A Discrete-Continuous Modeling Framework for Long-Distance, Leisure Travel Demand Analysis

by

Caleb Van Nostrand

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Civil Engineering
Department of Civil and Environmental Engineering
College of Engineering
University of South Florida

Major Professor: Abdul Pinjari, Ph.D.
Yu Zhang, Ph.D.
Steven Polzin, Ph.D.
John Lu, Ph.D.

Date of Approval:
March 23, 2011

Keywords: Long Distance Travel, Vacation Travel Demand, National Travel Model, Kuhn-Tucker Demand Model Systems, Destination Choice

Dedication

This thesis is dedicated to my parents, Al and Jeannine, for their constant support and encouragement in everything I do, and for always being there. It is from them that I learned what hard work is, and I am grateful to attribute my success thus far to them.

I would also like to dedicate this thesis to Danielle. She has continually supported me with patience and understanding despite all the late work nights and busy weekends, and I look forward to opportunities to return the favor in the future.

Acknowledgements

I would first like to thank my advisor, Dr. Abdul Pinjari, for his guidance and support in completing this thesis. He has consistently been available and supportive, and has provided invaluable insight throughout the process. I would also like to thank Dr. Yu Zhang, Dr. Steve Polzin, and Dr. John Lu for serving on my Master's thesis committee. They have all supported me in and out of classes during my tenure at USF. Thanks to the NBRTI team at CUTR, especially my supervisor Brian Pessaro, for their help and understanding during the thesis. Working at CUTR for the past few years has been pleasure. Finally, thanks to Vijay Sivaraman for his assistance throughout the project.

Table of Contents

List of Tables

List of Figures

ABSTRACT

This study contributes to the literature on national long-distance travel demand modeling by providing an analysis of households' annual destination choices and time allocation patterns for long-distance leisure travel purposes. An annual vacation destination choice and time allocation model is formulated to simultaneously predict the different destinations that a household visits and the time it spends on each of these visited destinations, in a year. The model takes the form of a Multiple Discrete-Continuous Extreme Value (MDCEV) structure (Bhat, 2005; Bhat, 2008). The model assumes that households allocate their annual vacation time to visit one or more destinations in a year to maximize the utility derived from their choices. The model framework accommodates variety-seeking in households' vacation destination choices in that households can potentially visit a variety of destinations rather than spending all of their annual vacation time for visiting a single destination. At the same time, the model accommodates corner solutions to recognize that households may not necessarily visit all available destinations. An annual vacation time budget is also considered to recognize that households may operate under time budget constraints. Further, the paper proposes a variant of the MDCEV model that avoids the prediction of unrealistically small amounts of time allocation to the chosen alternatives. To do so, the continuously non-linear utility functional form in the MDCEV framework is replaced with a combination of a linear and non-linear form.

The empirical data for this analysis comes from the 1995 American Travel Survey Data, with the U.S. divided into 210 alternative destinations. The empirical analysis provides important insights into the determinants of households' leisure destination choice and time allocation patterns.

An appealing feature of the proposed model is its applicability in a national, long-distance leisure travel demand model system. The annual destination choices and time allocations predicted by this model can be used for subsequent analysis of the number of trips made (in a year) to each destination and the travel choices for each trip. The outputs from such a national travel modeling framework can be used to obtain national-level Origin-Destination demand tables for long-distance leisure travel.

Chapter 1: Introduction

1.1 Background

In several countries, a significant portion of the travel comes from long distance travel, especially for leisure purposes. For example, in the United States, in the year 1977, Americans made approximately 521 million long distance person trips[1], totaling approximately 382 billion miles traveled (BTS, 1998). Within the next two decades, per the data in year 1995, the long distance travel more than doubled to about 1 billion person trips and 827 billion miles (BTS, 1998). While this increase may be attributed to an increase in travel for all purposes (business, social, and leisure, etc.), leisure travel is of particular importance due to several reasons. First, leisure travel constitutes a significant share of long distance travel (27% of all long distance trips made by US households in 1995 were for leisure; see BTS, 1997), as well as a significant share of the increase in long-distance travel (long-distance travel for leisure increased by 122% between 1997 and 1995; see BTS 1998, pp. 149). It also appears that the recent economic slowdown did not have a substantial impact on the vacation travel intentions of Americans. For instance, despite perceiving an increase in the vacation price, 84% of the respondents to a poll conducted by Priceline.com indicated that they still planned to travel (Hotel News Resource, 2007). Perhaps leisure travel is such an integral part of Americans' lifestyle (LaMondia and Bhat, 2008) that it is difficult to part with even in poor economic climates. Second, as the demographic makeup of several countries changes toward an increasingly ageing population, the amount of long-distance leisure travel is

[1] A long-distance trip is defined as roundtrip travel of at least 100 miles from home (BTS, 1998).

likely to continue to increase. Traveling and "exploring the world" appears to be an ambition that people pursue in their retirement years with substantial amounts of time and wealth at their discretion (Focalyst, 2007). On the same lines, several studies report that the baby boomers (those born between 1946 and 1964) allocate significant amounts of time and money to vacation travel (Mallet and McGuckin, 2000; Davies 2005). As the baby boomers have started to enter their late sixties, growth in vacation travel is likely to accelerate over the next several years. Third, leisure travel has a significant impact on the economy as it is highly consumption-oriented. For instance, a recent consumer expenditure report estimates that in the year 2008, U.S. households spent, on average, $1,415 per annum on activities such as dining, lodging, shopping, entertainment and recreation *while on* vacation and pleasure trips (BLS, 2010). It is not surprising that the economy of several destinations thrives on the tourism/leisure travel industry.

Due to the above-discussed and various other reasons, long-distance leisure travel behavior is one of the most studied topics in the tourism literature and is steadily gaining importance in the transportation literature. Several dimensions of leisure travel behavior have been studied to date, including whether to travel or not (Morley, 1992; Seddighi and Theocharous, 2002; Nicolau and Mas, 2005), travel purpose (LaMondia et al., 2008), length of stay and time/money budget allocation (Morley, 1992; Thornton et al., 1997; Money and Crotts, 2003; Nicolau and Mas, 2005), frequency of travel (Kubas et al., 2005), destination of travel (Train, 1998; Phaneuf and Smith, 2005) and mode of travel (LaMondia et al., 2009). Notable among these dimensions is the destination choice. From a tourism standpoint, a better understanding of where people travel for their vacation can aid in taking measures to enhance the attractiveness of the destinations and increase the tourism demand and revenue. Further, understanding the destination preferences of different types of travelers can help in devising targeted promotional campaigns to specific traveler segments. From a transportation planning perspective, understanding the vacation travel flow patterns helps in assessing national and local infrastructure needs and implementing appropriate transportation control policies.

This thesis contributes to the literature on long-distance leisure travel demand analysis by an analysis of households' long-distance, vacation travel destination choices in a year. Specifically a multiple discrete-continuous extreme value (MDCEV) model is used to analyze the different destinations that a household visits in a year and the time allocated to each of the visited destinations. The remainder of this section reviews the literature on long-distance leisure destination choice analysis and positions the current work vis-à-vis existing literature.

1.2 Literature on long distance leisure destination choice analysis

Leisure destination choice has been extensively studied in the tourism/leisure travel literature (Moutinho 1987; Eugenio-martin, 2003). A popular approach to analyze destination choices is the discrete choice analysis method using multinomial logit or nested logit models (Seddighi and Theocharous, 2002; Eymann and Ronning, 1997; Hong et al., 2006; Simma et al., 2001; and LaMondia et al., 2009). A variety of other methods have also been used to analyze various aspects related to destination choice. Examples include: (a) descriptive statistics (Bansal and Eislet, 2004; Crompton, 1979; Um et al., 1990) and regression analysis (Rugg, 1973; Molina and Esteban, 2006) , (b) factor analysis, determinant analysis and cluster analysis of destination image formation (Jiang et al., 2000; Castro et al., 2007), (c) structural equations modeling of beliefs, attitudes, and norms and past behavior on the intent to choose a destination (Lam and Hsu, 2006; Greenridge, 2001), (d) open ended surveys, cognitive mapping and qualitative analysis of the processes leading to destination choices (Woodside and

MacDonald, 1994; Woodside and Lyonski, 1989). Some of these studies[2] focus on analyzing the outbound tourism demand from one origin (usually a country) to multiple destinations, while others[3] analyze the inbound tourism demand from multiple origins to a single destination, such as a city or country. It appears that very few leisure studies analyze destination choices between multiple origins and multiple destinations. Specifically, LaMondia et al. (2009) analyzes vacation travel between several European Union countries, while Simma et al. (2001) analyzes leisure travel between the municipalities of Switzerland.

In the transportation planning/modeling literature, though several studies focus on short-distance leisure travel behavior within metropolitan areas (Yamamoto and Kitamura, 1999; Bhat and Gossen, 2004; Schlich et al., 2004; Lanzendorf, 2002), very little exists explicitly on long-distance leisure travel. Although long-distance travel analysis is a regular exercise in the form of statewide travel models[4] in the U.S. and intercity travel demand models[5], leisure travel is dealt with in very limited ways. For example, in statewide models, inter-state trips[6] are categorized as external, through, or visitor trips and the trip flows are estimated using aggregate, growth factor or gravity-based methods. Several national-level travel demand models also exist, predominantly

[2] Eymann and Ronning (1997), Gonzalez and Moral (1995), DeCrop and Snelders (2004), Lise and Tol (2001), Haliciolgu (2008)

[3] Greenridge (2001), Castro et al. (2005), Garin-Munoz, 2000; Chan et al. (2005)

[4] Horowitz (2006), Horowitz (2008), Cambridge Systematics (2007), Outwater et al. (2010)

[5] Thakuriah (2006), Koppelman and Sethi (2005), Bhat (1995), Baik et al. (2007), Yao and Morikawa (2005)

[6] A significant portion of long-distance leisure trips tend to be inter-state trips.

in the European context[7] and some for the US (Moeckel and Donnelly, 2010) and other nations. (see Zhang et al., 2010; Lundgvist and Mattsson, 2001 for extensive reviews). However, most models use aggregate trip distribution methods (couched within the traditional four-step modeling system) and/or do not pay explicit attention to vacation travel. This is not to say that disaggregate methods are not used or vacation travel is not paid any attention. Some statewide models in the U.S. (e.g., Outwater et al., 2010) and several European national models (e.g., Hackney, 2004) use disaggregate discrete choice MNL or nested logit models to analyze destination choices. A few studies analyze the destination choices with an explicit focus on vacation trips (LaMondia et al., 2010, Simma et al., 2002; Louviere and Timmermans, 1990). Furthermore, some models are built based on more behaviorally oriented activity-based and tour-based approaches (e.g., the Danish national model PETRA and the Dutch national model; Fosgerau, 2001) and agent-based methods (Parker and Epstein, 2008).

Despite all the advances, a drawback of most previous studies in both the travel demand literature and in the tourism literature is that their analysis is limited to smaller time frames such as a day (e.g., Cambridge Systematics, 2007; the Danish national model), a few weeks (e.g., the British national model) or months. Some studies (e.g., the Swiss national model) use a single trip, typically the most recent trip, as the unit of analysis, which restricts the ability to understand how the decisions pertaining to that trip are related to other vacation trips over longer time frames. Most data collection efforts also appear to collect travel information for smaller time frames other than a few exceptions such as the 1995 US American Travel Survey (ATS) and the DATELINE

[7] These include the national model systems for Denmark (PETRA, Fosgerau, 2001), Sweden (SAMPERS; Beser and Algers, 2001), Holland (LMS, HCG 1990), Germany (VALIDATE; Vortsih and Wabmuth, 2007), UK, Switzerland

survey[8] that collected respondent's travel information for one year. However, as indicated in Eugenio-martin's (2003) theoretical framework for tourism demand analysis and in Morley (1995), longer time frames such as a year may be more appropriate for vacation travel analysis (also see Little, 1979).

Existing studies with longer time frames such as a year use one of the two approaches: (1) Aggregate (e.g., gravity-based) methods for estimating annual vacation travel flows, (2) Employ disaggregate methods, but first predict the frequency of vacation trips for a given time frame and then perform a piecemeal analysis of the destination choices (and other decisions) for each trip. Studies belonging to the second category include van Middlekoop et al's (2004) microsimulation system for annual leisure activity/travel patterns and the long-distance holiday travel module in the recent version of the TRANS-TOOLS model for travel demand prediction in and between the European Union countries (see Rich et al., 2009). LaMondia et al.'s (2008) annual vacation time-use model is the only exception found that attempts a comprehensive analysis of the annual vacation time-use patterns by different vacation purposes. They do not, however, delve into destination choices.

1.3 Objectives of this thesis

In this paper, we propose an annual vacation destination choice and time allocation model to simultaneously analyze the different destinations that a household visits, and the time it spends on each of these visited destinations, in a year. Specifically, the recently emerging multiple discrete-continuous extreme value (MDCEV) model (Bhat, 2005; Bhat, 2008) is employed to analyze the factors influencing households' annual

[8] DATELINE Survey collects only holiday travel data for one year. This data is used estimate the travel models in the second version of the TRANS-TOOLS model for travel demand prediction in and between the European Union countries (see Rich et al., 2009). In this model, the total frequency of yearly long-distance holiday trips is first generated. These trips are then distributed to different destinations using a joint destination and mode choice model.

vacation destination choices and time allocation patterns. The model assumes that households allocate the annual vacation time available at their disposal to one or more destinations in a year in such a way as to maximize the utility derived from their choices. As described in LaMondia et al. (2008), the utility maximization framework is consistent with Iso-Ahola's (1983) *optimal arousal* concept of vacation behavior that people "suffer psychologically and physiologically from understimulating and overstimulating environments" and seek an "optimally arousing experience." The model framework accommodates variety-seeking in households' vacation choices in that households can potentially visit a variety of destinations rather than spending all of their annual vacation time for visiting a single destination. Households may seek variety in destination choices due to several reasons. First, different members of a household may have different preferences, leading to a variety in destinations choices. For example, children might prefer to spend a week at the Disney land while elderly might prefer a calm and warm winter resort. Second, households might visit multiple destinations due to satiation effects of increasing time allocation to a destination (i.e., they experience boredom and start seeking variety). Such satiation effects in vacation travel behavior have been noted in previous studies both in the context of visiting multiple destinations within a single vacation trip (Lue et al., 1993) as well as budgeting annual leisure time expenditures for different purposes (LaMondia et al., 2008). Third, people might take vacations for pursuing multiple types of activities (adventure, sightseeing, etc.) and/or during multiple seasons of the year but no single destination may be ideal for all purposes and/or during all time periods (hence a variety of destination choices over a year). The MDCEV model incorporates variety in destination choices by employing a non-linear utility framework that allows diminishing marginal utilities of increasing time allocation to a destination. At the same time, the model recognizes that households may not necessarily visit all available destinations, by incorporating corner solutions that allow zero time allocations to certain destinations. An annual vacation time budget is also considered to recognize that households may operate under time budget constraints.

The proposed model is couched within a larger vacation travel modeling framework as depicted in Figure 1. First, households are assumed to allocate annual time and money budgets for leisure travel. Next they are assumed to allocate the time and money budgets to visit one or more destinations. Subsequently, for each destination they choose to visit, they decide the number of trips to make to that destination, and travel choices for each trip, including mode choice, time (i.e., season) of the year, and length of stay. The analyst can apply this framework to all households in the nation and obtain a national-level Origin-Destination demand table for vacation travel. Of course, other decision elements, such as the travel party composition for each vacation trip, could be included in the framework. Further, the framework could be refined to include another step (between steps 1 and 2) where households allocate the annual vacation time to different purposes (recreation, sightseeing, etc.) and then decide the destinations to visit depending on the purposes they wish pursue. Alternatively, a slightly different framework that assumes an alternative hierarchy of decisions could be used (as shown in Figure 2). Specifically, in the second step the analyst can model the households' allocation of annual vacation time/money budgets into different purposes and different seasons (or times of the year). Subsequently, (s)he could model the destination choices and other travel decisions (e.g., mode choice) for each purpose and time of the year.

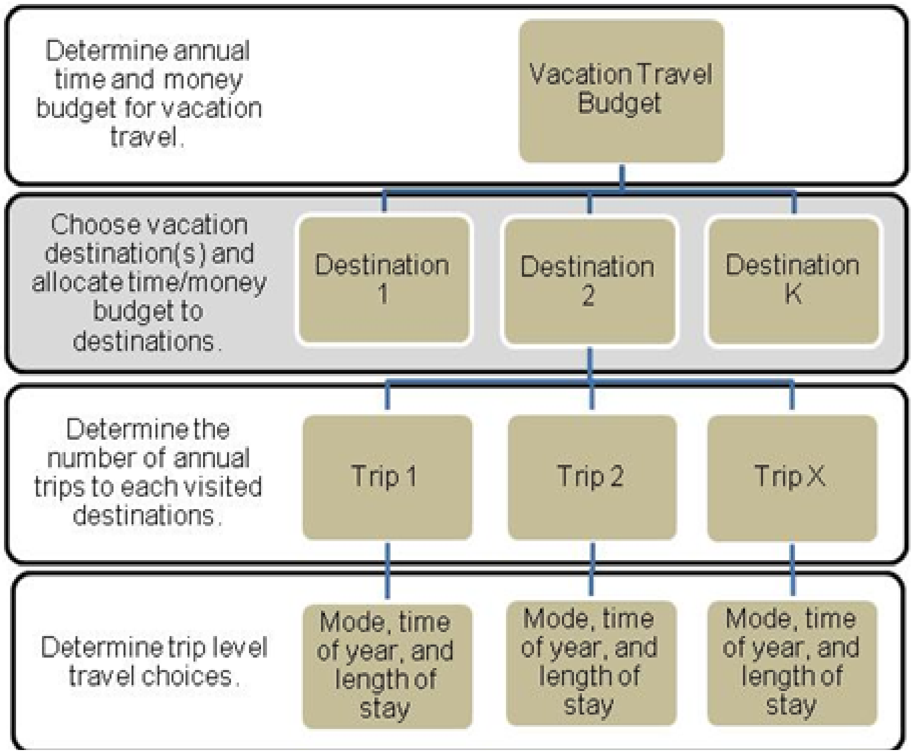

Figure 1: Modeling framework

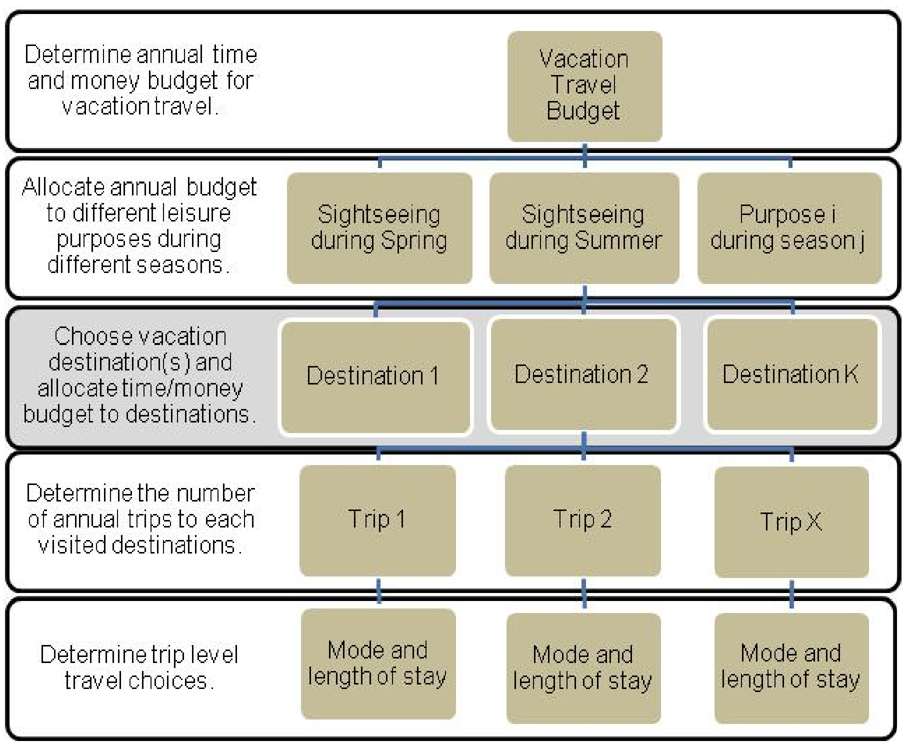

Figure 2: Alternative modeling framework

Notwithstanding which framework represents households' annual vacation decisions better (which is yet to be empirically tested), this thesis is focused on the annual vacation destination choice and time allocation decisions. Further, the thesis recognizes that mode choice decisions are generally closely tied to destination choices (Hackney, 2004) and estimates an auxiliary mode choice model that feeds the level of service characteristics into the destination choice model in the form of a log-sum variable. The empirical data used in this study comes from the 1995 American Travel Survey Data, with the U.S. divided into 210 destination choice alternatives. Thus, the study provides an opportunity to estimate, apply, and assess the performance of the MDCEV model for an empirical context with a large number of choice alternatives.

Finally, on the methodological front, we propose a variant of the MDCEV model that allows for the possibility that once a good is chosen, at least a certain reasonable amount of the good is consumed, as opposed to an unrealistically small amount of it. This is because satiation effects may start kicking in only after a certain amount of the good is consumed rather than right after the first infinitesimal consumption. In the current, long-distance vacation context, it is reasonable to expect that households allocate at least a certain minimum amount of time (say, at least half a day; as opposed to a few minutes or hours) to long-distance destinations. To accommodate such minimum required time allocation, the continuously non-linear utility functional form in the MDCEV framework is replaced with a combination of a linear and non-linear form, as described in Chapter 3.

1.4 Organization of this thesis

The remainder of this thesis is organized as follows. The next chapter will provide an extensive overview of the 1995 American Travel Survey (ATS), including a description of the household demographics and household trip file. Chapter 3 will provide a thorough explanation of the multiple discrete-continuous extreme value (MDCEV) model structure to be used for destination choice estimation in this thesis. Chapter 4 will provide a detailed methodology for the preparation of the 1995 ATS data set, including the leisure subset selection and selection of the 210 destination alternatives (4.1). The 1995 ATS does not provide level of service variables or variables indicating the attractiveness of a destination. The collection effort for these variables is also provided in Chapter 4 (4.2). Lastly, a descriptive analysis of the 1995 ATS leisure subset is provided in Chapter 4. Chapter 5 will provide the model estimation results and related discussion, followed by a model validation exercise. Finally, Chapter 6 concludes the thesis and identifies directions for possible future research.

Chapter 2: 1995 American Travel Survey

2.1 Survey description

The 1995 American Travel Survey (ATS) is the primary source of data used in this analysis. The 1995 ATS is an in-depth, long-distance nationwide travel survey of the United States that collects information on households' long-distance travel (i.e., trips of at least 100 miles) for an entire year. Admittedly, the data is a bit old, but no other recent dataset exists with information on one year worth of long-distance travel in the U.S. To be sure, a similar long-distance survey, the 2001 National Household Travel Survey (NHTS), was conducted recently to collect data on long-distance travel, although with a limited collection time per household (1 month) it is somewhat limited in the number of long-distance trips captured per household.

The ATS was conducted by the Bureau of Transportation Statistics between April 1995 and March 1996 and was designed to gather passenger flow data, as well as demographic information and other related data such as travel distance, trip purpose, mode used, length of the trip, and types of lodging used. The primary focus of the ATS is to examine long-distance trips, defined as trips with a round trip distance of 100 miles or more, excluding commuter trips (BTS, 1995). Similar data was previously collected in 1977 and so the 1995 ATS provided a much needed update.

Approximately 80,000 households taken from the 1980 Current Population Survey sample were selected to be interviewed for the ATS. Each household was interviewed three to four times, or every three months, over the course of the year to attempt to capture all long-distance trips. Computer aided telephone interviews (CATI) or computer aided personal interviews (CAPI) were utilized to attempt to limit respondent and interviewer burden. The sample for this survey consists of civilian households, group quarters (dormitories), religious group dwellings, and family-type housing on military bases. Military barracks and institutional group dwellings such as nursing homes and prisons are not included. The final number of responses is 62,609 households with 48,527 reporting at least one long-distance trip. A total of 337,520 household trips are recorded by the 1995 ATS (BTS, 1995).

Since the focus of the ATS is to provide passenger flow data, a detailed trip itinerary is included for each case. Additional details for each of the 12 potential side stops within the overarching trip including four stops to the final destination, four stops from the final destination, and four side trips originating at the final destination are provided. These include the side stop location at the metropolitan statistical area or state level, number of nights spent at the side stop, lodging accommodations utilized at the side stop, reason for the side stop, and transportation used to arrive at the side stop. This information is not provided in any later U.S. national travel survey and makes the 1995 ATS a valuable source of detailed information for long distance trip making. These additional stops are illustrated in Figure 1.

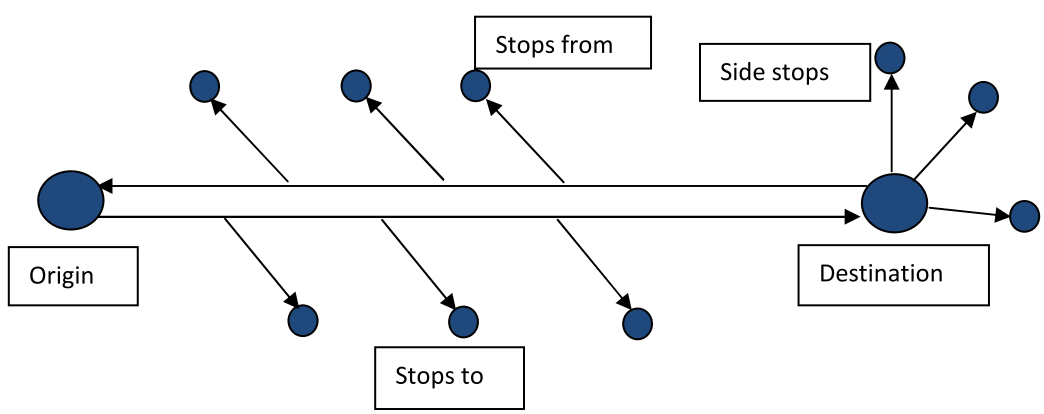

Figure 3: Description of additional stops

The 1995 ATS data is comprised of four different data sets; household trips, household demographics, person trips, and person demographics. For the purposes of this thesis, only the household data files will be used. The household demographic data set contains one record for each of the 62,609 households, of which 48,527 made at least one long distance trip during the survey year. The household demographic data set contains variables describing socio-economic characteristics (including race, education level, age, and income) and geographic characteristics (including the origin state and metropolitan statistical area). The household trip data set includes household demographic characteristics (such as age, education level, race, and household size) and trip characteristics (such as round trip distance, nights spent away, primary mode of transportation, origin, destination, and similar details on any side stops). Further details on the available variables contained within the 1995 ATS are provided in Table 1. Additionally, weights are provided within the household trip file to expand the contained trips to represent national totals. Unique household and trip identification variables are present in each of the household demographics and household trip files to allow for combining of files if necessary.

Table 1: Description of variables found in the 1995 American Travel Survey

Variable Name	Description of Variable
Race	The race of the householder or person.
Age	The age of the householder or person. This continuous variable was categorized.
Education Level	The education level of the householder or person.
Household Income	The combined annual income of the household.
Tenure	Determines if household lives in a rented or owned property.
Structure Type	The structure type of the household residence.
Household Size	The number of persons residing in the household.
Children in Household	Indicates the presence of children in the household by age.
Number of Vehicles Available	The number of personal vehicles available at the start of this trip.
Census Division Origin	The census division in which the household is located.
Census Division Destination	The census division of the primary destination of the trip.
U.S. Route Distance Traveled	The number of route miles traveled within the United States for this trip. Route miles are not available outside of the United States.
Travel Party Type	The composition of the travel party, i.e., the presence of adults and/or children.
Travelers in Party	The number of travelers in the party.
Trip Start Day	The originating day of the trip, either a weekday or weekend.
Primary Mode of Transportation	The primary mode of transportation used for this trip. This variable was re-coded as noted in the previous section.
Primary Purpose	The primary purpose for this trip. This variable was re-coded as noted in the previous section.

2.2 Description of the household demographics file

An overview of the demographic characteristics for all households recorded in the 1995 American Travel Survey is provided in Table 2. For comparison, the demographic characteristics for all households that made at least one long distance trip are provided. There are a total of 62,609 households recorded in the 1995 ATS, with

48,527 making at least one long distance trip. If applicable, the mean value of the characteristic is provided in bolded text.

The majority of surveyed householders are white (86.8 percent), while black travelers account for the second highest percentage at 8.1 percent. Approximately 70 percent of households surveyed are aged 25 to 64 with a mean age of 50.4 years. More than 85 percent of householders have attained at least a high school diploma, while just over one quarter have received a bachelor's degree or better. Almost 50 percent of those sampled for the survey make between $30,000 and $74,999 per year, falling into the middle-income category. The 1995 ATS does not provide income as a continuous variable and so a mean income is not provided. The majority of householders (58.3 percent) work full time. Retired householders make up the second largest portion of the sample, accounting for 22.8 percent of households. The average number of private vehicles available to a household is 1.89, with more than 85 percent of households having access to at least one vehicle.

The majority of survey respondents indicated they own their home, accounting for approximately three-quarters of the sample. Most (almost 80 percent) of households in the sample live in a house, duplex, or modular home. Household size is not provided as a continuous variable in the 1995 ATS, instead ending at 7 or more members of the household. Therefore, average household size could not be provided. Almost one-quarter (24.1 percent) of households consist of only one person, with two person households accounting for another 34.5 percent. This corresponds with the large proportion of households in the 1995 ATS with no children (there are no children, or no children under the age of 18 in 68.9 percent of households). This may have some impact on travel behaviors as effects of the presence of children can be very important as a result of their different needs and the additional variety needed to satisfy all members of the household. The census division variable indicates the region of the country in which the household is based. The most represented census division is the South Atlantic accounting for 16.7 percent of the households, while the Middle Atlantic accounts for 6.4 percent of the households. This appears to fairly represent the associated states, that is, the proportions seem to match the relative size of the census division. The South Atlantic division includes Delaware, Maryland, Washington, D.C., Virginia, North Carolina, South Carolina, Georgia, and Florida while the Middle Atlantic division is comprised of New York, New Jersey, and Pennsylvania.

The second column of Table 2 provides the household demographics for those households that made at least one long distance trip during the surveyed year. When comparing the entire sample of households, with sample of households that made at least one long distance trip during the survey year, there are several differences. The proportions of racial makeup between the two samples are very close to those seen in the sample of households that made at least one long distance trip. Elderly households (65 and older) tend not to make long distance trips, relative to middle aged households. This can be seen in the decrease in the elderly proportion of the sample from 24.4 percent to 19.3 percent and the decrease in the average age from 50.4 to 48.4 when comparing all households and trip making households. Households that reported at least one long distance trip tend to be better educated, with the proportion of householders with no high school diploma decreasing from 14.5 percent to 9.7 percent and the proportion of householders with a bachelor's degree or better increasing from 26.8 percent to 31.8 percent. Similarly, those households that made at least one long distance trip tend to have a higher yearly income and are more likely to be employed full time. The proportion of households that do not have a vehicle decreases in the sample of households that made a long distance trip, relative to the entire sample.

The housing characteristics (tenure and structure type) of the entire sample are only slightly different from the sample of households that made at least one long distance trip. In both cases, the majority of households owns their home and lives in a standalone house. The characteristics of the household slightly changes between the two samples. The typical household size is larger with an increase in 2 or more person households from 75.9 percent of the entire sample to 80.2 percent of the sample of households making long distance trips. Similarly, the proportion of households with no kids is lower in the sample of households that made at least one long distance trip decreasing from 68.9 percent to 65.9 percent. The shares of each census division do not change much between the entire sample and the sample of household trips with at least one long distance trip.

It is clear that the demographic characteristics of a household have some impact on the likelihood of making a long distance trip. Income and the household type (presence of kids and household size) are likely two of the major factors in the decision to travel during the year.

Table 2: Household demographics of the 1995 American Travel Survey

Characteristic	All Households	Households with at least one long distance trip
Sample Size	62,609	48,527
Race of Householder	---	---
White	86.8%	88.3%
Black	8.1%	6.5%
American Indian, Eskimo, Aleut	1.0%	1.0%
Asian or Pacific Islander	2.4%	2.5%
Other	1.7%	1.6%
Age of Householder	50.4	48.4
15 to 24	4.1%	4.5%
25 to 44	38.2%	41.1%
45 to 64	33.3%	35.1%
65 or older	24.4%	19.3%
Education of Householder	---	---
Less than high school	14.5%	9.7%
High school graduate	33.3%	31.2%
Some college, no degree	19.5%	20.9%
Associate's degree	5.8%	6.4%
Bachelor's degree	15.8%	18.4%
Some graduate or professional school, no degree	1.8%	2.2%
Graduate or professional degree	9.2%	11.2%
Household Income	---	---
Under $30,000	41.4%	34.1%
$30,000 to $74,999	49.1%	54.3%
$75,000 or more	9.5%	11.6%
Activity of Householder	---	---
Working full-time	58.3%	64.0%
Working part-time	6.4%	6.5%
Looking for work	1.4%	1.2%
In armed forces	0.5%	0.6%
Homemaker	6.0%	4.8%
Going to school	2.1%	2.3%
Retired	22.8%	18.6%
Doing something else	2.6%	1.9%
Mean number of vehicles	1.89	2.05
0	13.4%	10.7%
1	28.3%	24.9%
2	34.4%	36.9%
3	14.2%	16.1%
4 or more	9.7%	11.4%

Table 2: (continued)

Characteristic	All Households	Households with at least one long distance trip
Sample Size	62,609	48,527
Tenure	---	---
Owned or being bought	74.6%	76.5%
Rented for cash	23.5%	21.7%
No cash paid	1.9%	1.8%
Structure Type	---	---
House, townhouse, duplex, modular home	79.2%	81.4%
Apartment	13.8%	12.2%
Mobile home	5.7%	5.1%
Other	1.2%	1.3%
Household Size	---	---
1	24.1%	19.8%
2	34.5%	34.9%
3	16.5%	17.7%
4 or more	24.9%	27.7%
Presence of Children in Household	---	---
Children under 6	6.5%	7.1%
Children 6-17	18.7%	20.7%
Children under 6 and children 6-17	6.0%	6.4%
No Children	28.6%	24.7%
No children under 18	40.3%	41.2%
Census Division	---	---
New England	14.3%	13.8%
Middle Atlantic	6.4%	6.0%
East North Central	9.5%	9.3%
West North Central	12.5%	13.2%
South Atlantic	16.7%	16.1%
East South Central	9.5%	8.7%
West South Central	7.0%	6.9%
Mountain	15.2%	16.7%
Pacific	8.7%	9.4%

The aggregate household trip statistics are provided in Table 3. The mean number of trips taken annually by each household is 5.40, with 22.5 percent not taking any long-distance trips and 14.7 percent making more than 10 trips. The average household makes 2.34 long-distance trips per year, with 64.2 percent making at least one per year. The average total route distance traveled for long-distance trips within the United States is 4,572.82 per household. This does not include any travel overseas, as route distance traveled is not available outside of the United States.

Table 3: Household aggregate trip statistics

Characteristic	All Households	Households with at least one long distance trip
Sample Size	62,609	48,527
Number of trips taken	5.40	6.96
0	22.5%	0.0%
1	16.0%	20.7%
2	11.4%	14.7%
3	8.6%	11.1%
4	6.6%	8.5%
5 to 10	20.1%	26.0%
Greater than 10	14.7%	12.3%
Number of vacation trips taken	2.34	3.02
0	35.8%	17.1%
1	20.8%	26.8%
2	12.9%	16.7%
3	8.5%	11.0%
4	5.9%	7.7%
5	4.2%	5.4%
Greater than 5	11.9%	15.3%
Yearly long distance route miles traveled within US	4,572.82	5,899.80
Under 100 miles or no trips made	23.0%	0.7%*
100-2,000 miles	26.7%	34.4%
2,001-4,000 miles	16.1%	20.8%
4,001-6,000 miles	10.2%	13.1%
6,001-8,000 miles	6.7%	8.7%
8,001-10,000 miles	4.5%	5.8%
Over 10.000 miles	12.8%	16.5%

* Route distances less than 100 miles are due to international travel

2.3 Description of the household trip file

An overview of the household trips recorded in the 1995 ATS is provided in Table 4. The first numeric column provides the un-weighted descriptions for each variable and the second numeric column provides the weighted descriptions for each relevant variable. When relevant, the mean value for each variable is provided in bold text. There are 337,520 household trips provided in the 1995 ATS. When weights are applied, this represents over 684 millions trips.

Most travelers selected personal owned vehicles as the primary mode of transportation for their trip, accounting for 76.8 percent of all trips within the sample and 74.8 percent of all weighted trips. Air travel is the second most used mode of transportation, accounting for 19.3 percent of all trips within the sample and 21.0 percent of all weighted trips. The three most commonly provided reasons for taking a trip are work/business, visiting friends and relatives, and leisure accounting for 28.0 percent, 27.5 percent, and 24.8 percent of all trips within the sample respectively. Similar proportions are seen when weights are applied. The travel party composition, especially the presence of children, can potentially have some impact on travel behavior and influence the type of travel patterns used, due to additional variety seeking seen when kids and additional people are introduced to the travel party. Most trips are made by single adults with no children accounting for 58.8 percent of trips, while two adults with no children make up another 22.4 percent of all trips. Children are present on 17.4 percent of all long-distance trips. Again, similar proportions are seen for the travel party type when weights are applied. The mean number of travelers present per long-distance trip is 2.83 within the sample and 2.77 when weights are applied.

Typically, long distance household trips begin on a weekday (61.3 percent in the sample and 59.5 percent after applying weights) although when taking into account the fact that weekends constitute only 2 of 7 days per week, there does seem to be a tendency to start a trip on a weekend. The average number of nights spent away is 3.29 within the sample and 3.62 after weights are applied. In both the sample and weighted cases, approximately one quarter of all trips are completed in one day and no nights are spent away from home. The average length of all trips (excluding international travel, for which no route distances are available) is 848.25 miles, with 57.2 percent of trips ranging between 100 and 500 miles and another 18.8 percent of trips ranging between 500 and 1,000 miles. Only 3.5 percent of all recorded trips are to international destinations. Similar proportions are observed when weights are applied.

The origins and destinations show the greatest amount of variability between the un-weighted sample and the weighted total. When weights are not applied, the proportion of trips departing from a census division ranges from a low of 7.1 percent for West South Central (Oklahoma, Texas, Arkansas, and Louisiana) to 17.5 percent from the Mountain division (Idaho, Nevada, Arizona, Utah, Wyoming, Montana, Colorado, and New Mexico). The proportion of trips arriving at a census division ranges from a low of 7.3 percent to the East South Central division (Kentucky, Tennessee, Mississippi, and Alabama), to 17.7 percent to the Mountain division. When weights are applied, the proportion of trips departing from a census division ranges for a low of 4.8 percent from New England (Maine, New Hampshire, Vermont, Massachusetts, Rhode Island, and Connecticut) to 17.5 percent for the South Atlantic (Delaware, Maryland, Washington, D.C., Virginia, North Carolina, South Carolina, Georgia, and Florida). The proportion of trips arriving at a census division ranges from a low of 4.5 percent to New England, to 18.8 percent to the South Atlantic.

Table 4: Household trip statistics

Demographic Variable	Un-Weighted	Weighted
Sample Size/Population Size	337,520	684,661,562
Primary Mode of Transportation		
POV	76.8%	74.8%
Airplane	19.3%	21.0%
Bus	0.3%	0.4%
Intercity Rail	0.6%	0.6%
School Bus	0.6%	0.4%
Other	2.5%	2.7%
Purpose		
Work/Business	28.0%	27.0%
Combined Business and Pleasure	2.3%	2.2%
Shopping	2.4%	1.6%
School-related	3.1%	2.8%
Family/Personal Business	11.9%	11.0%
Visit friends or relatives	27.5%	29.4%
Leisure	24.8%	26.1%
Other	0.0%	0.0%
Travel Party Type	---	---
One adult, No children under 18	56.3%	58.8%
Two adults, No children under 18	24.6%	22.4%
Three or more adults, No children under 18	1.4%	1.3%
One adult, Children under 18	4.8%	4.4%
Two adults, Children under 18	9.5%	9.2%
Three or more adults, Children under 18	0.8%	0.8%
No adults, One child under 18	2.3%	2.6%
No adults, Two or more children under 18	0.3%	0.4%
Travelers in Party	2.83	2.77
1	35.3%	36.9%
2	32.8%	31.3%
3	12.1%	11.7%
4	9.4%	9.4%
5 or more	10.5%	10.7%

Table 4: (continued)

Demographic Variable	Un-Weighted	Weighted
Sample Size/Population Size	337,520	684,661,562
Trip Start Day		
Weekend	38.7%	40.5%
Weekday	61.3%	59.5%
Nights away from home	3.29	3.62
0	27.6%	24.3%
1 to 3	46.9%	48.0%
4 to 6	14.3%	15.4%
7 to 9	5.4%	6.0%
10 or more	5.8%	6.3%
Route Distance Traveled	848.25	867.24
100 to 500 miles	57.2%	54.0%
501 to 1,000 miles	18.8%	20.3%
1,001 to 2,000 miles	10.6%	11.1%
2,001 to 4,500 miles	7.5%	8.0%
Over 4,500 miles	2.4%	2.4%
International Destination	3.5%	4.1%
Householder Census Division Origin		
New England	12.8%	4.8%
Middle Atlantic	4.9%	11.3%
East North Central	8.8%	16.4%
West North Central	14.7%	8.9%
South Atlantic	15.0%	17.5%
East South Central	8.7%	6.4%
West South Central	7.1%	12.1%
Mountain	19.8%	7.7%
Pacific	8.1%	14.9%
Householder Census Division Destination		
New England	8.9%	4.5%
Middle Atlantic	7.7%	9.5%
East North Central	9.9%	14.2%
West North Central	13.7%	8.6%
South Atlantic	16.0%	18.8%
East South Central	7.3%	6.2%
West South Central	8.2%	11.4%
Mountain	17.7%	9.7%
Pacific	10.6%	12.9%
International	3.5%	4.1%

Chapter 3: Model Structure

The long-distance vacation travel destination choice model presented in this thesis is based on Bhat's (2005 and 2008) MDCEV framework. Thus, chapter 3.1 draws from Bhat (2008) to present the MDCEV framework for annual vacation destination choice and time allocation analysis. Chapter 3.2 extends the MDCEV framework to accommodate a given minimum amount of vacation time allocation to each of the chosen destinations.

3.1 The MDCEV model for vacation destination choice analysis

Let the U.S. be divided into K number of destination choice alternatives that a household considers for vacation travel. Let $\boldsymbol{t}$ be the vector of vacation time investments $(t_1, t_2, \ldots, t_K)$ by the household at each of the vacation destination alternatives k (k = 1,2,…,K). The time investments t_k can either be zero or some positive value expressed in number of nights spent. At least one element of $\boldsymbol{t}$ should be positive. Whether or not a specific t_k value (k = 1,2, …, K) is zero constitutes the discrete choice component, while the magnitude of each non-zero t_k value constitutes the continuous choice component.

Now, consider the following additive, non-linear, functional form[9] to represent the utility accrued by a household from its annual vacation destination choices (index for the household is suppressed in the notation):

$$U(\boldsymbol{t}) = \sum_{k=1}^{K} u(t_k) = \sum_{k=1}^{K} \gamma_k \psi_k \ln\left(\frac{t_k}{\gamma_k} + 1\right) \qquad (1)$$

---

[9] Some other utility function forms (as discussed in Bhat, 2008) were also considered, but the one presented here provided the best data fit. These alternative forms are not discussed here for conciseness.

In the above expression, the total utility $U(\boldsymbol{t})$ derived from the time allocation to the *K* destination choice alternatives is the sum of the sub-utilities $u(t_k)$ derived from the time allocation to each of the destinations *k*. Within the sub-utility function for an alternative k, ψ_k represents the marginal utility of unit vacation time investment for a destination alternative *k* at the point of zero time investment for the destination. ψ_k, labeled the baseline marginal utility parameter, controls the discrete choice decision of the household for alternative *k*. Specifically, at the point of zero time allocation to all destinations, the destination with the highest baseline marginal utility value is allocated the first unit of vacation time available to the household. Subsequently, with increasing time allocation to that destination, the marginal utility derived from spending time at that destination decreases (this diminishing marginal utility effect is called satiation). At some point, when the marginal utility for another destination becomes stronger, the next unit of time is allocated to that destination. This process of marginal time allocation to the destination with the highest marginal utility continues until the household runs out of its vacation time budget. As a result, the household derives the optimal utility from the destinations it visits and the time it allocates to each of the visited destinations. In summary, the household utility maximization problem can be viewed as an incremental time allocation process, with each additional unit of time allocated to the alternative with the highest marginal utility at that point of time allocation.

The satiation effect described above is captured in the model via a non-linear utility form with respect to the t_k terms (as in Equation (1)). In this context, the γ_k ($\gamma_k > 0, \forall k$) terms serve the role of satiation parameters by accommodating differential satiation rates across different alternatives. Specifically, the higher the γ_k value for an alternative *k*, the slower the satiation effect; hence, the amount of time allocated to alternative is larger (Bhat, 2008). Further, the γ_k terms serve as translation parameters that allow for the possibility that the household may not choose (or invest no time for) certain destinations.

In the utility function (1), socio-demographic and destination-specific attributes are introduced in the ψ_k and γ_k terms as: $\psi_k = \exp(\beta' z_k + \varepsilon_k)$ and $\gamma_k = \exp(\theta' w_k)$. z_k is the vector of exogenous variables influencing the baseline marginal utility for alternative *k*. z_k includes destination specific variables (e.g., leisure/tourism industry employment, temperature, and whether at the destination), transportation level of service variables (e.g., distance, travel times, costs), and interactions of these variables with household socio-demographic attributes. w_k is also a similar vector of variables influencing the satiation rate for alternative *k*. β and θ are parameter vectors corresponding to the explanatory variables in z_k and w_k, respectively. Finally, ε_k (*k* = 1,2,…,*K*) are the random error terms representing the unobserved factors influencing the baseline preference for each of the destination alternatives *k* (*k* = 1,2,…,*K*).

From the analyst's perspective, a household maximizes the overall utility $U(\boldsymbol{t})$ subject to the vacation time budget constraint: $\sum_k t_k = T$, where T is the annual vacation time (in number of days) available to that household.[10] The optimal time investments t_k^* (*k* = 1,2,...,*K*) can be determined by forming the Lagrangian function corresponding to the households' utility maximization problem and applying the Kuhn-Tucker (KT) conditions, as below:

[10] The reader will note here that we assume the total annual household vacation time, *T*, to be known *a priori* and focus only on households who undertake some amount of vacation travel each year. As indicated in Section 1.3, the total annual vacation time *T* could be modeled in a separate (prior) step, where the 365 days in a year would be split into non-leisure time, non-vacation leisure time (i.e., leisure time spent within the neighborhood/urban area of residence), and vacation leisure time.

$$\text{Lagrangian, L } = \sum_k \gamma_k \left[\exp(\beta' z_k + \varepsilon_k)\right] \ln\left(\frac{t_k}{\gamma_k} + 1\right) - \lambda\left[\sum_{k=1}^{K} t_k - T\right], \quad (2)$$

where λ is the Lagrangian multiplier associated with the time constraint. The Kuhn-Tucker (KT) first-order conditions for the optimal vacation time allocations (the t_k^* values) are given by:

$$\left[\exp(\beta' z_k + \varepsilon_k)\right]\left(\frac{t_k^*}{\gamma_k} + 1\right)^{-1} - \lambda = 0 \text{, if } t_k^* > 0 \text{, } k = 1, 2, \ldots, K \quad (3)$$

$$\left[\exp(\beta' z_k + \varepsilon_k)\right]\left(\frac{t_k^*}{\gamma_k} + 1\right)^{-1} - \lambda < 0 \text{, if } t_k^* = 0 \text{, } k = 1, 2, \ldots, K$$

The optimal vacation destination choices and time allocations satisfy the above KT conditions and the vacation time budget constraint $\sum_k t_k^* = T$. The budget constraint implies that only K-1 of the t_k^* values need to be estimated, since the vacation time invested for any one destination is automatically determined from the time invested for all the other destinations. To accommodate this constraint, designate destination 1 as a vacation destination to which the household allocates some non-zero amount of time. The KT condition for this destination may then be written as:

$$\lambda = \left[\exp(\beta' z_1 + \varepsilon_1)\right]\left(\frac{t_1^*}{\gamma_1} + 1\right)^{-1} \quad (4)$$

Substituting for λ from above into Equation (3) for the other destinations (k = 2, 3,…,K), and taking logarithms, the K-T conditions can be rewritten as:

$V_k + \varepsilon_k = V_1 + \varepsilon_1$ if $t_k^* > 0$ (k = 2, 3, …, K)

$V_k + \varepsilon_k < V_1 + \varepsilon_1$ if $t_k^* = 0$ (k = 2, 3, …, K), where (5)

$$V_k = \beta' z_k - \ln\left(\frac{t_k^*}{\gamma_k} + 1\right) \quad (k = 1, 2, \ldots, K)$$

Assuming that the error terms ε_k (k = 1, 2, …, K) are independent and identically distributed across alternatives with a type 1 extreme value distribution, the probability that the household allocates vacation time to the first M of the K destinations (for duration t_1^* in the first alternative, t_2^* in the second, … t_M^* in the M^{th} alternative) is (see Bhat, 2008):

$$P(t_1^*, t_2^*, t_3^*, \ldots t_M^*, 0,0,0..0) = \left[\prod_{i=1}^{M} c_i\right] \left[\sum_{i=1}^{M} \frac{1}{c_i}\right] \left[\frac{\prod_{i=1}^{M} e^{V_i}}{\left(\sum_{k=1}^{K} e^{V_k}\right)^M}\right] (M-1)!, \qquad (6)$$

where $c_i = \left(\frac{1}{t_i^* + \gamma_i}\right)$ for i = 1, 2, …, M.

3.2 The MDCEV model with minimum required consumptions

In the above discourse, the vacation time t_k (k = 1,2,…,K) is treated as a continuous variable. Thus it can potentially take a very small value (e.g., a few minutes or a few hours) that may not necessarily be realistic in a long-distance vacation travel context. As indicated earlier, it is reasonable to expect that households allocate at least a minimum amount of time (say, half a day) as opposed to a few minutes or hours for visiting long-distance destinations. However, the MDCEV model, in its original formulation, does not accommodate this and can potentially result in unrealistically small amounts of time spent for certain destinations. To address these issues, the continuously non-linear utility function of the MDCEV model (as in Equation (1)) is replaced with a combination of a linear and non-linear utility form, as below:

$$U(\boldsymbol{t}) = \sum_{k=1}^{K} u(t_k)$$

$$\text{where } \; u(t_k) = \psi_k t_k \quad \text{if } t_k \le t_0, \qquad (7)$$

$$= \psi_k t_0 + \gamma_k \psi_k \ln\left(\frac{t_k - t_0}{\gamma_k} + 1 \right) \quad \text{if } t_k \ge t_0.$$

In the above equation, t_0 is the minimum amount of time allocated to the destinations that the household chooses to visit. Thus, the utility derived from the time allocation to a destination alternative *k* (if that destination is chosen) increases in a linear fashion until the minimum required amount of time is allocated to that destination, after which the functional form takes a non-linear shape. This is depicted in Figure 4, with the linear and non-linear parts of the sub-utility functional form. The figure depicts the sub-utility profiles for ψ_k = 5 and different values of γ_k. As can be seen from the figure, The functional form of the sub-utility profiles is such that the marginal utility (*i.e.,* the slope) takes a constant value of ψ_k until the consumption reaches t_0 = 0.5, and then starts decreasing to capture the diminishing marginal returns[11]. For any chosen destination, households are assumed to experience satiation only after spending t_0 amount of time, as opposed to immediate satiation after the first unit consumption. This assumption ensures that at least t_0 amount of time is spent at any chosen destination, and helps avoid destination choices with unrealistically small amounts of time allocation.

[11] Note: marginal utility at $t_k = t_0$ is equal to ψ_k for both the linear and non-linear parts of the sub-utility curve.

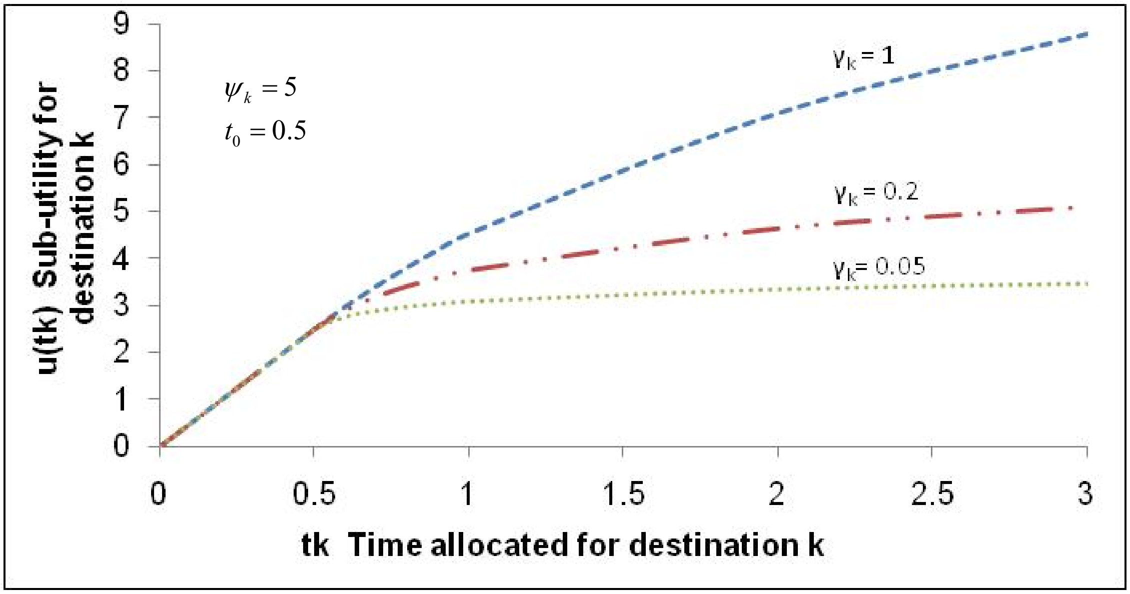

Figure 4: Sub-utility curves with a combined linear and non-linear form

To understand this, recall the incremental time allocation process discussed in Chapter 3.1. At the point of zero time allocation to all destinations, the first unit of time is allocated to the destination with the highest marginal utility (ψ_k) value. Subsequently, unlike in the case of the MDCEV model, the marginal utility of time allocation to this destination does not diminish until the time allocation reaches t_0. Given that the marginal utility of this destination remains the same (and so remains greater than the baseline marginal utility of other goods) until t_0, additional units of time are allocated to this same destination until the cumulative time allocation for this destination reaches t_0. It is only after a cumulative time allocation of t_0 that the other destinations start competing for the vacation time. As the marginal utility of time allocation for the first chosen destination diminishes (after t_0 amount of time is allocated to it), the destination with the next higher baseline utility becomes stronger (in marginal utility) and gets its first unit of time allocation. Again, until this next destination gets the minimum amount of time

(t_0) allocated, no other destination competes for vacation time. This process continues until the annual vacation time budget is exhausted. In summary, the sub-utility functional form in Equation (7) with a linear form at the corner (until a minimum amount, t_0 of time allocation), followed by a non-linear form, relaxes the assumption of immediate satiation effects at the corner (i.e., after first unit consumption). This helps ensure a minimum amount (t_0) of time allocation for each chosen destination and thus, reduce the possibility of unrealistically short vacation durations.[12]

As in chapter 3.1, the analyst can parameterize ψ_k as $\psi_k = \exp(\beta' z_k + \varepsilon_k)$ and γ_k as $\gamma_k = \exp(\theta' w_k)$, form a Lagrangian for the household's constrained utility maximization problem, and arrive at the KT conditions that form the basis for deriving the vacation destination choice and time allocation probability expressions. The Lagrangian is given by:

$$\mathrm{L} = \sum_k u(t_k) - \lambda \left[\sum_{k=1}^{K} t_k - T \right], \tag{8}$$

where $u(t_k)$ is as defined in Equation (7), and all other terms are as defined before. The KT conditions for the optimal vacation time allocations are given by:

$$u'(t_k^*) - \lambda = 0 \text{, if } t_k^* > 0 \text{, } k = 1, 2, \ldots, K \tag{9}$$

$$u'(t_k^*) - \lambda < 0 \text{, if } t_k^* = 0 \text{, } k = 1, 2, \ldots, K$$

where, $u'(t_k^*) = \dfrac{\partial}{\partial t_k}\big(U(\boldsymbol{t})\big) = \psi_k \quad \text{if } t_k^* \le t_0,$

[12] The reader will note a subtlety here that not all chosen destinations may be allocated the required minimum amount of time. Specifically, at the end of the incremental time allocation process, the last chosen destination can potentially be allocated less than required minimum amount of time simply because there is not enough time left. Thus, the model does not completely preclude destination choices with less than required amounts of time allocated. However, it should help significantly reduce such unrealistic time allocations.

$$= \psi_k \left(\frac{t_k^* - t_0}{\gamma_k} + 1 \right)^{-1} \quad \text{if } t_k^* \geq t_0.$$

Next, without any loss of generality, designating alternative 1 as a vacation destination to which the household allocates some non-zero amount of time and following the steps in chapter 3.1, the above KT conditions can be rewritten as:

$\bar{V}_k + \varepsilon_k = \bar{V}_1 + \varepsilon_1$ if $t_k^* > 0$ (k = 1, 2, …, K)

$\bar{V}_k + \varepsilon_k < \bar{V}_1 + \varepsilon_1$ if $t_k^* = 0$ (k = 1, 2, …, K), (10)

where, $\bar{V}_k = \beta' z_k \quad \text{if } t_k^* \leq t_0,$

$$= \beta' z_k - \ln\left(\frac{t_k^* - t_0}{\gamma_k} + 1 \right) \quad \text{if } t_k^* \geq t_0.$$

Assuming that the error terms ε_k (k = 1,2,…,K) are IID type 1 extreme value distributed, the probability that the household allocates vacation time to the first M of the K destinations (for duration t_1^* in the first alternative, t_2^* in the second, … t_M^* in the M^{th} alternative) is:

$$P(t_1^*, t_2^*, t_3^*, ...t_M^*, 0,0,0..0) = \left[\prod_{i=1}^{M} \bar{c}_i \right] \left[\sum_{i=1}^{M} \frac{1}{\bar{c}_i} \right] \left[\frac{\prod_{i=1}^{M} e^{\bar{V}_i}}{\left(\sum_{k=1}^{K} e^{\bar{V}_k} \right)^M} \right] (M-1)! \qquad (11)$$

The $\bar{V}_k$ terms in the above equation take an expression $\beta' z_k$ for all non-chosen destinations (i.e., alternatives for which zero time is allocated), and the expression

$\beta' z_k - \ln\left(\frac{t_k^* - t_0}{\gamma_k} + 1\right)$ for all chosen destinations. The $\overline{c}_k$ terms for all *k* = 1,2,…,*M* take an expression $\left(\frac{1}{(t_i^* - t_0) + \gamma_i}\right)$.[13]

The above probability expression can be used to form the likelihood and use the familiar maximum likelihood estimation method to estimate the parameter vectors β and θ. In this paper, the model estimation was performed using a maximum likelihood estimation code written in the GAUSS mathematical system version 9.0 (Aptech Systems, 2008).

A few notes before we move to the empirical analysis. First, we do not estimate the minimum amount of vacation time t_0 allocated to a chosen destination, but assume it *apriori* as half a day. Limited experiments to estimate t_0 with the current and other cross-sectional datasets indicate that it is unnecessary to estimate t_0. One could simply constrain t_0 as the minimum time allocated to the chosen alternatives in the data.[14] Second, the concept of minimum required consumption is not new to the consumer demand analysis literature. For example, Pollak and Wales (1992, pp. 3) discus a linear expenditure system (LES): $U(\boldsymbol{Y}) = \sum_k a_k \ln\left(y_k - b_k\right)$, in which the consumption quantities y_k must always be greater than a minimum amount b_k. Note that their LES utility function is not defined for consumption quantities below b_k. The indifference

[13] The reader will note the minor differences between the terms used in the above probability expression (i.e., $\overline{V}_k$ and $\overline{c}_k$) and the terms (V_k and c_k) in the probability expression for the original MDCEV model in Equation (6).

[14] We assume half a day as the minimum required amount of time for any chosen vacation destination. However, this is not to assert that no household ever allocates less than 0.5 days of time to visiting a long-distance destination. By specifying a certain minimum required consumption, we are building a model framework that can reduce the likelihood of unrealistically small consumptions (or time allocations).

curves implied by such an LES system are asymptotic to the consumption axes at b_k, avoiding consumptions below b_k (Pollak and Wales 1992, pp. 7). In this context, Deaton and Muellbauer (1981, pp. 65) interpret b_k as subsistence quantities or minimum required quantities that are consumed first. It is important to note, however, that the subsistence quantities discussed by them were in the context of only numeraire outside goods that are always consumed. On the other hand, our discussion is for a general case that includes inside goods that may not be consumed by some consumers (in fact, our empirical context does not have an outside good). If a good is consumed, we are able to accommodate a minimum required amount of its consumption. Besides, instead of specifying undefined utility functions for consumptions below a subsistence amount, we provide a basis for why a minimum amount is consumed by employing a combined linear – non-linear utility functional form that avoids immediate satiation effects at the corners. Third, although the proposed variant of the MDCEV model attempts to accommodate a minimum required consumption of the chosen goods, the model does not necessarily provide integer outputs for the consumptions. Vacation time is still treated as a continuous entity. However, the concept we propose here can potentially be extended to incorporate integer outputs from the utility maximization problem. Specifically, instead of combining a linear utility piece at the corner with a subsequent non-linear utility form, one can combine several linear utility pieces to form a piece-wise linear, convex utility function that provides count data outcomes from the consumers' utility maximization problem. This extension is beyond the scope of this thesis, but an important topic for further exploration.

Chapter 4: Data

4.1 1995 American Travel Survey

The 1995 American Travel Survey (ATS) is the primary source of data used in this analysis. The 1995 ATS collected information from 62,609 American households on all long-distance trips of 100 miles of more over the course of an entire year (BTS, 1995). Admittedly, the data is a bit old. However, no other recent dataset exists with information on one year worth of long-distance travel in the U.S. A similar long-distance survey was conducted along with the 2001 National Household Travel Survey (NHTS). However, the 2001 NHTS elicits long-distance travel information over the period of only four weeks.

4.1.1 Data set preparation

The first step was to recode certain continuous variables into categorical variables for ease of interpretation. The chosen categories are based on previous literature and intuition. Income was re-coded into the same high (greater than $75,000), middle ($30,000 to $75,000), and low (less than $30,000) categories used by LaMondia and Bhat (2008). Round trip distance traveled for each trip was also recoded into several categories including: 100 to 500 miles, 501 to 1,000 miles, 1,001 to 2,000 miles, 2,001 to 4,500 miles, and greater than 4,500 miles. These categories were selected based on the paper by LaMondia and Bhat (2008) which also utilized the 1995 ATS to study long distance leisure travel. Age was divided into five groups; under 15, 15 to 24, 25 to 44, 45 to 64, and over 64. These age ranges were selected to match the U.S. 2000 Census. Transportation mode and primary purpose for the trip have been re-coded into more generalized, easier to interpret categories using similar methodology as that used by Hu and Young (2001) although with some modification to the trip purpose. Tables 5 and 6 below indicates the methods used. By dividing these variables into more aggregated groups, the values can provide greater meaning to the reader.

Table 5: Reclassification of modes from 1995 ATS

New Mode	1995 American Travel Survey Mode
Private ground	Car, pickup truck, or van
	Other truck
	Rental car, truck, or van
	Recreational vehicle or motor home
	Motorcycle, moped, or motor bicycle
Commercial Air	Commercial Airplane
Other	City to City Bus
	Charter bus or tour bus
	School bus
	Train
	Taxi
	Ship or boat
	Cruise ship
	Passenger line or ferry
	Recreational boat, sailboat, pleasure boat, or yacht
	Bicycle
	Other

Table 6: Recoding methodology for trip purpose and transportation

Purpose	Transportation
Business	**POV**
Business	Car, pickup truck, or van
Convention, seminar, or conference	Other truck
Combined Business and Pleasure	Rental car, truck, or van
Combined Business and Pleasure	Recreational vehicle
Shopping	Motorcycle
Shopping	**Airplane**
School-related	Commercial airplane
School-related	**Bus**
Personal, family, or medical	City to City bus
Personal, family, or medical	**Intercity Rail**
Visit relatives or friends	Intercity train
Visit relatives or friends	**School bus**
Leisure	School bus
Rest or relaxation	**Other**
Sightseeing, or to visit a historic or scenic attraction	Corporate/personal airplane
Outdoor recreation	Charter bus or tour bus
Entertainment	Ship or boat
Change trans/Spend Night/Passenger	Cruise Ship
Spend the night	Passenger line or ferry
Transfer from one airplane to another	Recreational boat, sailboat, or yacht
Change to a different type of transportation	Taxi
Drop off or pick up a passenger	Bicycle
Other	Other
Other	

4.1.2 Leisure subset selection

Out of all the surveyed households in the 1995 ATS sample, 48,527 reported at least one long-distance trip[15]. As such, a total of 337,520 trips were reported, along with the information on the purpose, mode, and destination of travel and other travel attributes.

The scope of analysis of the current thesis consists of long distance leisure/vacation travel within the United States. Therefore, only households that made at least one long-distance trip for the purpose of relaxation, sightseeing, outdoor recreation, or entertainment were considered. Trips for visiting friends and family were not

[15] The reader will note here that in this survey a trip is defined as a travel out of home that eventually returns home (which is usually called a tour in traditional metropolitan area travel modeling context).

considered in this study. This is because the factors that underlie the destination choice decisions for this type of trips are quite different from the trips for other purposes. Specifically, the primary determinants of a household's destination choices for visiting purpose may be the location of family and friends (i.e., social networks), rather than the destination characteristics themselves. Unfortunately, the data does not contain any social networks information.

Of the 337,520 trips reported in the 1995 ATS, 25% were for leisure purposes (i.e., relaxation, sightseeing, outdoor recreation, or entertainment) made by 28,210 households. From these households, a small percentage of those who traveled to destinations outside the United States were removed for the purpose of the current study (3.5% of all leisure trips were made to international destinations).[16] Next, only households that used private ground (i.e., auto) and commercial air modes of travel were considered (this was approximately 94% of the data as seen in Table 7). While it is desirable to include these other households as well (especially those that use the inter-city bus and rail modes and water modes), it was very difficult to gather the transportation network and level of service characteristics for these modes for the year 1995. For this same reason, Hawaii was excluded as a destination (or origin). Thus, the analysis is limited to the contiguous states of the U.S. After further processing to clean households with missing information on important variables (income, travel information for a big chunk of the year), the dataset was still sizeable with 22,215 households that made 57,989 long-distance leisure trips. 6000 of these households were randomly sampled to estimate the destination choice MDCEV model, while another 715 (again randomly sampled) were kept for validation purposes.

[16] Considering international destinations adds a layer of complexity to the model in terms of increasing the number of alternative destinations in the choice set. Besides, the data does not contain information on which country the trip was made to.

Table 7: Primary mode used for long distance leisure travel - 1995 ATS

	Frequency	Percentage
Private ground	55,606	83.8%
Commercial air	6,704	10.1%
Other	4,014	6.1%
Total	66,324	100.0%

4.1.3 Destination alternatives

For the current analysis, the United States was divided into 210 alterative destinations. Specifically, each of the Metropolitan Statistical Areas (MSAs) from each state was counted as a destination alternative, resulting in 162 MSA destinations. Then, the remaining non-MSA area in each state was counted as a single destination (one non-MSA area for each state, with the exception of Rhode Island which was entirely included in the Falls River-Warwick MSA). This resulted in 48 non-MSA destinations. All together, the U.S. was divided into 210 destinations (162 MSAs plus 48 non-MSAs). While it is desirable to divide the non-MSAs into smaller and more meaningful geographies, the destinations reported in the data did not provide any further information other than MSAs or non-MSAs. The labels of the destinations are provided in Table 8 below.

Table 8: Destination alternatives

Origin State	Origin MSA	Origin State	Origin MSA
AL	Birmingham, AL MSA	CT	Connecticut - Not in MSA
AL	Huntsville, AL MSA	DE	Wilmington, DE PMSA
AL	Mobile, AL MSA	DE	Delaware - Not in MSA
AL	Montgomery, AL MSA	DC	Washington, DC-MD-Va PMSA
AL	Alabama - Not in MSA	FL	Daytona Beach, FL MSA
AK	Anchorage, AK MSA	FL	Fort Lauderdale, FL PMSA
AK	Alaska - Not in MSA	FL	Fort Myers-Cape Coral, FL MSA
AZ	Phoenix-Mesa, AZ MSA	FL	Jacksonville, FL MSA
AZ	Tucson, AZ MSA	FL	Lakeland-Winter Haven, FL MSA
AZ	Arizona - Not in MSA	FL	Melbourne-Titusville-Palm Bay, FL MSA
AR	Little Rock-North Little Rock, AR MSA	FL	Miami, FL PMSA
AR	Arkansas - Not in MSA	FL	Orlando, FL MSA
CA	Bakersfield, CA MSA	FL	Pensacola, FL MSA
CA	Fresno, CA MSA	FL	Sarasota-Bradenton, FL MSA
CA	Los Angeles-Long Beach, CA PMSA	FL	Tallahassee, FL MSA
CA	Modesto, CA MSA	FL	Tampa-St. Petersburg-Clearwater, FL MSA
CA	Oakland, CA PMSA	FL	West Palm Beach-Boca Raton, FL MSA
CA	Orange County, CA PMSA	FL	Florida - Not in MSA
CA	Riverside-San Bernardino, CA PMSA	GA	Atlanta, GA MSA
CA	Sacremento, CA PMSA	GA	Augusta, GA MSA
CA	Salinas, CA MSA	GA	Macon, GA MSA
CA	San Diego, CA MSA	GA	Georgia - Not in MSA
CA	San Francisco, CA PMSA	ID	Boise City, ID MSA
CA	San Jose, CA PMSA	ID	Idaho - Not in MSA
CA	Santa Barbara-Santa Maria-Lompoc, CA MSA	IL	Chicago, IL PMSA
CA	Santa Rosa, CA PMSA	IL	Peoria-Pekin, IL MSA
CA	Stockton-Lodi, CA MSA	IL	Rockford, IL MSA
CA	Vallejo-Fairfield-Napa, CA PMSA	IL	St. Louis, MO-IL MSA
CA	Ventura, CA PMSA	IL	Illinois - Not in MSA
CA	California - Not in MSA	IN	Fort Wayne, IN MSA
CO	Boulder-Longmont, CO PMSA	IN	Gary, IN PMSA
CO	Colorado, CO MSA	IN	Indianapolis, In MSA
CO	Denver, CO PMSA	IN	South Bend, IN MSA
CO	Colorado - Not in MSA	IN	Indiana - Not in MSA
CT	Bridgeport, CT PMSA	IA	Des Moines, IA MSA
CT	Hartford, CT MSA	IA	Iowa - Not in MSA
CT	New Haven-Meriden, CT PMSA	KS	Wichita, KS MSA
CT	New London-Norwich, CT MSA	KS	Kansas - Not in MSA
CT	Stamford-Norwalk, CT PMSA	KY	Lexington, KY MSA

Table 8: (continued)

Origin State	Origin MSA	Origin State	Origin MSA
KY	Louisville, KY MSA	NJ	Middlesex-Somerset-Hunterdon, NJ PMSA
KY	Kentucky - Not in MSA	NJ	Monmouth-Ocean, NJ PMSA
LA	Baton Rouge, LA MSA	NJ	Newark, NJ PMSA
LA	New Orleans, LA MSA	NJ	Trenton, NJ PMSA
LA	Shreveport-Bossier City, LA MSA	NJ	New Jersey - Not in MSA
LA	Louisiana - Not in MSA	NM	Albuquerque, NM MSA
ME	Maine - Not in MSA	NM	New Mexico - Not in MSA
MD	Baltimore, MD PMSA	NY	Albany-Schenectady-Troy, NY MSA
MD	Maryland - Not in MSA	NY	Binghamton, NY MSA
MA	Boston, MA PMSA	NY	Buffalo-Niagara Falls, NY MSA
MA	Lowell, MA PMSA	NY	Dutchess County, NY PMSA
MA	Springfield, MA MSA	NY	Nassau-Suffolk, NY PMSA
MA	Worcester, MA PMSA	NY	New York, NY PMSA
MA	Massachusetts - Not in MSA	NY	Newburgh, NY PMSA
MI	Ann Arbor, MI PMSA	NY	Rochester, NY MSA
MI	Detroit, MI PMSA	NY	Syracuse, NY MSA
MI	Flint, MI PMSA	NY	Utica-Rome, NY MSA
MI	Grand Rapids-Muskegon-Holland, MI MSA	NY	New York - Not in MSA
MI	Kalamazoo-Battle Creek, MI MSA	NC	Charlotte-Gastonia, NC MSA
MI	Lansing-East Lansing, MI MSA	NC	Fayetteville, NC MSA
MI	Saginaw-Midland, MI MSA	NC	Greenboro-Winston-Salem-High Point, NC MSA
MI	Michigan - Not in MSA	NC	Hickory-Morganton, NC MSA
MN	Minneapolis-St. Paul, MN MSA	NC	Raleigh-Durham-Chapel Hill, NC MSA
MN	Minnesota - Not in MSA	NC	North Carolina - Not in MSA
MS	Jackson, MS MSA	ND	North Dakota - Not in MSA
MS	Mississippi - Not in MSA	OH	Akron, OH PMSA
MO	Kansas City, MO-KS MSA	OH	Conton-Massillon, OH MSA
MO	Springfield, MO MSA	OH	Cincinnati, OH-KY PMSA
MO	Missouri - Not in MSA	OH	Cleveland-Lorain-Elyria, OH PMSA
MT	Montana - Not in MSA	OH	Columbus, OH MSA
NE	Omaha, NE MSA	OH	Dayton-Springfield, OH MSA
NE	Nebraska - Not in MSA	OH	Hamilton-Middletown, OH PMSA
NV	Las Vegas, NV MSA	OH	Toldeo, OH MSA
NV	Reno, NV MSA	OH	Youngstown-Warren, OH MSA
NV	Nevada - Not in MSA	OH	Ohio - Not in MSA
NH	New Hampshire - Not in MSA	OK	Oklahoma City, OK MSA
NJ	Atlantic-Cape May, NJ PMSA	OK	Tulsa, OK MSA
NJ	Bergen-Passaic, NJ PMSA	OK	Oklahoma - Not in MSA
NJ	Jersey City, NJ PMSA	OR	Eugene-Springfield, OR MSA

Table 8: (continued)

Origin State	Origin MSA	Origin State	Origin MSA
OR	Portland-Vancouver, OR-WA PMSA	TX	Beaumont-Port Arthur, TX MSA
OR	Salem, OR PMSA	TX	Corpus Christi, TX MSA
OR	Oregon - Not in MSA	TX	Dallas, TX PMSA
PA	Allentown-Bethlehem-Easton, PA MSA	TX	El Paso, TX MSA
PA	Erie, PA MSA	TX	Fort Worth-Arlington, TX PMSA
PA	Harrisburg-Carlisle, PA MSA	TX	Houston, TX PMSA
PA	Lancaster, PA MSA	TX	Mcallen-Edinburg-Mission, TX MSA
PA	Philadelphia, PA-NJ PMSA	TX	San Antonio, TX MSA
PA	Pittsburgh, PA MSA	TX	Texas - Not in MSA
PA	Reading, PA MSA	UT	Provo-Orem, UT MSA
PA	Scranton-Wilkes Barre-Hazleton, PA MSA	UT	Salt Lake City-Ogden, UT MSA
PA	York, PA MSA	UT	Utah - Not in MSA
PA	Pennsylvania - Not in MSA	VT	Vermont - Not in MSA
RI	Providence-Fall River-Warwick, RI MSA	VA	Norfolk-Virginia Beach-Newport News, VA MSA
RI	Rhode Island - Not in MSA	VA	Richmond, VA MSA
SC	Charleston-North Charleston, SC MSA	VA	Virginia - Not in MSA
SC	Columbia, SC MSA	WA	Seattle-Bellevue-Everett, WA PMSA
SC	Greenville-Spartanburg, SC MSA	WA	Spokane, WA MSA
SC	South Carolina - Not in MSA	WA	Tacoma, WA PMSA
SD	South Dakota - Not in MSA	WA	Washington - Not in MSA
TN	Chattonooga, TN MSA	WV	Charleston, WV MSA
TN	Johnson City-Kingsport-Bristol, TN MSA	WV	West Virginia - Not in MSA
TN	Knoxville, TN MSA	WI	Appleton-Oshkosh-Neenah, WI MSA
TN	Memphis, TN MSA	WI	Madison, WI MSA
TN	Nashville, TN MSA	WI	Milwaukee-Waukesha, WI PMSA
TN	Tennessee - Not in MSA	WI	Wisconsin - Not in MSA
TX	Austin-San Marcos, TX MSA	WY	Wyoming - Not in MSA

4.2 Secondary data sources

4.2.1 Level of service variables

In addition to the data provided by the 1995 American Travel Survey (ATS), several secondary data sources were utilized to compile other required information such as: (1) the level of service variables, included as travel times and costs between each origin-destination pair via air and auto modes, (2) destination size and attraction variables for the year 1995, including land area, number of employees in different sectors (leisure and/or hospitality, retail, etc.), total population, total gross domestic product, and gross domestic product for amusement and recreation, and (3) destination climate variables including mean monthly temperatures for different months in a year, miles of coastline at the destination, and the annual number of freezing days experienced at the destination. Gathering all this information required a significant amount of effort from multiple data sources.

Ground travel times and costs were derived as a function of ground route distances between each metropolitan statistical area (MSA) or non-MSA area. It is assumed that route distances would not significantly change in the context of long-distance travel between 1995 and 2010. Microsoft MapPoint 2010 software, in conjunction with its Mile Charter add-on, was used to plot route distances between each origin-destination pair (Microsoft, 2009; Winwaed Software Technology, 2009). The Mile Charter add-on provides both the route distance, and travel times between all origins and destinations in a simple matrix format. Some additional work was required to aggregate origins and destinations from the city level, to the metropolitan statistical area (MSA). When an MSA is made up of more than one city, the average distance between each city to all possible destinations is taken. A similar methodology applies to a destination MSA comprised of more than one city. For example, take MSA X as being comprised of two cities, City 1 and City 2, while MSA Y is comprised of two cities, City 3 and City 4. The route distance between these cities is the average of the distance between City1 and Cities 3 and 4 and City2 and Cities 3 and 4. It is not known from the data which city within the MSA is the origin or destination so this provides the closest proxy. When considering non-MSA areas, the level of service variables are more difficult to derive. For MSA to non-MSA travel (or vice versa), the non-MSA area is taken as the centroid of the state. The exception to this rule is the case where the origin and destination are within the same state, in which case, the average travel distance between the MSA area and the opposite borders of the state are used. When both the origin and destination are non-MSA areas, the averages of all MSA to MSA routes within that state are taken. This applies to both same state and different state combinations. For some of these non-MSA to non-MSA cases, this is not possible as there is 1 or fewer MSA areas within the state as defined by the 1995 ATS. In these cases, the route distances and travel times are taken as those between a city near the border and a centrally located city as shown in Table 9. While these distances, especially for non-MSA areas, are not perfectly accurate, they do provide a reasonable assumption of travel distance.

Table 9: Non-MSA to non-MSA area proxies for select states

State	Origin city	Destination city
Alaska	Fairbanks	Anchorage
Maine	Dover-Foxcroft	Portland
Montana	Hobson	Libby
New Hampshire	Concord	Gorham
North Dakota	Underwood	Marmarth
South Dakota	Fort Pierre	Sioux Falls
Vermont	Morrisville	Bennington
Wyoming	Casper	Jackson

Travel costs were derived as a function of travel distance, and average vehicle miles per gallon. Lim (1997) found that private gasoline costs between origin and destination are often used as a proxy for surface travel in tourism demand models. While the cost of the vehicle, insurance, and maintenance are all a part of the trip, these costs are paid separately from the cost of the trip and likely would not be considered. The average vehicle fuel efficiency in 1996 was 19.7 miles per gallon (Grush, 1998).The cost of gas is taken from the Energy Information Administration which provides gas prices for 1995 by region within the United States (Energy Information Administration, 1995). It is assumed that while gas prices do vary by geographic area, they would not vary as much within each region. To find the gas price paid by the traveler, an average gas price between the origin region and the destination region was taken.

Air fare and air travel times were both taken from the Airline Origin and Destination Survey (DB1B) provided through Transtats from the Research and Innovation Technology Administration (RITA) at the USDOT (BTS, 1995). Due to smaller sample sizes for lesser traveled routes the years 1994, 1995, and 1996 were used to expand the sample size. The DB1B survey is comprised of three main data sources: the market, itinerary, and coupon data sets. The two data sets used for this thesis are the market and coupon surveys. The market survey provides market fare, market distance traveled (actual distance traveled), nonstop miles (GCD distance traveled), and the airport group (airport codes of all airports within the itinerary including origin and destination).The coupon survey provided the fare class (coach, business, etc.) for each ticket. To eliminate any fares that only cover tax, but not the base fare, all fares less than fifty dollars were eliminated (National Transportation Library, 2010). Secondly, all first class/business fares were eliminated. This was done to reduce the variance amongst traveler costs and since the majority of travelers typically travel coach class it was considered reasonable. The average cruising speed of a Boeing 757 (500 miles per hour) was taken as the average air speed (Boeing, 2011).

Four "ODPAIR" variables were created as a concatenation of the origin and destination using SPSS within the DB1B survey. Table 10 provides the methodology used to create each "ODPAIR" variable. These variables were created to mimic the methodology used for finding ground distance for between destinations.

Table 10: Origin-destination pair variables created from DB1B survey

Variable	O-D Pair Type	Origin	Destination
ODPAIR1	MSA to MSA	airport code	airport code
ODPAIR2	non-MSA to MSA	state code	airport code
ODPAIR3	MSA to non-MSA	airport code	state code
ODPAIR4	Non-MSA to non-MSA	state code	state code

In addition to the "ODPAIR" variables, a variable indicating the number of layovers was created for each case. The airport group variable consists of each airport code for the trip, including the origin and destination, each separated by a colon. Since each airport code is three characters long, the number of characters in this variable is directly related to the number of layovers. SPSS provides a function to compute the length of a given variable and so a new variable called "layovers" was created based on the character length of the airport group variable. For example, if the airport group variable was seven, then there were no layovers. For every four characters beyond the first seven there was one additional layover (airport code plus colon). The aggregate function in SPSS using each of the "ODPAIR" variables as the break variable was used to find the mean market fare, mean market distance, mean nonstop distance, and mean layovers.

Each airport code is associated with one or more MSAs using the Places Rated Almanac (Savageau and Loftus, 1997). Each state code is used as a proxy for the given state as an origin or destination and accounts for all airports within the given state. The only state with no airports identified in the DB1B for the 1994, 1995, and 1996 years was Delaware. It was decided that Philadelphia International Airport (PHL) should be used as this is the airport assigned to both Wilmington and Dover, the two MSA areas identified in the Places Rated Almanac (Savageau and Loftus, 1997). Several MSAs are served by multiple airports. This was dealt with in a similar manner to the ground distance and travel times. The average of all possible connections was taken and used a proxy for the given origin-destination pair.

The actual cost associated with commercial air travel is a function of both the ticket prices, and the party size. Unlike the private car where the marginal cost of another passenger can be considered negligible, the cost increases by the amount of an individual's airfare for each additional party member. In an exploratory analysis of the 1995 ATS, it was found that party size varies for only 10 percent of household trips to a given destination. For these cases in which party size does vary, the variation is typically quite low, with the difference typically being within one or two people. To avoid the issue of determining the party size for all potential trips, the average party size was taken for each household and used to compute air costs in the destination choice model.

4.2.2 Destination attraction variables

Data for several indices for the attractiveness of a destination were selected. These include the number of leisure and/or hospitality employees, the number of retail employees, the number of total non-farm employees, the total population, land area, gross domestic product (including individual industries), miles of coastline, mean monthly temperature, number of freezing days per annum.

State and local employment levels were taken from the Bureau of Labor Statistics (BLS) (Bureau of Labor Statistics, 1995). The number of employees within each industry in a given metropolitan statistical area (MSA) was taken as the sum of all cities within the MSA. Statewide totals of employment, less the number of employees for each MSA within the given state, were used for non-MSA areas. The statewide employment data for Rhode Island is taken as zero since the Providence MSA covers the entire state. In some cases, more than one value is given for a large MSA. In these cases, the metropolitan division was used as it does not overlap with adjacent, but separate, MSAs. Five MSAs cross state borders including Philadelphia, PA-NJ, Kansas City, MO-KS, St. Louis MO-IL, Portland OR-WA, and Providence-Fall River-Warwick, RI-MA. With the exception of Kansas City, the majority of the MSA falls within a single state and so the MSA was assumed to fall completely within that state. The employment values for Kansas City were provided separately for Kansas and Missouri. In a few rare cases, the definition of an individual MSA was different from those defined in the 1995 ATS. Bridgeport, CT and Stamford-Norwalk, CT are considered as one MSA and so the same employment totals were used for both. Cincinnati, OH and Hamilton-Middletown are also considered as one MSA and so the same employment totals were used in this case as well. In each of these cases, the employment totals were only subtracted from the statewide totals once.

Population and land areas for each MSA and non-MSA area were taken from the 2000 Census (U.S. Census Bureau, 2000). In addition to employment within key leisure related industries, the gross domestic product at the state level for all industries, amusement and recreation services, and hotels and other services was taken from the Bureau of Economic Analysis (Bureau of Economic Analysis, 1995) of the U.S. Department of Commerce.

Miles of coastline and several climate variables were also included in the destination attraction data set. The total miles of coastline, including the Great Lakes, was taken from the National Oceanic and Atmospheric Administration's Ocean and Coastal Resource Management (NOAA, 2011). Information on the mean monthly temperatures for both January and June and the total number of freezing days was obtained from the *Places Rated Almanac* (Savageau and Loftus, 1997) for the year 1995 for all MSA areas. The same information for non-MSA areas was considered as the average values of the MSA areas within the state (Savageau and Loftus, 1997).

4.3 1995 ATS leisure subset data description

4.3.1 Household demographics

Table 11 provides an overview of the socio-demographic makeup and the leisure travel characteristics of all households surveyed within the 1995 ATS, the 1995 ATS leisure subset (the subset of households who made at least one leisure trip, obtained after the cleaning process explained in the “Primary Data Source” section), and a random sample of 6,000 households from the leisure subset utilized for the destination choice model estimation. There are a total of 62,609 households in the 1995 ATS, 22,215 households within the leisure subset, and 6,000 households were sampled from the leisure subset for the destination choice model estimation.

The average age of households in the 1995 ATS is 50, and drops to approximately 46 in the leisure subset and the estimation sample. The elderly (65 or older) are less represented in the leisure datasets; suggesting that the elderly are less likely to take long-distance leisure trips. In terms of annual income, households who made leisure trips appear to be slightly more affluent than the general ATS sample. This makes intuitive sense as long-distance leisure travel can be considered as a luxury which those with very low incomes are unlikely to be able to afford. Two person households account for the highest proportion of household size in the data. Both the leisure subset and the estimation sample tend to have somewhat larger households than the overall 1995 ATS sample. There may be several reasons for this, including the presence of children, for whom a household may tend to make leisure trips. The majority of households within the leisure subset are married, accounting for nearly 70 percent of the sample. Approximately half of these married couples have children, which likely has some impact on the leisure travel pattern and destination choices.

Table 11: Household demographics and leisure travel characteristics in 1995 ATS

Household Characteristics	**1995 ATS**	**Leisure Subset***	**Model estimation Sample**
Sample size	62,609	22,215	6,000
Age of householder	50.4	46.4	46.7
15 to 24	4.1%	4.5%	3.9%
25 to 44	38.2%	45.4%	45.4%
45 to 64	33.3%	35.9%	36.0%
65 or older	24.4%	14.2%	14.7%
Household yearly income			
Under $30,000	33.1%	27.1%	26.1%
$30,000 to $74,999	57.4%	60.2%	60.7%
$75,000 or more	9.5%	12.7%	13.2%
Household size			
1	24.1%	15.5%	15.7%
2	34.5%	34.3%	34.7%
3	16.5%	18.8%	18.5%
4 or more	24.9%	31.4%	31.1%
Household type			
Married couple family – with children under 18	25.3%	33.5%	32.9%
Married couple family – no children	33.7%	35.2%	35.4%
Other family – with children under 18	5.8%	5.4%	5.2%
Other family – no children	6.6%	5.2%	5.3%
Non family – not living alone	4.4%	5.2%	5.6%
Non family – living alone	24.2%	15.5%	15.8%
Household Leisure Travel Characteristics	**1995 ATS**	**Leisure Subset***	**Model estimation Sample**
Number of long distance leisure trips	---	2.61	2.64
1	---	47.9%	46.9%
2	---	21.3%	22.0%
3	---	11.6%	11.7%
4	---	6.5%	7.3%
5 or more	---	12.7%	12.1%
Number of destinations visited	---	---	---
1	---	60.7%	60.1%
2	---	24.3%	25.1%
3	---	9.6%	9.3%
4	---	3.3%	3.5%
5 or more	---	2.1%	2.0%
Number of trips made to a destination**	---	---	---
1	---	78.3%	78.5%
2	---	11.5%	10.9%
3	---	4.2%	4.7%
4	---	2.0%	1.9%
5 or more	---	4.0%	4.0%

*Leisure subset: Subset of households who made at least one leisure trip in the year.
**These proportions are of all destinations visited by each household.

The next set of rows provides an overview of the leisure travel characteristics of those households who made at least one leisure trip in the year. Several observations can be made from the leisure subset column. First, on average, these households (who made at least one leisure trip) made 2.61 leisure trips per year, with 52.1% making more than one trip per year. Second, close to 40% of these households visited more than one destination. Third, 78.3% of the households visit a destination (if they do so) only once. That is, several households are likely to visit multiple destinations per year, but less likely to re-visit a destination. This suggests multiple discreteness (choosing multiple destinations) and variety-seeking behavior in households' destination choices. As discussed earlier, the multiple discreteness or variety in households' leisure destination choices comes from several reasons, including the satiation effects of increasing time allocation to one destination, and the presence of different persons with a variety of preferences in the household. Similar inferences can be made from the model estimation households' sample as well (the last column in the data).

Traditional discrete choice models assume that the destination choice alternatives are perfect substitutes of each other. Thus, it is difficult to use the framework for the current situation with multiple destination choices. This is not to say that one cannot use discrete choice models for the current situation (e.g., a repeated discrete choice framework can be used; see Herriges and Phaneuf, 2002). However, it is cumbersome to do so. Further, such approaches are not based on a unifying utility maximizing framework. The multiple discrete-continuous extreme value (MDCEV) model (Bhat, 2005; Bhat, 2008) on the other hand, is based on a unifying utility maximizing framework for modeling multiple discreteness. Given the total number of days per year a household allocates to vacation, the analyst can use the MDCEV model to simultaneously analyze all the destinations visited by the household in a year, and the time allocations to each destination. In addition, the model accommodates satiation effects (hence variety seeking) through a non-linear utility framework (Kim et al., 2002), and recognizes that households operate under time budgets via a constrained utility maximization framework.

4.3.2 Household trips

Table 12 provides an overview of the trip level characteristics of the 1995 ATS data for the entire data, for the trips made by the leisure subset, and for the trips made by the households n the estimation sample). The ATS contains records for 337,520 household trips, of which 57,889 were leisure trips made by households in the leisure subset and 15,826 by the 6,000 household model estimation sample. The vast majority of all trips made in the 1995 ATS utilize either private ground modes, or commercial air modes, accounting for 96.1 percent of all trips. It is for this reason that the scope of this thesis is confined to these two dominant modes of transportation. In both the leisure subset and model estimation sample, approximately 90 percent of trips are made using private ground modes and approximately 10 percent are made using commercial air modes of transportation. In terms of trip distance, the majority of leisure trips are less than 500 miles, with the proportions of trips declining as distance increases. This makes intuitive sense as longer distances typically equate to higher costs and travel times. The average number of nights (away from home) spent on each vacation trip is 3.29 for the 1995 ATS, increasing slightly to 3.39 for leisure trips. Just under one-quarter of leisure trips are day trips, which do not involve spending a night away from home. The highest proportion of nights spent at the destination is two with a greater share of trips relative to any other number of nights spent away. While additional nights spent at the destination would increase the cost of the trip, travelers may also be less likely to spend too little time at a destination due to the already expended time and cost involved with traveling.

For the current analysis and modeling purposes, the number of nights variable was effectively considered as the number of days (away from home) spent on the trip. For all day trips, it was considered that half a day was spent on the trip. For each household, the sum of all the days spent across all the visited destinations was considered as the annual household long-distance vacation time budget, *T*. This annual long-distance vacation time budget varied from 0.5 (i.e., a single day trip) to as much as 352.50 days, with an average value of 9.11 days in the leisure subset data (and similar values in the model estimation data).

Table 12: Leisure trip characteristics in 1995 ATS

Characteristic	1995 ATS	Leisure Subset	Model Estimation Sample
Sample size	337,520	57,989	15,826
Primary mode of transportation	---	---	---
Private ground	76.8%	89.3%	89.5%
Commercial air	19.3%	10.7%	10.5%
Other	3.9%	---	---
Round trip U.S. route distance (miles)	848.25	780.13	770.86
International Destination	3.5%	---	---
100 to 500 miles	57.2%	59.9%	61.1%
501 to 1,000 miles	18.8%	20.6%	19.9%
1,001 to 2,000 miles	10.6%	10.1%	9.6%
2,001 to 4,500 miles	7.5%	8.0%	7.9%
Over 4,500 miles	2.4%	1.5%	1.5%
No. of nights away from home on trip*	3.29	3.39	3.39
0 (day trip)	27.6%	21.0%	21.3%
1	15.7%	14.8%	16.1%
2	20.2%	24.2%	23.2%
3	11.0%	12.6%	12.4%
4	7.7%	8.4%	8.0%
5	4.1%	4.1%	4.3%
6	2.5%	3.2%	3.2%
7	3.1%	4.4%	4.2%
8	1.3%	1.5%	1.6%
9	0.9%	1.2%	1.0%
10 or more	5.9%	4.6%	4.6%

* For current analysis, the number of nights variable was effectively considered as the number of days (away from home) spent on the trip. For day trips, it was assumed that half a day (0.5 days) was spent on the trip.

Finally, we conducted an exploratory analysis of the mode choices for long-distance leisure trips in the data (not shown in the Tables). Specifically, we explored if households changed their mode choices across the different destinations they visited, as well as across the different trips they made to a single destination. The analysis indicates, as expected, that households did change their mode choices across the different destinations they visited. That is, a household's mode choices may vary across the different destinations they visit, depending on the transportation level of service characteristics to the destinations by different modes (and household characteristics). However, if households visited a destination more than once a year, a vast majority of the times (99.5% of the times) the same mode was used to travel across all the different trips made to that same destination. This suggests that long-distance leisure trip mode choices depend primarily on the destination choices, and exhibit little variation (or multiple-discreteness) across the different trips made to the same destination. Taking advantage of this finding, we estimated a traditional discrete mode choice model with data on all leisure destinations visited by the households (i.e., 36,263 destinations visited by 22,215 households). This auxiliary mode choice model was used to construct the log-sum variable to be fed into the destination choice MDCEV model as a composite impedance measure that considers the travel times and costs by both air and auto modes.

Chapter 5: Results and Discussion

This section presents and discusses the model estimation results. First the auxiliary mode choice model results are discussed (Section 5.1), and then the main destination choice MDCEV model results are discussed (Section 5.2).

5.1 Auxiliary mode choice model specification

The results of the auxiliary mode choice estimation are provided in Table 13. The binary choice (for choice between air and auto modes) includes an alternative specific constant, household income categories, travel cost and travel time variables (between household residential locations and their visited destinations) by alternative modes, and dummy variables for origin or destination being an MSA. The first, income variable effects indicate, as expected, that higher income households are more likely to travel via the air mode while lower income households are least likely to do so. The next variable is the travel cost variable, computed as the cost of travel for all persons in the travel party. Several specifications were explored on the travel cost variable, including a simple linear form, Box-Cox transformation (Mandel et al., 1997; Gaudry, 2002), logarithmic transformation (Gunn, 2001), and a piece-wise linear specification (Pinjari and Bhat, 2006). The linear specification provided the worst model data fit, while all non-linear specifications improved the model fit and suggested a dampening trend in the sensitivity to costs (i.e., a decrease in the marginal disutility cost as costs increased). This trend is widely noted in the long-distance travel literature (see, for example, Daly 2008). Box-Cox transformation improved the model fit, but provided an unintuitive interpretation when travel cost was interacted with household income category variables. Piece-wise linear specification resulted in sudden discontinuities in the sensitivities from large values to small values (see Daly, 2010 for warning on this same issue). The logarithmic transformation on the cost variable provided the best model fit as well as an intuitive interpretation, while not losing the generality when compared to the Box-Cox transformation (the cost sensitivity vs. cost profiles of both log-cost and Box-Cox transformations were very similar). To account for income-based heterogeneity in households' sensitivities to travel costs, the travel cost variable (in its logarithmic form) was interacted with income category variables. The corresponding coefficients indicate, as expected, that the low income households are most sensitive to travel costs, while high income households are least sensitive. It was difficult to get such intuitive interpretations from the Box-Cox specification.

The next, travel time variable was specified in the linear form because non-linear specifications resulted in interpretation difficulties (e.g., Box-Cox resulted in an unintuitive positive sign) as well as insignificant model improvements. Besides, several long-distance mode choice studies in the past used a linear specification on travel time (e.g., Gunn, 2001).

Dummy variables to indicate whether the origin and destinations are part of a metropolitan statistical area (MSA) were introduced to the utility function for the air mode. The positive coefficients on these variables indicate that the air more is more attractive for those travelers who are departing from (i.e., reside in) an MSA or traveling to destinations in an MSA, when compared to those who reside in or travel to a non-MSA. This is a reasonable result as major airports (with good connectivity and cheaper airfares) are generally closer to metropolitan statistical areas. Finally, the alternative specific constant for the auto mode is positive, reflecting the higher auto mode share in the sample. Overall the model results are reasonable and provide an understanding of the factors influencing mode choice for long-distance leisure travel.

Table 13: Auxiliary mode choice model specification

Explanatory Variables	Ground/Auto		Air	
	Parameter	t-stat	Parameter	t-stat
Low income (< $30k/year) medium income is base	--	--	-0.389	-2.81
High income (>$75k per year) medium income is base	--	--	0.581	8.99
LogCost = Log(Travel Cost in 100s of Dollars)	-2.020	-55.50	-2.020	-55.50
Low income <$30k per year) dummy* LogCost	-0.166	-1.98	-0.166	-1.98
Very high income (>$100k per year) * LogCost	0.180	4.36	0.180	4.36
Travel Time in hours	-0.028	-14.53	-0.028	-14.53
Dummy if origin is an MSA	--	--	0.606	14.400
Dummy if destination is an MSA	--	--	1.400	31.320
Alternative Specific Constant	0.764	8.19	0.000	fixed
Number of Cases	36263			
Log Likelihood at Convergence	-8024.04			
Log Likelihood - Constants Only	-14583.22			
Adjusted Rho Square	0.449			

5.2 Destination choice model specification

The empirical specification of the vacation destination choice and time allocation model is provided in Table 14 for both the basic MDCEV model (as in section 2.1) as well as the MDCEV model that incorporates minimum required time allocations (as in section 2.2). The table is divided into three main parts including the baseline marginal utility function specification, satiation function specification, and model goodness of fit measures, as discussed next.

5.2.1 Baseline marginal utility specification

As discussed earlier, the baseline marginal utility function governs the discrete choices, since it represents the marginal utility derived at zero time investment before any satiation effects begin to occur. A destination alternative with a higher baseline marginal utility is more likely to be visited than that with a lower baseline marginal utility.

Between the two models (i.e., the MDCEV and the MDCEV with minimum required time allocations), there are no significant differences in the baseline marginal utility parameter estimates as well as the corresponding interpretations of the variable effects. Thus, we discuss the variable effects for only one model without any comparisons to the other model.

The first set of variables in the baseline marginal utility specification corresponds to the transportation level of service characteristics. The first, log-sum variable, provides a measure of composite impedance for the modes in the mode choice model. The smaller the log-sum value is (i.e., the higher negative value it takes), greater is the impedance between the origin (household's residential location) and the alternative destination. Thus, a positive and statistically significant coefficient of the log-sum variable, as expected, indicates a lower attractiveness of destinations with higher impedance to travel. The next variable is the highway travel distance between household residential MSA/non-MSA and the destination MSA/non-MSA. As expected, farther away destinations are less likely to be visited. At the same time, as shown by the demographic interactions with the distance variable, demographic heterogeneity exists in households' sensitivity to travel distance. Households with children are more sensitive to distance (i.e., less likely to visit farther away destinations) than households without children, perhaps due to the difficulty of traveling farther distances with children. The distance variable was interacted with the annual income of the household. The hypothesis was that higher income travelers would not be as sensitive (as lower income households would be) to additional travel distances, as they can better afford the additional costs associated with farther travel distances. The positive parameter associated with this interaction variable confirmed the hypothesis. Lastly, householder age group dummy variables were interacted with the distance variable, with the middle age group (25-64 years) as the base category. These householder age-group variables represent the life cycle stage of the household. The corresponding coefficients indicate that both younger (<25 years) and older (>64 years) age groups are likely to travel farther distances than the middle age group households. These results make intuitive sense as both the younger and older age groups may have lower time constraints and hence can potentially visit farther away vacation destinations. The younger age group typically comprises students and young adults with fewer time demands associated with a family and career, while the older age group is typically in retirement and less likely to have the time constraints associated with a full time career. For the middle age group, on the other hand, career and familial responsibilities may impose time constraints that make them less likely to travel farther away for vacation purposes. The next two variables in the level of service characteristics correspond to indicators for the destination to be in the same state (as the household is), and the adjacent state. The coefficients of these variables are positive and significant, indicating a higher propensity of households to visit familiar (and perhaps close by) locations within their state and adjacent states.

The first of the destination characteristics is a size measure (logarithm of the area of the destination MSA or non-MSA) and used as a control to account for the differences in the areas across the destinations. The coefficient of the size variable is positive and smaller than one. This can be explained based on the spatial aggregation of several elemental destination alternatives in the model. For example, several MSAs defined in the model may include multiple destination cities (e.g., the Tampa–St. Petersburg–Clearwater MSA with three different cities) and most non-MSAs defined in the model are an aggregation of different individual destinations. As explained in Daly (1982), a smaller than one coefficient on the size variable indicates a significant presence of unobserved attributes that vary across these elemental destination alternatives (i.e., non-homogeneity across the elemental destination alternatives within a destination).

The next variable, MSA dummy, controls for differences between MSA destinations and non-MSA destinations. The coefficient suggests that MSA destinations tend to be more attractive than non-MSA destinations for long-distance leisure travel purposes, perhaps due to the presence of more opportunities for recreation, entertainment, and other leisure activities in MSAs.

The next variable "density of employment in the leisure and hospitality industry" includes the employment levels in food services, arts, entertainment, recreation, and accommodation sectors. As such, the variable is a surrogate measure for leisure activity opportunities at the destination. A positive and statistically significant coefficient for this variable indicates, as one would expect, that places that offer higher leisure activity opportunities are more attractive as vacation destinations.[17]

The length of coastline was also included as a destination attractor. The coefficient on this variable is positive and statistically significant, indicating, as expected

---

[17] Other employment variables, including a total employment variable and a retail employment variable were also explored in the model. A population density variable was explored too. Several of these variables are highly correlated with leisure and hospitality employment and with each other. Thus, the variables were introduced separately as well as together in different specifications. The signs on the coefficients of these variables reversed and provided unintuitive results when introduced together rather than separately. Such explorations were performed for each combination of variables and by using alternative functional forms such as natural log of the employment variables as well as per area density of employment. After extensive exploration, it was decided that using only the leisure/hospitality employment density variable (with no other employment or populations variables) provided most intuitive interpretation for long distance leisure travel without any substantial impact on the model fit to the data.

that destinations with longer coastlines are more attractive. This is because destinations with longer coastlines offer a variety of leisure activity opportunities such as swimming, fishing, boating, or sightseeing.

The next set of variables in the baseline utility function is associated with the climate at the destination. First of these is the difference in the number of freezing days per year between the destination and the origin. A freezing day is defined as a day in which the temperature drops below 32 degrees Fahrenheit. The negative coefficient on this variable suggests that households are less likely to visit destinations with more freezing days per year than what they experience at their residential end. Colder destinations are less attractive for vacations because freezing temperatures limit many of the activities for which a household may want to travel. Besides, a greater number of freezing days per year result in fewer available days for most vacation activities. In addition to the annual freezing days variable, the mean temperatures for the destination during the months of January and June were included in the model as a way to understand the influence of winter and summer temperatures. Several specifications were explored before arriving at the final specification that provided the best data fit and offered an intuitive explanation.[18] The January temperatures ranged from a maximum of 75 to a minimum of below freezing temperatures. The corresponding variables and coefficients indicate that households prefer to visit destinations that offered the warmest winter temperatures. As the winter temperatures drop below the 65-75 range, the attractiveness of the destinations decreases. Specifically, *ceteris paribus*, destinations with temperatures near or below freezing point are likely to be the least preferred. For

[18] Note here that the temperatures used in the data are daily maximum temperatures averaged over a month. Daily minimum and average temperatures values were also explored in the model, but the maximum daily temperatures data provided a better model fit (albeit slightly better). Other explorations included, specifying an annual average temperature variable (as opposed to separate, winter and summer temperatures), which yielded a poor model fit and coefficients that were difficult to interpret.

summer temperatures, the results indicate that the utility of a destination does not vary monotonously with temperature. Rather, a moderate temperature range might exist that is comfortable for most people (Savageau and Loftus, 1997), and an increase or decrease of temperatures beyond the moderate ranges may reduce the attractiveness of destinations. We explored different temperature ranges and the best fitting model suggested 65-75 degrees Fahrenheit as a comfortable temperature range. Temperatures above or below this range were included as dummy variables of 5 degree increments. Comparing coefficients of the 60-64 degree dummy variable with those of the other variables suggests that destinations with temperatures below the comfort range (65-75) in June have a higher disutility than those destinations with temperatures above the comfort range. Comparison of the coefficients across January and June temperature variables also suggests that the disutility associated with colder (than moderate) climates is higher in magnitude than that of hotter (than moderate) climates.

5.2.2 Satiation (γ_k) function specification

The satiation function coefficients in Table 4 refer to the elements of the θ vector, where the satiation parameter γ_k for vacation type *k* is written as $\exp(\theta' w_k)$. A higher value of the γ_k parameter implies lower satiation for the destination alternative *k* (hence, larger amount of time allocated for that destination). Thus, a positive θ coefficient on a positive valued variable increases the satiation parameter, implying a slower rate of satiation (or higher time allocation).

While there are perceivable differences between the satiation parameter estimates of the two models (i.e., the MDCEV and the MDCEV with minimum required time allocations), there are no significant differences in the interpretations of the variable effects. Thus, we discuss the variable effects for only the latter model.

The coefficient for travel distance has a positive sign and is significant. This suggests that as the distance to a traveled destination increases, and thus the travel time and costs associated with reaching the destination increase, travelers will be more likely to allocate more time to that destination. That is, travelers will likely not make a very long (and costly) trip for a very short stay. Perhaps they take advantage of the time and money spent for the transport to farther away (and more exotic) destinations by staying longer at those destinations. Another possibility is that farther away destinations simply require longer travel times (hence longer time allocated). Travel distance was also interacted with different levels of annual income of the household. The corresponding coefficients indicate that high income households spend smaller portions of time, where as low income households spend larger portions of annual vacation time for farther away destinations. These income differences may be due to the differences in the travel mode choices between different income groups. High income households may travel by air which helps reduce their overall time spent on the vacation trip. Low income households, on the other hand, may travel by slower modes and hence need more time for their vacation trips. Besides, low income households might want to take advantage of the money spent on longer trips by staying longer, while high income households might not feel the same need to stay longer at a destination.

The next set of variables corresponds to household demographics – age of householder and household size. Householder age was introduced in the form of categorical variables with the 25-45 age range as the base category. The coefficients on these age category variables suggest that older (age 46 and above) age groups are likely to allocate relatively more time to a vacation destination than other age groups. The relative magnitude of the coefficients indicate that households belonging to the oldest age group (65 and above) tend to allocate the largest proportion of their time to a destination followed by the older middle age (46-64), the youngest (15-24) age groups, and finally the younger middle age group (25-45). This order makes intuitive sense as it is reflective of the different levels of time constraints faced by households in different life cycle stages (represented by the householder age groups). The oldest (65 and above) householders include those in their retirement years with the least familial and career oriented time constraints and a higher amount of time (and perhaps money) at their disposal. Hence this age group is likely to spend longer vacation times at the destinations they visit. The youngest (15-24) age group is also likely to have lesser time constraints (hence spend more vacation time). The younger middle age (25-45) group householders, on the other hand, are typically at an early state in their professional career and with family related time constraints. Older middle age (46-64) group householders are likely to be well established in their careers and not likely to have young children. So their time constraints may not be as tight as those earlier in their careers and at an early stage of their family life cycle.

The last variable in the satiation function is household size, which is a surrogate measure for the number of travelers (i.e., the travel party size) on vacation trips. The positive coefficient on this variable suggests that a larger household is likely to spend a larger amount of time for a destination than a smaller household. A plausible reason for this result is that larger households (hence larger travel party sizes) tend to travel by slower ground modes than by expensive air modes, hence take longer time for visiting a destination. Another reason is that larger households, typically with children, might prefer to take more time at a destination for a relaxing vacation than making a quick and tiring trip.

In summary, the MDCEV model estimates are reasonable and provide important insights into the impact of the travel level of service attributes, destination characteristics, and household socio-demographic characteristics on households' annual vacation destination choices. These results demonstrate the usefulness of the MDCEV model framework for modeling annual vacation destination choices and time allocation patterns.[19] The model fit measures are reported in the last set of rows. The log-likelihood values of both the models show significant improvement over a naïve model with no explanatory variables. The Rho-squared value for the model is 0.260, an acceptable value for an ambitious model framework that attempts to model all the annual destination choices and the time allocations of households with a large choice set of 210 alternatives. Further, while the proposed variant of the MDCEV model does not offer significantly different interpretations compared to the original MDCEV model, it provides a better fit to the data.

[19] It took about 90 minutes to estimate the parameters of the final model specification presented here (on a 2.6 GHz, 3.25 GB RAM, dual core processor desktop machine, with default starting values for the parameters). The MDCEV model estimation code available at Bhat's website was used as a starting point for this study. His code was modified so that the model estimation input data could be stacked into as many rows as the number of households times the number of destination choice alternatives, as opposed to the usual way of stacking model estimation data into as many rows as the number of households (with one row containing information on all the 210 destination choice alternatives for a household).

Table 14: Destination choice model specification

	MDCEV		MDCEV w/ minimum required consumption	
Baseline Utility Function (Ψ) Specification	**Coeff**	**t-stat**	**Coeff**	**t-stat**
Distance and level of service characteristics				
Log-sum variable from the mode choice model	0.3043	29.12	0.3034	29.03
Highway distance to Destination (100's of miles)	-0.0578	-33.13	-0.0576	-33.10
Highway distance* Presence of Children (0-17)	-0.0196	-10.25	-0.0203	-10.61
Highway distance*High income (> $75k) dummy	0.0186	8.75	0.0185	8.73
Highway distance* Householder age 15 to 24 (25-64 as base)	0.0109	2.63	0.0105	2.56
Highway distance* Householder age 65 or older (25-64 as base)	0.0129	5.35	0.0126	5.24
Dummy if destination in same state as household residence	1.5444	37.09	1.5426	37.02
Dummy if destination in adjacent state to household residence	0.8914	29.52	0.8900	29.46
Destination Characteristics				
Log(Land area of the destination in sq. miles)	0.5453	35.91	0.5415	35.63
Destination is an MSA (Dummy variable)	1.3098	14.85	1.2848	14.55
Leisure Employment Density in 100's of employees/sq. mile	0.0953	45.09	0.0949	44.88
Length of coastline in 1000's of miles	0.0731	10.50	0.0733	10.54
Difference in number of freezing days (destination minus origin)	-0.0092	-20.77	-0.0092	-20.76
Winter (January) temperatures (monthly avg of max daily values)				
55 to 65 degrees Fahrenheit (65-75 degrees as base)	-0.8337	-16.93	-0.8314	-16.88
45 to 55 degrees Fahrenheit (65-75 degrees as base)	-1.5288	-26.71	-1.5349	-26.80
35 to 45 degrees Fahrenheit (65-75 degrees as base)	-2.2623	-31.97	-2.2711	-32.06
Less than 35 degrees Fahrenheit (65-75 degrees as base)	-2.3046	-27.68	-2.3182	-27.83
Summer (June) temperatures (monthly avg of max daily values)				
60 to 65 degrees Fahrenheit (65-75 degrees as base)	-3.4599	-12.89	-3.4246	-12.75
75 to 80 degrees Fahrenheit (65-75 degrees as base)	-0.6770	-17.05	-0.6615	-16.66
80 to 85 degrees Fahrenheit (65-75 degrees as base)	-0.4292	-9.74	-0.4254	-9.65
85 to 90 degrees Fahrenheit (65-75 degrees as base)	-0.7987	-16.02	-0.8025	-16.10
More than 90 degrees Fahrenheit (65-75 degrees as base)	-0.5090	-10.34	-0.5162	-10.49
Satiation Function (γ) Specification				
Highway distance to Destination (100's miles)	0.0910	27.81	0.0991	31.31
Distance*Low income (under $30k) dummy ($30k-$75k is base)	0.0432	8.21	0.0361	7.00
Distance*High income (over $75k) dummy ($30k-$75k is base)	-0.0392	-8.72	-0.0381	-8.66
Householder age 15 to 24 (25 to 45 is base)	0.5102	4.39	0.2555	2.32
Householder age 46 to 64 (25 to 45 is base)	0.4091	9.42	0.2830	7.07
Householder age 65 or older (25 to 45 is base)	0.9966	17.09	0.8883	16.11
Household size	0.2680	28.46	0.1850	20.83

Table 14: (continued)

	MDCEV	MDCEV w/ minimum required consumption
Model Fit Measures		
Log-likelihood at convergence: L $(\hat{\beta},\hat{\theta})$	**-46891.91**	**-46027.30**
Log-likelihood with no variables in the model: L(0)	**-63796.70**	**-62206.33**
Rho-squared = 1- {L $(\hat{\beta},\hat{\theta})$ / L(0)}	**0.265**	**0.260**

5.3 Destination choice model validation

This section provides a validation analysis of the annual destination choice and time allocation MDCEV models discussed earlier. The validation exercise was performed using a sample of 715 households from the 1995 American Travel Survey that were not a part of the 6000 household-sample used for model estimation.

Validation of an empirical MDCEV model requires the application of the MDCEV modeling framework to simulate (or predict) households' annual vacation destination choices and time allocation patterns. In this study, we used a simple and computationally very fast prediction algorithm that Pinjari and Bhat (2010) presented for using the MDCEV model for prediction purposes. For each of the 715 households under consideration, we used 50 sets of random draws from independent type-1 extreme value distributions to simulate the unobserved heterogeneity (i.e., the ε_k terms) in the model.[20] For each household and each set of random draws, conditional upon the total annual vacation time available to the household, the MDCEV model estimates were used to predict the annual vacation destination choices and the time allocation to each predicted destination. The prediction exercise was carried out for both the basic MDCEV model and the proposed variant of the MDCEV model. Subsequently, histograms were plotted

[20] Using the prediction procedure proposed in Pinjari and Bhat (2010), it took less than 1 minute to complete the prediction simulation for all 715 households over all 50 sets of random draws. The Pinjari and Bhat (2010) forecasting procedure was slightly modified to apply the proposed variant of the MDCEV model that accommodates minimum required time allocations. Details are suppressed here to save space, but available from the authors.

to obtain the distributions of the predicted choices over all 715 households and all 50 random draws for both the models. Such predicted distributions were compared to the observed distributions over all the 715 households in the data. Figures 5, 6, and 7 provide both observed and predicted distributions (for both the models) and are discussed next.

Figure 5 provides the distributions of the home-to-destination distances for the destinations observed in the data as well as the destinations predicted by the models. Both the models provide similar distributions that are reasonably consistent with the observed distribution. However, the models seem to slightly under-predict destinations within 1000 miles from the household locations, over-predict destinations in the 1000-3500 mile range from the household locations, indicating a lower sensitivity of the model to level of service variables. One way to improve these results is to jointly estimate the destination choice and mode choice models. The travel time and travel cost sensitivities embedded in the current destination choice MDCEV model (through the log-sum variable) are based on households' mode choice decisions. A joint model may help incorporate the sensitivities (to the level of service variables) that are based on both mode choices and destination choices and thereby improve the distance-based validations.

Figure 6 provides the observed and predicted distributions of the total number of destinations visited by households in the year. Note that the MDCEV framework does not directly model the number of chosen destinations. Nonetheless, both the models provide similar distributions that are consistent with the observed distribution. There are minor differences in that the models slightly under-predict households that visited one destination in a year, and slightly over-predict the households that visited more than 2 destinations. A few (although very small percentage) households were predicted to visit as many as 16 destinations, where as the observed choices indicate a maximum of 7 destinations visited.

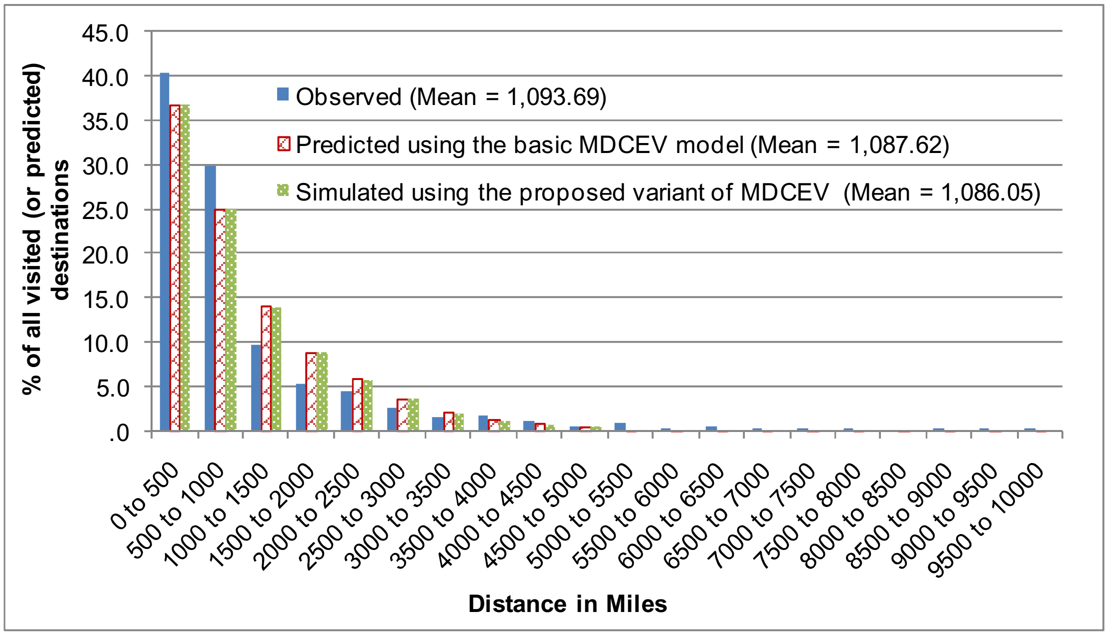

Figure 5: Model validation results based on distances to chosen destinations

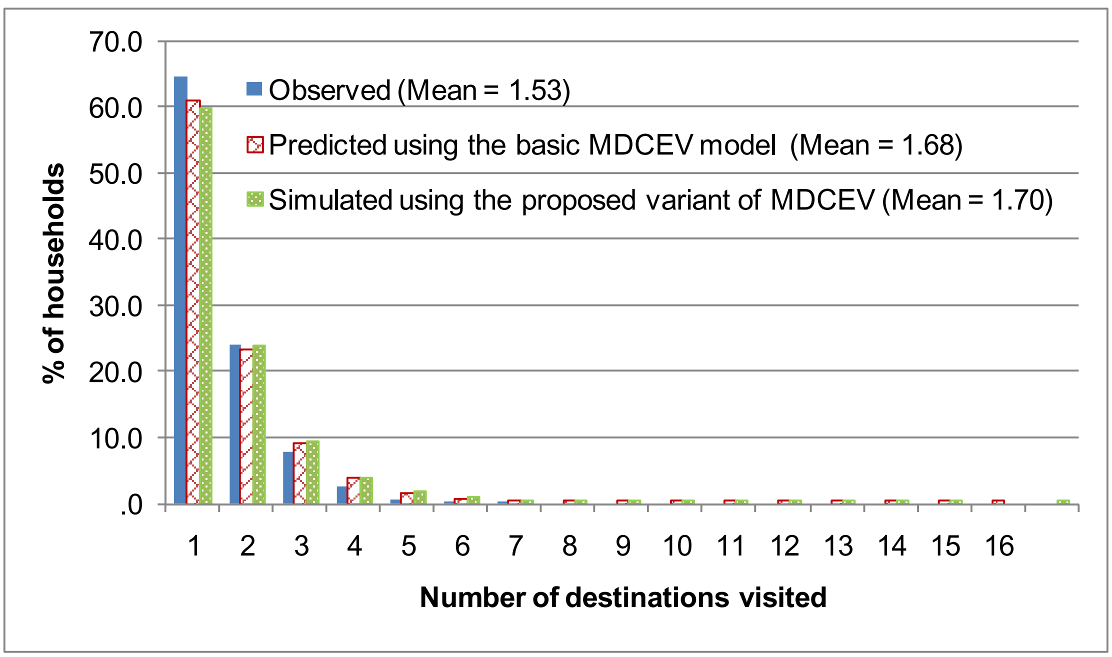

Figure 6: Model validation results based on the number of destinations visited in a year

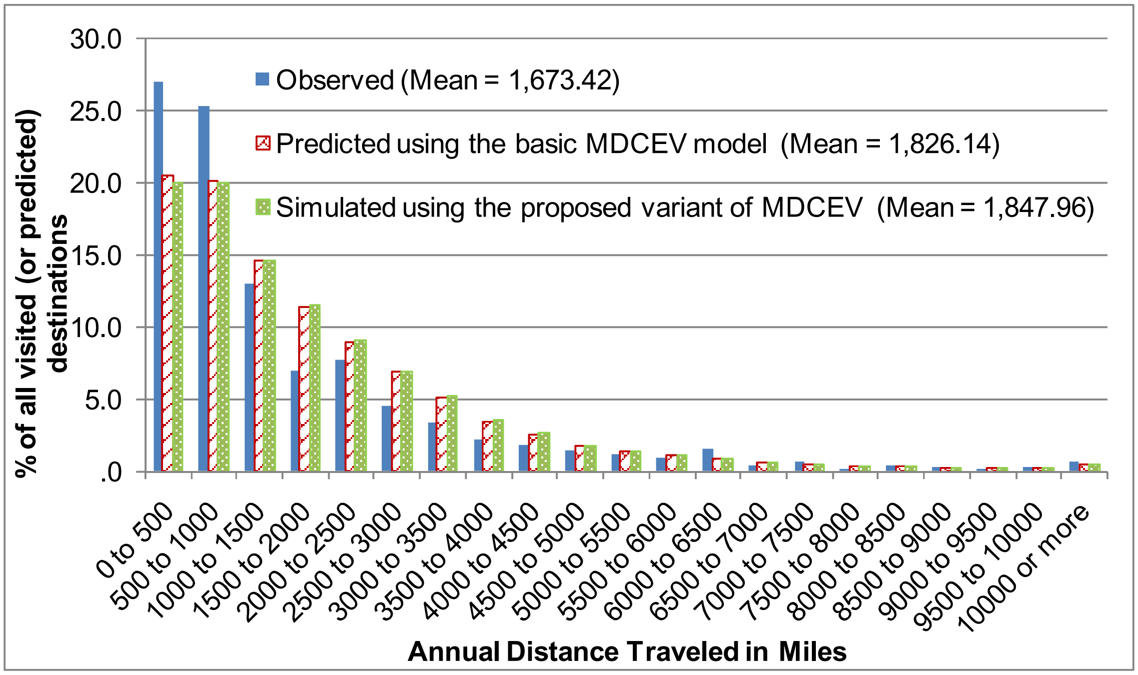

Figure 7: Model validation results based on the total distance to the chosen destinations

Figure 7 provides the observed and predicted distributions of the total distance from home location to all destinations visited in the year. Again, both the models provide similar results, with under-predictions in the shorter distance ranges and over-prediction in the longer distance ranges. This may be due to a combination of lower model sensitivity to level of service variables (as discussed in the context of Figure 5) and the over-prediction of the number of destinations visited (hence longer distances) for a small percentage of households.

We also compared the observed and predicted distributions of the time (no. of days) allocated to chosen destinations (figure not shown). By design, no household in the data is observed to have spent less than 0.5 days for any chosen destination. However, about 10% of the predicted destinations from the basic MDCEV model were allocated less than half a day of time. The proposed variant of the MDCEV model reduces such predictions with less than minimum amount of time allocation to only 2%, although it doesn’t completely preclude very small time allocations.

In summary, the validation results demonstrate the models' ability to provide reasonable predictions, at the least in the aggregate level.[21] The results also provide leads to improve the model specification. The basic MDCEV model and the proposed variant of the MDCEV model provided similar validation results. However, the proposed variant of MDCEV helped in reducing the percentage of choices with smaller than minimum required amount of consumption. Thus, the proposed framework can potentially be useful in situations where it is important to avoid predicting unrealistically small amounts of consumption.

[21] This is not to claim that reproducing aggregate observed distributions (even if in a validation sample) is a sole yard stick for measuring model performance. It is important that the model demonstrate appropriate sensitivity to changes in policy variables and the socio-demographic makeup.

Chapter 6: Conclusions and Future Research

This thesis contributes to the literature on national travel demand modeling by providing an analysis of households' annual destination choices and time allocation patterns for long-distance leisure travel purposes. More specifically, an annual vacation destination choice and time allocation model is formulated to simultaneously predict the different destinations that a household visits in a year, and the time it allocates to each of the visited destinations. The model takes the form of a Multiple Discrete-Continuous Extreme Value (MDCEV) structure. Given the total annual vacation time available for a household, the model assumes that households allocate the annual vacation time to visit one or more destinations in a year in such a way as to maximize the utility derived from their choices. The model framework accommodates variety-seeking in households' vacation destination choices in that households can potentially visit a variety of destinations rather than spending all of their annual vacation time for visiting a single destination. At the same time, the model accommodates corner solutions to recognize that households may not necessarily visit all available destinations. An annual vacation time budget is also considered to recognize that households operate under time budget constraints.

The empirical data for this analysis comes from the 1995 American Travel Survey (ATS) data, with the U.S. divided into 210 alternative destinations. Thus, the study provides an opportunity to estimate, apply, and assess the performance of the MDCEV model for an empirical context with a large number of choice alternatives. The empirical analysis provides important insights into the determinants of long-distance leisure destination choice and time allocation patterns. Select findings are summarized here: (a) Destinations with larger impedance to travel are less attractive in general, but especially so for households with children, low and medium income households, and middle age group (25-64 years) householders. (b) Leisure and hospitality employment, length of coastline, number of annual freezing days (relative to the origin), and winter and summer temperatures are important determinants of travelers' attractiveness to a destination. Specifically, destinations that offer a greater number of leisure activity opportunities, longer coastline, and moderate temperatures (65-75 degree Fahrenheit) are more attractive than other destinations. (c) Low income households tend to spend a longer time for vacations to farther destinations followed by medium income and high income households, in that order. (d) Households with older (>64 years) householders and those with larger number of individuals tend to spend longer time at a vacation destination compared to other households.

On the methodological front, the paper proposes a variant of the MDCEV model that helps reduce the prediction of unrealistically small amounts of time allocation to the chosen alternatives. To do so, the continuously non-linear utility functional form in the MDCEV framework is replaced with a combination of a linear and non-linear form. The proposed variant of the MDCEV model provides a better model fit than the original MDCEV model, and reduces the likelihood of destination choices with unrealistically small amounts of time allocation.

The annual destination choice and time allocation models estimated in this study were validated using a validation sample of 715 households. The validation results demonstrated the models' prediction ability in terms of producing reasonable aggregate-level distributions of the predicted distances traveled and the number of destinations visited in a year.

An appealing feature of the proposed model is its applicability in a national, long-distance leisure travel demand model system. While the proposed destination choice model does not explicitly provide a nationwide origin-destination trip distribution table, the knowledge of the annual destination choices and time allocations predicted by this model can be used for subsequent analysis of the number of trips made (in a year) to each destination and the travel choices for each trip, including mode choice, time (i.e., season) of the year, and length of stay. Thus, the models developed in this study can be incorporated into a larger national travel modeling framework for predicting the national-level, origin-destination flows for vacation travel. This larger national level travel modeling framework would be of particular use to national and regional level tourism boards and national level transportation agencies.

This study paves way to several avenues for further work. First, it will be useful to implement a larger, national-level vacation travel demand system as described in Figures 1 and 2. Additionally, an expanded modeling framework that includes additional travel purposes, beyond leisure travel, can provide further improvements over traditional modeling techniques. Many travelers will likely combine trips (e.g. travel for business, but also incorporate some leisure activities) and so capturing the details of these trips, and their impacts, will provide valuable insight to both transportation planners and regional tourism agencies. Second, the current empirical study can be enhanced in many ways, including: (a) a joint estimation of the mode choice and destination choice models, (b) inclusion of inter-city bus and rail modes in the analysis, and (c) performing policy simulations to assess model sensitivity to important policies. Third, the model does not consider short-distance leisure travel (i.e., leisure travel within the residential neighborhood such as going to a mall, a nearby beach etc.), because the 1995 ATS data does not collect information on short-distance travel. It would be useful to understand the potential substitution patterns between short-distance leisure travel and long-distance leisure travel. Fourth, the current model considers time budget constraints and allocation, but ignores money budgets both due to the unavailability of the data and the lack of methods to do so. This is another important aspect for future research.

References Cited

Aptech Systems. (2008). GAUSS (Version 9.0).

Baik, H., Trani, A., Hinze, N., Ashiabor, S., & Seshadri, A. (2008). Forecasting model for air taxi, commercial airline, and automobile demand in the United States. *Journal of the Transportation Research Record, 2052*, 9-20.

Bansal, H., & Eiselt, H. (2004). Exploratory research of tourist motivations and planning. *Tourism Management, 25*(3), 387-396.

Beser, M., & Algers, S. (1999). The SAMPERS models. In L. Lundqvist & L. G. Mattsson (Eds.), *National Transport Models: Recent Developments and Prospects*. Stockholm: The Swedish transport and communications research board.

Bhat, C. (1995). A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research Part B: Methodological, 29*(6), 471-483.

Bhat, C. (2005). A multiple discrete-continuous extreme value model: formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological, 39*(8), 679-707.

Bhat, C. (2008). The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions. *Transportation Research Part B: Methodological, 42*(3), 274-303.

Bhat, C., & Gossen, R. (2004). A mixed multinomial logit model analysis of weekend recreational episode type choice. *Transportation Research Part B: Methodological, 38*(9), 767-787.

Boeing (Producer). (2011). 757 Commercial Transport History. Retrieved from http://www.boeing.com/history/boeing/757.html

Bureau of Economic Analysis. (1995). Gross domestic product by state. Retrieved January 10, 2011 http://www.bea.gov/regional/gsp/

Bureau of Labor Statistics. (1995). Employment, hours, and earnings - state and metro area. Retrieved December 2010 http://www.bls.gov/data/#employment

Bureau of Labor Statistics. (2010). BLS Spotlight on Statistics Travel.

Bureau of Transportation Statistics. (1995a). 1995 American Travel Survey (U. S. D. o. Transportation, Trans.). Washington, D.C.

Bureau of Transportation Statistics. (1995b). 1995 American Travel Survey Technical Documentation. Washington, D.C.

Bureau of Transportation Statistics. (1995c). Airline Origin and Destination Survey (DB1B). Washington, D.C.

Bureau of Transportation Statistics. (1997). 1995 American Travel Survey Profile, from http://www.bts.gov/publications/1995_american_travel_survey/us_profile/entire.pdf

Bureau of Transportation Statistics. (1998). Transportation statistics annual report 1998 long-distance travel and freight. Washington, DC.

Bureau of Transportation Statistics. (2001). 2001 National Household Travel Survey. http://nhts.ornl.gov/tools.shtml

Bureau of Transportation Statistics. (2010). Annual U.S Domestic Average Itinerary Fare in Current and Constant Dollars. Retrieved February 23, 2011 http://www.bts.gov/programs/economics_and_finance/air_travel_price_index/html/annual.html

Cambridge Systematics. (2007). National travel demand forecasting model phase I final scope.

Castro, C., Martín Armario, E., & Martín Ruiz, D. (2007). The influence of market heterogeneity on the relationship between a destination's image and tourists' future behaviour. *Tourism Management, 28*(1), 175-187.

Castro, C. B., Armario, E. M., & del Río, M. E. S. (2005). Consequences of market orientation for customers and employees. *European Journal of Marketing, 39*(5/6), 646-675.

Chan, F., Lim, C., & McAleer, M. (2005). Modelling multivariate international tourism demand and volatility. *Tourism Management, 26*, 459-471.

Crompton, J. (1979). Motivations for pleasure vacation. *Annals of Tourism Research, 6*(4), 408-424.

Daly, A. (1982). Estimating choice models containing attraction variables. *Transportation Research Part B: Methodological, 16*(1), 5-15.

Daly, A. (2008). *The relationship of cost sensitivity and trip length.* Paper presented at the Educational Travel Community, Noordwijk, Netherlands.

Davies, C. (2005). 2005 Travel and adventure report: A snapshot of boomers' travel and adventure experiences. Washington, DC: AARP.

Deaton, A., & Muellbauer, J. (1981). Functional forms for labor supply and commodity demands with and without quantity restrictions. *Econometrica: Journal of the Econometric Society*, 1521-1532.

Decrop, A., & Snelders, D. (2004). Planning the summer vacation: An adaptable process. *Annals of Tourism Research, 31*(4), 1008-1030.

Energy Information Administration. (1995). Retail gasoline historical prices. Retrieved December 2010 http://www.eia.doe.gov/oil_gas/petroleum/data_publications/wrgp/mogas_history.html

Environmental Protection Agency. (2005). Emission Facts: Calculating Emissions of Greenhouse Gases, from http://www.epa.gov/otaq/climate/420f05003.htm

Epstein, J. M., Parker, J., Cummings, D., & Hammond, R. A. (2008). Coupled contagion dynamics of fear and disease: mathematical and computational explorations. *PLoS One*, 3(12), 3955.

Eugenio-Martin, J. (2003). Modelling determinants of tourism demand as a five stage process: A discrete choice methodological approach. *Tourism and Hospitality Research, 4*(4), 341.

Eymann, A., & Ronning, G. (1997). Microeconometric models of tourists' destination choice. *Regional Science and Urban Economics, 27*(6), 735-761.

Focalyst. (2007). The sky's the limit: travel trends among the baby boom generation and beyond *Leisure Travel Trends*: AARP.

Fosgerau, M. (2001). PETRA - an activity based approach to travel demand analysis. In L. Lundqvist & L. G. Mattsson (Eds.), *National transport models: Recent developments and prospects*. Stockholm: The Swedish transport and communications research board.

Garin-Munoz, T., & Amaral, T. (2000). An econometric model for international tourism flows to Spain. *Applied Economics Letters, 7*(8), 525-529.

Gaudry, M. (2002). Test of nonlinearity, modal captivity and spatial competition within the STEMM multicountry application for passengers. In L. Lundqvist & L. G. Mattsson (Eds.), *National Transport Models - Recent Development and Proposects, Advances in Spatial Science*. Berlin: Springer-Verlag.

González, P., & Moral, P. (1995). An analysis of the international tourism demand in Spain. *International Journal of Forecasting, 11*(2), 233-251.

Greenridge, K. (2001). Forecasting tourism demand: An STM approach. *Annals of Tourism Research, 28*(1), 98-112.

Grush, W. (1998). Usage and Vehicle Miles of Travel (VMT) per Capita. *Highway Information Quarterly, 5*(4).

Gunn, H. (2001). Spatial and temporal transferability of relationships between travel demand, trip cost and travel time. *Transportation Research Part E: Logistics and Transportation Review, 37*(2-3), 163-189.

Hackney, J. (2004). *Discrete choice models for long-distance travel based on the DATELINE survey.* Paper presented at the 4th Swiss Transport Research Conference, Monte Verita/Ascona.

Haliciolgu, F. (2008). *An econometric analysis of aggregate outbound tourism demand of Turkey.* Paper presented at the 6th DeHaan tourism management conference proceedings.

HCG and TOI. (1990). A model system to predict fuel use and emissions from private travel in Norway from 1985 to 2025: Norwegian Ministry of Transport.

Herriges, J. A., & Phaneuf, D. J. (2002). Inducing patterns of correlation and substitution in repeated logit models of recreation demand. *American Journal of Agricultural Economics, 84*(4), 1076-1090.

Hong, S., Kim, J., Jang, H., & Lee, S. (2006). The roles of categorization, affective image and constraints on destination choice: An application of the NMNL model. *Tourism Management, 27*(5), 750-761.

Horowitz, A. (2006). NCHRP Synthesis 358: Statewide Travel Forecasting Models. *Transportation Research Board of the National Academies, Washington, DC*.

Horowitz, A. (2008). *White paper: Statewide travel demand forecasting*. Paper presented at the Conference on meeting federal surface transportation requirements in statewide and metropolitan transportation planning.

Hotel News Resource. (2007). Consumers Aren't Letting Record Gasoline Prices Get In The Way of Their Next Vacation from http://www.hotelnewsresource.com/article27802.html

Hu, P., & Young, J. (2000). *Using the NPTS and the ATS Together*. Paper presented at the Personal travel: The long and short of it, Washington, D.C.

Iso-Ahola, S. (1983). Towards a social psychology of recreational travel. *Leisure Studies, 2*, 45-56.

Jiang, J., Havitz, M. E., & O'Brien, R. M. (2000). Validating the international tourist role scale. *Annals of Tourism Research, 27*(4), 964-981.

Kim, J., Allenby, G. M., & Rossi, P. E. (2002). Modeling consumer demand for variety. *Marketing Science, 21*(3), 229-250.

Koppelman, F., & Sethi, V. (2005). Incorporating variance and covariance heterogeneity in the generalized nested logit model: an application to modeling long distance travel choice behavior. *Transportation Research Part B: Methodological, 39*(9), 825-853.

Kubas, A., Yilmaz, F., Aktas, Y., & Metin, N. (2005). Analysis of visitor decision making system when visiting natural recreation sites by multinomial logit model. *Quality and Quantity, 39*(5), 615-623.

Lam, T., & Hsu, C. (2006). Predicting behavioral intention of choosing a travel destination. *Tourism Management, 27*(4), 589-599.

LaMondia, J., Bhat, C., & Hensher, D. (2008). An annual time use model for domestic vacation travel. *Journal of Choice Modelling, 1*(1), 70.

LaMondia, J., Snell, T., & Bhat, C. (2009). Traveler Behavior and Values Analysis in the Context of Vacation Destination and Travel Mode Choices: A European Union Case Study.

LaMondia, J. J., & Bhat, C. R. (2010). A Study of Visitors' Leisure Travel Behavior in the Northwest Territories of Canada.

Lanzendorf, M. (2002). Mobility styles and travel behavior: Application of a lifestyle approach to leisure travel. *Transportation Research Record: Journal of the Transportation Research Board, 1807*(-1), 163-173.

Lim, C. (1997). Review of international tourism demand models. *Annals of Tourism Research, 24*(4), 835-849.

Lise, W., & Tol, R. S. J. (2002). Impact of climate on tourist demand. *Climatic Change, 55*(4), 429-449.

Little, J., & Boston, F. R. B. o. (1979). *International travel in the US balance of payments*: Research Dept., Federal Reserve Bank of Boston.

Louviere, J., & Timmermans, H. (1990). Stated preference and choice models applied to recreation research: a review. *Leisure Sciences, 12*(1), 9-32.

Lue, C., Crompton, J., & Fesenmaier, D. (1993). Conceptualization of multi-destination pleasure trips. *Annals of Tourism Research, 20*(2), 289-301.

Lundqvist, L., & Mattsson, L. G. (2002). *National transport models: recent developments and prospects*: Springer.

Mallett, W., & McGuckin, N. (2000). Driving to distractions: Recreational trips in private vehicles. *Transportation Research Record: Journal of the Transportation Research Board, 1719*(-1), 267-272.

Mandel, B., Gaudry, M., & Rothengatter, W. (1997). A disaggregate Box-Cox Logit mode choice model of intercity passenger travel in Germany and its implications for high-speed rail demand forecasts. *The Annals of Regional Science, 31*(2), 99-120.

Moeckel, R., & Donnelly, R. (2010). *Nationwide estimate for long distance travel (NELDT)* Paper presented at the Third International Conference on Innovations in Travel Modeling (ITM) of the Transportation Research Board, Tempe, AZ.

Molina, A., & Esteban, A. (2006). Tourism brochures: usefulness and image. *Annals of Tourism Research, 33*(4), 1036-1056.

Money, R., & Crotts, J. (2003). The effect of uncertainty avoidance on information search, planning, and purchases of international travel vacations. *Tourism Management, 24*(2), 191-202.

Morley, C. L. (1992). A microeconomic theory of international tourism demand. *Annals of Tourism Research, 19*, 250-267.

Moutinho, L. (1987). Consumer behavior in tourism. *European Journal of Marketing, 21*(10), 3-44.

National Oceanic and Atmospheric Administration. (2011). Ocean and Coastal Resource Management. Retrieved January 10, 2011, from National Oceanic and Atmospheric Administration http://coastalmanagement.noaa.gov/

National Transportation Library. (2010). Calculate average fares between airports using DB1B market, 2010, from https://ntl.custhelp.com/app/answers/detail/a_id/416/~/calculate-average-fares-between-airports-using-db1b-market

Nicolau, J., & Más, F. (2005). STOCHASTIC MODELING:: A Three-Stage Tourist Choice Process. *Annals of Tourism Research, 32*(1), 49-69.

Outwater, M. L., Tierney, K., Bradley, M., Sall, E., Kuppam, A., & Modugula, V. (2010). California Statewide Model for High-Speed Rail. *Journal of Choice Modelling, 3*(1), 58.

Phaneuf, D. J., & Smith, V. K. (2005). Recreation demand models. In J. R. Vincent (Ed.), *Handbook of enviornmental economics*. North Holland.

Pinjari, A., & Bhat, C. (2010). An efficient forecasting procedure for Kuhn-Tucker consumer demand model systems. *Technical paper, Department of Civil & Environmental Engineering, University of South Florida*.

Pollak, R. A., & Wales, T. J. (1992). *Demand system specification and estimation*. Oxford: Oxford University Press.

Rich, J., Brocker, J., Hanson, C., Korchenewych, A., Nielsen, O., & Vuk, G. (2009). Report on scenario, traffic forecast and analysis of traffic on the TEN-T, taking into consideration the external dimension of the Union-Trans Tools version 2.

Rugg, D. (1973). The choice of journey destination: A theoretical and empirical analysis. *Review of Economics and Statistics, 55*(1), 64-72.

Savageau, D., & Loftus, G. (1997). *Places rated almanac*: Macmillan General Reference.

Schlich, R., Schonfelder, S., Hanson, S., & Axhausen, K. (2004). Structures of leisure travel: Temporal and spatial variability. *Transport Reviews, 24*(2), 219-237.

Seddighi, H., & Theocharous, A. (2002). A model of tourism destination choice. A theoretical and empirical analysis. *Tourism Management, 23*, 475-487.

Simma, A., Schlich, R., & Axhausen, K. (2001). Destination choice modelling of leisure trips: The case of Switzerland. *Arbeitsberichte Verkehrs-und Raumplanung, 99*.

Simma, A., Schlich, R., Axhausen, K. W., & Raumplanung, S. A. V. (2002). Destination choice modelling for different leisure activities. *Arbeitsberichte Verkehrs-und Raumplanung, 99*.

Thakuriah, P., Virmani, D., Yun, S., & Metaxatos, P. (2001). Estimation of the Demand for Inter-City Travel Issues with Using the American Travel Survey. *Transportation Research E-Circular, 255Á269*.

Thornton, P., Shaw, G., & Williams, A. (1997). Tourist group holiday decision-making and behavior: The influence of children. *Tourism Management, 18*(5), 287-297.

Train, K. (1998). Recreation demand models with taste differences over people. *Land Economics, 74*(2), 230-239.

U.S. Census Bureau. (1995). Land area, population, and density for places and county subdivisions: 2000. Retrieved December 2010 http://www.census.gov/population/www/censusdata/2000places.html

U.S. Census Bureau. (2000). Demographic profiles: 100-percent and sample data. http://www.census.gov/census2000/demoprofiles.html

U.S. Census Bureau. (2001). *Statistical abstract of the United States*. Washington, D.C.

U.S. Energy Information Administration. (2011). Weekly Retail Gasoline and Diesel Prices. Retrieved February 23, 2011, from U.S. Energy Information Administration http://www.eia.gov/dnav/pet/pet_pri_gnd_dcus_nus_w.htm

Um, S., & Crompton, J. (1990). Attitude determinants in tourism destination choice. *Annals of Tourism Research, 17*(3), 432-448.

United States Department of Transportation. (2004). 2001 National Household Travel Survey User's Guide. Washington, D.C.

United States Department of Transportation. (2009). 2009 NHTS User Notes.

van Middelkoop, M., Borgers, A., & Timmermans, H. (2004). Merlin: Microsimulation system for predicting leisure activity travel patterns. *Transportation Research Record 1894*, 20-27.

von Haefen, R. H., Phaneuf, D. J., & Parsons, G. R. (2004). Estimation and welfare analysis with large demand systems. *Journal of Business and Economic Statistics, 22*(2), 194-205.

Vortish, P., & Wabmuth, V. (2007). *VALIDATE - A nationwide dynamic travel demand model for Germany*. Paper presented at the Transportation research board planning applications conference, Daytona, FL.

Winwaed Software Technology. (2009). Mile Charter (Version 2.12).

Woodside, A., & Lysonski, S. (1989). A general model of traveler destination choice. *Journal of Travel Research, 27*(4), 8-14.

Woodside, A., & Macdonald, R. (1994). General systems framework of customer choice processes of tourism services. In G. Gasser & K. Weiermair (Eds.), *Spoilt for choice. Decision making processes and preference change of tourists: Intertemporal and intercountry perspectives*. Thaur, Germany.

Yamamoto, T., & Kitamura, R. (1999). An analysis of time allocation to in-home and out-of-home discretionary activities across working days and non-working days. *Transportation, 26*(2), 231-250.

Yao, E., & Morikawa, T. (2003). *A study on intercity travel demand model*. Paper presented at the 10th conference on travel behavior research Lucerne.

Zhang, L. (2010). *Multimodal inter-regional origin-destination demand estimation: A review of methodologies and their applicability to national-level passenger travel analysis in the U.S.* Paper presented at the 2010 World Conference on Transport Research, Lisbon.