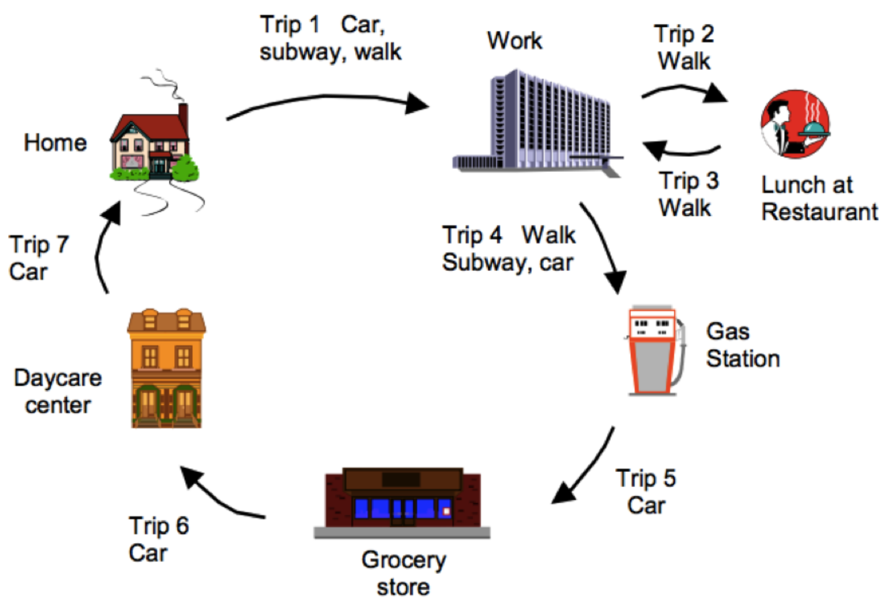

Figure 1.1: Daily travel trips of a hypothetical person in a weekday. [1]

Figure 1.1 shows a hypothetical person's travel during the day which contains 7 separate trips. Trip 1 is from home to work with no stops in between. The destination is work place, so trip 1 is not a tour. Then, he walks for a lunch at noon and go back to work later (trip 2 and 3). The length of the restaurant stay is important. The two trips consist of a tour if it is a fast food restaurant and he spends less than 30 minutes there. If he sit in the restaurant and have a meal for longer than 30 minutes, trip 2 and 3 will not be considered part of a tour. At the end of the day, trips 4 through 7 consist his way from work to home, when he stops in the middle for tasks such as filling gas at gas station, picking up food at grocery store and picking up a child at daycare center. If none of the intervening stops take over 30 minutes, then trip 4 to 7 will consist of a single tour.

As some literatures indicated, the longer the trip chain, the more complex the tour is. The tour made up by trip 4 through 7 is more complex than the one formed by trip 2 and 3.

1.3 Literature Review

Although the literatures in trip chaining have focused on a lot of distinct problems, they can be classified into three main aspects: trip dependency, tour complexity and trip chaining behavior.

In 1970s, researchers first discovered the dependency between trips that cause bias and inaccuracy in travel demand generation modeling. Early studies were focused on understanding urban impact and the interrelationship between trips, especially shopping trips (Hanson, 1980 [14]; Takahashi, 1986 [23]; Goulias and Kitamura, 1991[13]). Trip chaining were studies and added into travel forecast models to gain a better demand forecast (Adler and Ben-Akiva, 1979 [2]). Kitamura (1984) [18] pointed out that if dependency of trips is not accounted for, the coefficient of estimates were biased. Golob (1999) [12] developed a trip chaining generation model to forecast activity participation and travel time.

Tour complexity was also a key study topic regarding to trip chaining. Hensher (2000)[15], Cirillo and Axhausen (2002) [7] and Ye et al. (2006) [28] analyzed the relationship between trip chaining and mode of choice. They used number of stops of a tour to describe trip chain complexity, and found out that with the increment of tour complexity, the usage of public transportation significantly dropped. Complex tours are dominant by private vehicles. However, these works failed to explain why and the formation of complex tours, rather than focused on modeling. Noland and Thomas (2010) [22] discussed how population densities influence trip chaining behavior. They concluded that sparse population densities encouraged tour complexity than well-developed city areas. In addition, Ye et al. (2007) [28] studied the tours primary purpose and trip chain complexity. They showed that the complexity of trip chains related with characteristics of traveler. They also found out that older people would like to chain shopping trips.

The third main research topic of trip chaining is behavioral analysis. Characteristics of individuals' sociodemographics and land use were frequently studied by researchers (Yang et al., 2010 [27]; Walle and Steenberghen, 2006 [25]). Wallace (1999) [24] found that number of adults in household, count of employees and number of children have impact on household trip chaining behavior. Larger number discouraged household trip chaining. Ye et al. (2007)[28] found out the negative relation of household size and trip chaining, and the negative relation of family income and trip chaining. Clarke et al. (1981) [8] analyzed child effect on trip chaining. He discovered that young worker family with no child chains their trips in large proportion. Household with preschool children took more simple tours and have fewer complex working commute tours. Household with school age children, however, tended to have more complicated trip chains. Golob (1986) [11] and McGuckin et al. (2005) [21] both found out that life cycle is the most important factor regarding to trip chaining activities. McGuckin et al. also found out that expect for life cycle and gender, other sociodemographic variables had insignificant impact on trip chaining behavior. Wallace et al. (2000) [24] and Krizek (2003) [19] indicated that high density of service facilities discouraged complex tours but encouraged the total number of tours. Zhao et al. (2012) [29] focused on evolution of trip chaining patterns.

Some researches have focused their studies on a specific people segments. McGuckin and Murakami (1999) [20] studies women and their trip chaining behavior when there are children in the household. Noland and Thomas (2010) [22], Schmcker et al. (2010) [22], drew attention to old peoples' trip chaining behavior and found out that medical condition and limitation of mobility had impact on trip chaining. Wegmann (1998) [26] studied worker's trip chaining behavior.

Table 1.1 summarize some of the famous studies and the variables they focused on. We can see that life cycle impact have been confirmed by most studies. Although different research groups may have studied on the same variable, there are no contradiction in their conclusion regarding to the same variable.

Literature of shopping trips has been widely studied by researchers. Doti and Sharir (1981) [9] developed a theoretical constrained utility maximization model to study grocery shopping behavior including time in trip, shopping size and number of trips within a given period and Arentze et al. (1993) [3] added interdependency of chained stores in it. Kahn (1989) [16] studies shopping behavior empirically, while Bhat (1997) [4] have conducted several quantitative multinomial logit models to analyze travel model and departure time choice

Table 1.1: Summary of variables that have an impact on trip chaining

	Life cycle	Worker	Age	Income	Gender	HH Size	Land Use
Clarke, 1981	*						
Wallace, 1999	*	*					*
Golob, 1986	*		*	*			
Noland, 2010	*		*		*		*
McGuckin, 2005	*				*		
Ye, 2007				*		*	
Krizek, 2003							*

for urban shopping trips. Kim and Park (1998)[17] found out that shoppers are heterogeneous in terms of shopping trip regularity and frequency. 70% are random shoppers, while 30% have fixed regular intervals. Some recent group of researches also linked traditional shopping trips with online shopping behavior (Farag, 2007[10]; Cao, 2009 [5]; Cao et al., 2013 [6]; Zhou and Wang, 2014 [30]).

1.4 Problem of Interest

There are three main limitations of previous behavioral analysis works. First, most studies are descriptive analysis. They neither failed to quantify variable effect nor explained why. Second, almost every group was using a different dataset in an area. There was not a national wide study on trip chaining behavior due to the limitation of data. Finally, there is not a study linked shopping trip with trip chaining. To fill this gap, this thesis tries to answer these questions in the following chapters: what are the factors influencing shopping trip chaining behavior and how to quantify them? What are the difference in results of national study and area wide study? Are there any region behavior differences?

CHAPTER 2

DATA

This thesis use 2009 National Household Travel Survey (NHTS) data, conducted for Federal Highway Administration (FHWA) by the U.S. Bureau of the Census. The survey of NHTS family (NPTS before 2001) is conducted every five to seven years, with the first one conducted on 1969. It contains 6 data files: day trip file, trip chaining file, personal file, household file and vehicle file. The data of interest in this thesis is shopping trips in day trip file and trip chaining file. The key (in Database Management, each entry has a unique key) to access the same entry in those two files is made up by personal ID, household ID and travel day trip number. Referring to the key, information in two files can be linked together.

2009 NHTS contains daily trips of all family members in all households in a randomly assigned travel survey day, and it is designed to catch trend of individual as well as household travel patterns. It is the only national wide travel information source in the United States. 2009 NHTS provides information to assist transportation planners and policy makers who need comprehensive data on travel and transportation patterns in the United States. The 2009 NHTS updates information from 2001 NHTS and prior Nationwide Personal Transportation Surveys (NPTS) in 1995, 1990, 1983, 1977 and 1969.

2.1 Day Trip Data

The 2009 NHTS serves as a national inventory of daily travel, which contains 205,062 shopping trips, 19.5% of all 1,048,575 trips. Day trip file records all of the respondent's daily trips taken in a 24-hour period. Each household was assigned a "travel day", and the respondent were required to record all trips taken during that date. Households are also randomly assigned a travel day over the year, so that any seasonal effect see in travel, like holiday effect, or weekend travel is taken in to account. The time frame for the 2009 NHTS is a 13-month window from April 2008 to April 2009. Households in 2009 NHTS were selected in a stratified random sample, and the national sample includes households in all 50 states.

Information recorded on 2009 NHTS are trip purpose (work, recreation, etc.), transportation mode used of the trip (car, bus, bicycle, walk, etc.), trip length, travel time, time of day and day of week when trip took place, number of people traveled with respondent, traveler sociodemographic characteristics (gender, education level, working status, etc.), respondent's household information (household income, life cycle, urban or rural area, population, number of drivers or workers in the household, household race, etc.). The dataset is created for all trips, all purposes, all modes, and all areas of the country of all respondents in their assigned travel day.

2.2 Trip Chaining Data

Trip chain data file is generated from day trip file by FHWA for user's convenience. It contains one record for each day trip record. It is the same size with day trip file, over 200,000 shopping trips in total. Household ID, personal ID and travel day trip number serve as record key in both files. The trip-chain file records variables of tour flag (whether the trip entry belong to a tour), number of stops of the tour, dwell time at the stop, type of tour and final trip weight. There are only 3 kinds of anchors recorded, home, work and other, which generate 9 types of tours: home-to-home, home-to-work, work-to-other, etc.

2.3 Dataset of Interest

The dataset collected over 1,048,575 trips in United States, and 205,062 of them are recorded with shopping purpose. Only shopping trips are of this thesis interest. 205,062 shopping trips are selected from both trip chain file and day trip file, and the two files are joined with key (personal ID, household ID and travel day trip number). The created shopping file is used for national study. Besides national analysis, NHTS divided the Unite States into 9 divisions, see Figure 2.1, according to US Census definition. Division wide studies for all areas have also been conducted. Referring to household ID, personal ID and travel day trip number, information of the same trip in trip chaining file and day trip file can be combined together.

2009 NHTS's advantage over other datasets is great. It is the only national wide travel pattern resource from government; it is a relatively new survey (the

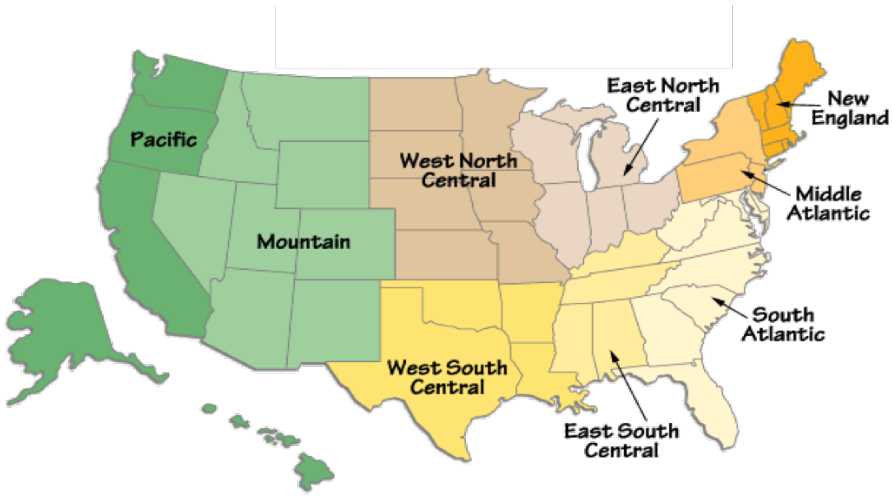

Figure 2.1: US Census division.

trip chain file is released on 2012); and it contains all the necessary information needed. 2009 NHTS is extremely powerful and thus preferred by researchers. I used this dataset to relate travel behavior with sociodemographics variables.

CHAPTER 3

METHOD

This thesis applied econometric method to model the relation of shopping trip chaining tendency and individual's sociodemographics, household and land use characteristics. Coefficients are estimated and used to quantify trip chaining behavior. In particular, a binary logit model is applied to study the tendency of shopping trip chaining, served as the main model of this thesis. A supplementary negative binomial model is used to model the respondent shopping frequency, with the purpose of comprehensively explaining the results. Model type is chosen by properties of their dependent variables. All individual's sociodemographics, household and land use characteristics variables are tried and final selected independent variables are chosen by not violating multicolinearity, i.e. each independent variable is not dependent with all other variables.

3.1 Binary Logit Model

This thesis applied a binary logit model to study shopping trip chaining tendency. Binary logit model, also known as binomial logistic regression model, is suitable for modeling binary outcomes. The model is used because of the way dependent variable is define: y describes whether the respondent chain the current shopping trip into a trip chain. y equals to 1 when the respondent chain the shopping trip into a tour, with probability p; y equals to 0 with probability $1 - p$ when not chained. As in equation (3.1)

$$y = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1 - p. \end{cases} \tag{3.1}$$

As equation (3.2) shows, p_i is the conditional probability of $y_i = 1$ given $\boldsymbol{x_i}$, where $\boldsymbol{x_i}$ is a $m \times 1$ vector indicating individual i's m characteristics, y_i is the indicator that whether this person chain his trip or not, and p_i is the probability of person i chaining the trip.

$$p_i = \Pr(y_i = 1|\boldsymbol{x_i}). \tag{3.2}$$

In logistic regressions, odds is defined as

$$\text{Odds} = \frac{p}{(1 - p)}, \tag{3.3}$$

and odds ratio of event 1 and event 2 is

$$\text{Odds ratio} = \frac{p_1/(1 - p_1)}{p_2/(1 - p_2)}. \tag{3.4}$$

The logistic regression assumes that the natural logarithm of odds is linearly related with independent variables:

$$\text{logit}(p) = \ln\frac{p}{1 - p} = \ln\frac{\Pr(y = 1|x)}{\Pr(y = 0|x)} = \boldsymbol{x'\beta}, \tag{3.5}$$

where $\boldsymbol{\beta}$ is the matrix of parameter of estimates, an $m \times n$ matrix for n observations.

Having estimated coefficients, it is very easy to compute odds and probability of success in a binary logit model. Transform equation 3.5, we get

$$\frac{p_i}{1 - p_i} = \exp(x_i'\beta), \tag{3.6}$$

and

$$p_i = \frac{e^{x_i'\beta}}{1 + e^{x_i'\beta}}. \tag{3.7}$$

Using equation 3.6 and 3.7, we can analyze single variable effect on odds. If full information of the individual is available, it is possible to forecast their probability of chaining a shopping trip.

Independent variables considered for binary logit model are age, education, gender, worker, driver, number of people with respondent in the trip, urban or rural living area, weekend, etc. Table 3.1 shows how variables are defined, their type (dummy, categorical or discrete), descriptions and whether the variable is selected.

Dummy variable, also called indicator variable, is a zero-one variable that describe a single effect. For example, the dummy variable for gender is defined as: 1 - women, 0 - men. Categorical variable, for example, education, is classified into 5 categories, it equals to 1 when the respondent education level is less than high school, 2 when he has high school degree, etc. Discrete variable such as age, takes all natural numbers from 0 to 100 (or larger).

Table 3.1: Selected variables for the binary logit model

Variable	Type	Description	Selection
Gender	Dummy	1-women, 0-men	✓
Education	Categorical	5 categories, 1-less than high school, 2- high school, 3-college, etc.	
Age	Discrete	Adult only, 18-65 years old, otherwise 0	
Urban	Dummy	1-urban, 0-rural	✓
Driver	Dummy	1-is driver, 0-not	✓
Weekend	Dummy	1-is weekend, 0-is not	✓
Family with Child	Dummy	1-R is adult and the family has at least one child, 0-otherwise	✓
Public Transportation	Dummy	1-use public transportation, 0-not	✓
Household Life Cycle	Categorical	10 types of life cycles, see Appendix	✓
Household Race	Categorical	8 categories, 1-white, 2-African American, 3-Asian, etc.	
Household Size	Discrete	Number of people in the family	
Family Income	Categorical	18 categories, 1-<$5000,18->$100,000.	✓
Household Vehicle Count	Discrete	Number of vehicles in the household	
# of People with R on Trip	Discrete	Number of people with R on the trip	✓
Area Indicator	Categorical	1-New England, 2-Middle Atlantic, etc., see Figure 2.1	✓

In table 3.1, education, age and household race are not selected finally because their coefficient of estimates are not significantly different from zero. Household vehicle count, household size is deleted because they have multicolinerity problems with existing variables. The multicolinerity is measured by **variance inflation factor** (VIF). VIF of factor β_i is calculated with the following formula:

$$\mathrm{VIF} = \frac{1}{1 - R_i^2}, \tag{3.8}$$

where R_i^2 is the coefficient of determination of the regression equation. VIF = 1 means there is absolutely no multicolinearity and VIF > 5 indicates serious multicolinearity problem.

3.2 Negative Binomial Model

A supplementary negative binomial model is used to model people's shopping frequency, because its results can provide better understanding of the trip chaining tendency, which will be illustrated in detail on the next chapter.

Shopping frequency, or number of shopping trips the respondent take during the survey day, served as the dependent variable y for this model. Shopping frequency is a count variable that statistically derived from the day trip dataset. There are several good models designed specially for count data, but I applied the negative binomial model for the following reasons.

Linear regression model is not proper here because the dependent variable is discrete. Linear regression best works with continuous dependent variables. Poisson regression model is often used, but it requires that the sample distribution of dependent variable should follow a Poisson distribution; mean and variance of sample should be equal. The distribution of individual's shopping frequency, however, is by no means Poisson distributed. Negative binomial regression model relax this constraint and allows that mean and variance is not equal.

Instead of assuming

$$y \sim \text{Poisson}(\lambda), \tag{3.9}$$

in the Poisson distribution, negative binomial model assumes that λ is also a random variable that follows Gamma distribution:

$$y \sim \text{Gamma}(\alpha, \beta). \tag{3.10}$$

That is y conditional on λ is Poisson distributed:

$$y|\lambda \sim \text{Poisson}(\lambda). \tag{3.11}$$

With simple transformation, we derived that the unconditional distribution of y is negative binomial distribution:

$$y \sim \text{Negbin}(\mu, \kappa). \tag{3.12}$$

Apply a log link to y and $\boldsymbol{x}$, the regression model is

$$log(\mu) = \boldsymbol{x}'\boldsymbol{\beta}, \tag{3.13}$$

where $\boldsymbol{\beta}$ and κ in equation 3.12 and 3.13 are the parameter of estimates.

The considered independent variables are the same with that of binary logit model, and they are selected using VIF. The finally selected ones are marked in Table 3.2.

Table 3.2: Selected variables for the negative binomial regression model

Variable	Type	Selection
Gender	Dummy	✓
Education	Categorical	✓
Age	Discrete	
Urban	Dummy	✓
Driver	Dummy	✓
Weekend	Dummy	✓
Family with child	Dummy	✓
Public Transportation Usage	Dummy	✓
Household Life Cycle	Categorical	✓
Household race	Categorical	
Household size	Discrete	
Family Income	Categorical	✓
Household vehicle count	Discrete	✓
Number of people with R on the trip	Discrete	✓
Area Indicator	Categorical	✓

CHAPTER 4

RESULTS

This chapter is organized in the following structure: first, the mathematical method that is used to estimate parameters is illustrated; then tells the coefficient of estimates table of both models; finally, interpretations and insights from the results are given.

Both models are estimated using Akaike Information Criterion (AIC). AIC is computed as

$$\mathrm{AIC} = 2k - 2\log(L), \tag{4.1}$$

where k is number of parameters, L is the maximum likelihood. AIC method not only maximize log-likelihood, but also draw a penalty when there are too many parameters. The prefered model is the one with a minimum AIC value.

4.1 Coefficient of Estimate

The coefficient of estimates of final binary logit model and negative binomial model is show in Table 4.1 and Table 4.2 respectively.

Table 4.1: Coefficient of estimates for binary logit model (trip chaining tendency)

Variable	Para. of Estimate	Std. Error	Significance	VIF	1/Odds: $(1-p)/p$
Constant	1.613	0.037	***		
Gender	-0.247	0.014	***	1.238	1.280
Weekend	-0.343	0.015	***	1.050	1.409
Urban	-0.115	0.016	***	1.001	1.212
Driver	0.174	0.024	***	1.234	1/1.190
Family with Children	-0.100	0.016	***	1.082	1.105
Public Transportation	-1.153	0.074	***	1.001	3.168
Household Life Cycle	-0.008	0.002	***	1.046	×
Family Income	0.007	0.001	***	1.183	×
Number of People with R on the Trip	-0.142	0.011	***	1.050	×
Area Indicator	-0.022	0.003	***	1.044	×

Level of significance code: 0.0001 '***'; 0.001 '**';0.01'*'.

× means not applicable. Odds of categorical variable is a $m \times m$ matrix, where m is number of categories. It is unable to show it in the table.

Table 4.2: Coefficient of estimates for negative binomial model (shopping frequency)

Variable	Parameter Estimate	Std. Error	Significance	VIF
Constant	-0.408	0.037	***	1.204
Gender	0.047	0.015	**	1.056
Education	0.033	0.004	***	1.042
Weekend	0.229	0.009	***	1.010
Urban	-0.093	0.010	***	1.109
Worker	0.219	0.007	***	1.302
Driver	0.023	0.002	***	1.463
Household Size	-0.074	0.002	***	1.842
Household Vehicle Count	0.015	0.004	***	1.105
Family with Children	-0.100	0.016	***	1.082
Online Shopping	0.004	0.000	***	1.001

Level of significance code: 0.0001 '***'; 0.001 '**';0.01'*'.

× means not applicable.

4.2 Quantitative Behavioral Analysis

4.2.1 Gender

Gender effect can be quantified by odds of trip chaining tendency, and gender odds is computed by coefficient of estimate of binary logit model. As equation 3.6 suggested, the odds = $\exp(-coef)$. The estimated coefficient for gender is -0.247 in binary logit model, and recall that $y = 1$ represents women and $y = 0$ for men, odds:

$$\frac{p_{gender}}{1 - p_{gender}} = \exp(\boldsymbol{x_{gender}}'\boldsymbol{\beta_{gender}}) = \exp(-0.247) = \frac{1}{1.28}, \tag{4.2}$$

This suggests that women chained less than men. The odds that men chaining a shopping trip is 1.28 times than that of women's. Other other hand, the negative binomial model gives positive estimates of gender, indicating that women actually travel more shopping trips than men! This can be explained by the statistics results that women on average spend more time on shopping than men. The average women shopping time is 32 minutes, while it is 29 minutes for men. Figure 4.1 gives the distribution of peoples' shopping time in US. Longer shopping time will terminate a trip chain. If shopping is the first purpose of a travel and he stays over 30 minutes, this shopping trip will not be counted as part of tour; but this reducing effect loses its power when the shopping purpose is the second, third or later trips, the shopping trip will still be counted as part of tour.

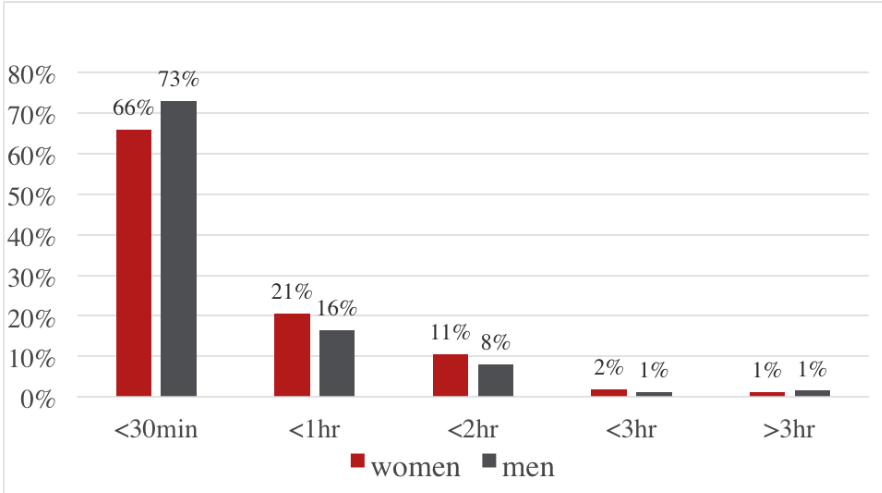

Figure 4.1: Self reported shopping time for women and men in US

4.2.2 Weekend

Weekend variable exhibits the same principle with gender. The estimated coefficient for weekend is -0.343 in binary logit model, and recall that $y = 1$ means the trip is taken on weekend and 0 on weekday, so odds:

$$\frac{p_{weekend}}{1 - p_{weekend}} = \exp(\boldsymbol{x_{weekend}}'\boldsymbol{\beta_{weekend}}) = \exp(-0.343) = \frac{1}{1.409}, \tag{4.3}$$

This implies that shopping trip chaining happens more on weekdays. What about the shopping time and frequency? Table 4.3 provides statistics directly that are derived from 2009 NHTS. It shows that although the traveling distance for shopping has no difference on weekday or weekend, the shopping time is longer for on weekend. Longer shopping time will terminate a trip chain some times.

Table 4.3: Average shopping time and distance on weekday and weekend

	Weekend	Weekday
Avg shopping distance (miles)	5.26	5.26
Avg shopping time (minutes)	34.82	29.08

4.2.3 Child

Child also has a significant impact on adults' traveling behavior. The indicator for family with child is carefully defines to 1 only when the respondent is an adult of age 18 to 65, and there is at least one child (age 0 to 17) in the household. When the respondent is a child, elder person or adults with no children, this variable is zero. So this variable only quantifies child's influence on adult's shopping chaining behavior. The estimated coefficient for child is −0.1 in trip chaining binary logit model, so odds:

$$\frac{p_{child}}{1 - p_{child}} = \exp(\boldsymbol{x_{child}}'\boldsymbol{\beta_{child}}) = \exp(-0.1) = \frac{1}{1.105}, \tag{4.4}$$

This means that existance of child in household significantly reduce the chance of chaining a shopping trip. Notice that two alternative variables have been added to test whether there is difference of child influence on women or men. But the results show that there is not.

4.2.4 Urban Effect

This thesis uses a simple urban/rural indicator variable to dispict land use effect instead of population density or other land use variables. Urban effect can be quantified by odds of trip chaining. As equation 4.5 suggested, the odds of urban area can be calculated by $\exp(-coef)$. The estimated coefficient for urban is -0.115, so odds:

$$\frac{p_{urban}}{1-p_{urban}} = \exp(\boldsymbol{x_{urban}}'\boldsymbol{\beta_{urban}}) = \exp(-0.115) = \frac{1}{1.212}, \tag{4.5}$$

It shows that the odds of shopping trip chaining in rural area is 1.212 times than that in urban areas. This result is actually quite reasonable if explained in this way: there is relatively limit living facilities in rural areas, so rural people's cost of traveling is higher than that of urban people. To save cost, trip chaining is a good way especially for shopping trips. As expected, empirical results show that rural people travel longer distance to shopping, as Table suggested. The difference in average shopping time in urban or rural area is not significant though.

Table 4.4: Average shopping time and distance in urban, rural area in US

	Urban	Rural
Avg shopping distance (miles)	3.58	5.26
Avg shopping time (minutes)	31.3	30.8

4.2.5 Driver

Driver status, in common sense, indicates high mobility of the person. The estimated coefficient for driver is 0.147 in trip chaining binary logit model, so odds:

$$\frac{p_{driver}}{1-p_{driver}} = \exp(\boldsymbol{x_{driver}}'\boldsymbol{\beta_{driver}}) = \exp(0.147) = 1.19, \tag{4.6}$$

which implies a positive effect of driver status in trip chaining.

4.2.6 Public Transportation

It is surprising to find out that the use of public transportation modes when chaining a shopping trip drops so fast rather than non-public modes. This coincide with other people's research that car is still the dominant mode in shopping and other trips. Public transportation effect can be quantified by odds of trip chaining. As equation 3.6 suggested, the odds of public transports can be calculated by $\exp(-coef)$. The estimated coefficient for gender is -1.153. Odds:

$$\frac{p_{pubtrans}}{1 - p_{pubtrans}} = \exp(\boldsymbol{x_{pubtrans}}'\boldsymbol{\beta_{pubtrans}}) = \exp(-1.153) = \frac{1}{3.168}, \tag{4.7}$$

which means the odds that people using private vehicles are 3.168 times than that of public transportations. It is reasonable because shopping usually associate with carrying stuff, which will make public transportation very inconvenient for later trips in the tour.

4.2.7 Other Effects

Other variables like area indicator are also significant in binary logit model, but the binary logit model fails to help us quantify their effect. The odds matrix for a 10-catetory variable, for example, is a 10×10 matrix. But we know that area difference does exist and there is a minor, subtle trend that East people generally chained more than people in the West (the areas are numbered from East to West, and we got a negative coefficient, -0.022). Further research regarding area difference may provide more reasonable insights to this problem, but that is beyond the scope of this problem.

CHAPTER 5

CONCLUSION

This thesis studies travelers' tendency to chain a shopping trip into daily travel tour. A comprehensive analysis of personal shopping trip-chaining probability and its relation to sociodemographic characteristics, household status and land use information is conducted both nationally and area wide in United States. Using the 2009 National Household Travel Survey (NHTS) data, a binary logit model is chosen to analyze the probability of trip chaining, and a negative binomial regression model is used to model individuals daily shopping frequency. Results show that gender, household life cycle, family income, driver status, rural living environment and weekend have significant impact on people's chaining propensity, while influences of age, education, and worker status is insignificant. The same studying method and models are applied to all US census divisions as well. Area behavior, as expected, is consistent with national behavior.

By quantifying trip chaining tendency, we end up with the following conclusions. Women are less likely to chain shopping trips, but they on average take more shopping trips than men and spend more time shopping. Weekend shopping take longer time than that on weekdays, and there is less shopping trip chaining. Adults with children are less likely to chain shopping trips or make complex tours. People in rural area spend more time on their way to stores because of the low density of living facilities, so they chain shopping trips more often than people in metropolitan area. Most shopping trips are done by private vehicles because of their convenience and high mobility. Driver status also has a positive effect on chaining propensity.

The limitations of this thesis remain. The definition of trip chain by FHWA requires that dwell time at intervening stop should be less than 30 minutes. In practice, however, people may think they are forming a trip chain when they spend one hour shopping, because they have other desitinations to visit after shopping. Due to this definition, we see some suprising results that women chained less shopping trips than men. With a different definition, we might have very contradictory results. The boundary of dwell time should be discussed in future research. While this thesis gives insights to transportation planners and policy makers of how likely people will chain a shopping trip, there is still a long way to apply the results directly to practical use.

BIBLIOGRAPHY

[1] U.s. department of transportation, federal highway administration, 2009 national household travel survey. url: http://nhts.ornl.gov.

[2] Thomas Adler and Moshe Ben-Akiva. A theoretical and empirical model of trip chaining behavior. *Transportation Research Part B: Methodological*, 13(3):243–257, 1979.

[3] Theo Arentze, Aloys Borgers, and Harry Timmermans. A model of multi-purpose shopping trip behavior. *Papers in Regional Science*, 72(3):239–256, 1993.

[4] Chandra R Bhat. Analysis of travel mode and departure time choice for urban shopping trips. *Transportation Research Part B: Methodological*, 32(6):361–371, 1998.

[5] Xinyu Cao. E-shopping, spatial attributes, and personal travel: a review of empirical studies. *Transportation Research Record: Journal of the Transportation Research Board*, (2135):160–169, 2009.

[6] Xinyu Jason Cao, Zhiyi Xu, and Frank Douma. The interactions between e-shopping and traditional in-store shopping: an application of structural equations model. *Transportation*, 39(5):957–974, 2012.

[7] Cinzia Cirillo, Kay W Axhausen, Kay W Axhausen, and Kay W Axhausen. *Mode choice of complex tours: A panel analysis*. ETH, Eidgenössische Technische Hochschule Zürich, Institut für Verkehrsplanung, Transporttechnik, Strassen-und Eisenbahnbau IVT, 2002.

[8] MI Clarke, MC Dix, PM Jones, and IG Heggie. Some recent developments in activity-travel analysis and modeling. *Transportation Research Record*, (794), 1981.

[9] James L Doti and Shmuel Sharir. Households'grocery shopping behavior in the short-run: Theory and evidence. *Economic Inquiry*, 19(2):196–208, 1981.

[10] Sendy Farag, Tim Schwanen, Martin Dijst, and Jan Faber. Shopping online and/or in-store? a structural equation model of the relationships between e-shopping and in-store shopping. *Transportation Research Part A: Policy and Practice*, 41(2):125–141, 2007.

[11] Thomas F Golob. A nonlinear canonical correlation analysis of weekly trip chaining behaviour. *Transportation Research Part A: General*, 20(5):385–399, 1986.

[12] Thomas F Golob. A simultaneous model of household activity participation and trip chain generation. *Transportation Research Part B: Methodological*, 34(5):355–376, 2000.

[13] Konstadinos G Goulias and Ryuichi Kitamura. Recursive model system for trip generation and trip chaining. 1991.

[14] Susan Hanson. The importance of the multi-purpose journey to work in urban travel behavior. *Transportation*, 9(3):229–248, 1980.

[15] David A Hensher and April J Reyes. Trip chaining as a barrier to the propensity to use public transport. *Transportation*, 27(4):341–361, 2000.

[16] Barbara E Kahn and David C Schmittlein. Shopping trip behavior: An empirical investigation. *Marketing Letters*, 1(1):55–69, 1989.

[17] Byung-Do Kim and Kyungdo Park. Studying patterns of consumer's grocery shopping trip. *Journal of retailing*, 73(4):501–517, 1998.

[18] Ryuichi Kitamura, Lidia P Kostnyiuk, and Michael J Uyeno. *Basic properties of urban time-space paths: empirical tests*. Number 794. 1981.

[19] Kevin J Krizek. Neighborhood services, trip purpose, and tour-based travel. *Transportation*, 30(4):387–410, 2003.

[20] Nancy McGuckin and Elaine Murakami. *Examining Trip-Chaining Behavior: A Comparison of Travel by Men and Women*. 1998.

[21] Nancy McGuckin, Johanna Zmud, and Yukiko Nakamoto. Trip-chaining trends in the united states: understanding travel behavior for policy making. *Transportation Research Record: Journal of the Transportation Research Board*, (1917):199–204, 2005.

[22] Jan-Dirk Schmöcker, Fengming Su, and Robert B Noland. An analysis of trip chaining among older london residents. *Transportation*, 37(1):105–123, 2010.

[23] S Takahashi. Effects of multiple stops on the distance travelled to a grocery store. *Geogr. Rev. Jpn*, 59(2):119–127, 1986.

[24] Brett Wallace, Jennifer Barnes, and G Rutherford. Evaluating the effects of traveler and trip characteristics on trip chaining, with implications for transportation demand management strategies. *Transportation Research Record: Journal of the Transportation Research Board*, (1718):97–106, 2000.

[25] Stefaan Vande Walle and Therese Steenberghen. Space and time related determinants of public transport use in trip chains. *Transportation Research Part A: Policy and Practice*, 40(2):151–162, 2006.

[26] Frederick J Wegmann and Tae Youn Jang. Trip linkage patterns for workers. *Journal of Transportation Engineering*, 124(3):264–270, 1998.

[27] Min Yang, Wei Wang, Gang Ren, Rui Fan, Bing Qi, and Xuewu Chen. Structural equation model to analyze sociodemographics, activity participation, and trip chaining between household heads: survey of shangyu, china. *Transportation Research Record: Journal of the Transportation Research Board*, (2157):38–45, 2010.

[28] Xin Ye, Ram M Pendyala, and Giovanni Gottardi. An exploration of the relationship between mode choice and complexity of trip chaining patterns. *Transportation Research Part B: Methodological*, 41(1):96–113, 2007.

[29] Zhan Zhao, Glen Chua, and Jinhua Zhao. Evolution of trip chaining patterns in london from 1991 to 2010. In *Innovations in Travel Modelling Conference, Tampa*, 2012.

[30] Yiwei Zhou and Xiaokun Cara Wang. Explore the relationship between online shopping and shopping trips: An analysis with the 2009 nhts data. *Transportation Research Part A: Policy and Practice*, 70:1–9, 2014.