

11-17-2015

A Theoretical and Methodological Framework to Analyze Long Distance Pleasure Travel

Vijayaraghavan Sivaraman

University of South Florida, sivaraman@mail.usf.edu

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [Economics Commons](#), [Mathematics Commons](#), and the [Urban Studies and Planning Commons](#)

Scholar Commons Citation

Sivaraman, Vijayaraghavan, "A Theoretical and Methodological Framework to Analyze Long Distance Pleasure Travel" (2015).
Graduate Theses and Dissertations.
<http://scholarcommons.usf.edu/etd/6026>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

A Theoretical and Methodological Framework to Analyze Long-Distance Pleasure Travel

by

Vijayaraghavan Sivaraman

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Civil and Environmental Engineering
College of Engineering
University of South Florida

Co-Major Professor: Abdul R. Pinjari, Ph.D.
Co-Major Professor: Steven E. Polzin, Ph.D.
Xuehao Chu, Ph.D.
Yu Zhang, Ph.D.
Gabriel Picone, Ph.D.

Date of Approval:
November 20, 2015

Keywords: MDCEV, VFR, Multiple Constraint Choice Models, Kuhn Tucker Systems

Copyright © 2015, Vijayaraghavan Sivaraman

DEDICATION

I dedicate this dissertation to Nithya and Anaswara. It would have been impossible to pursue this effort without their support and indomitable spirit.

ACKNOWLEDGMENTS

I would like to sincerely thank my major professors Dr. Abdul R. Pinjari and Dr. Steven E. Polzin for their support and guidance in pursuing this research. Their advice and help extends beyond academics, and has helped me hone myself as a person as well as a professional. I would also like to convey my gratitude to my committee members: Dr. Xuehao Chu, Dr. Yu Zhang, and Dr. Gabriel Picone for providing with their review and recommendations.

In addition and above all, I would like to thank my parents, siblings and in-laws for their continued support in due course of this effort. Finally yet importantly, thank you to all friends and colleagues at the Center for Urban Transportation Research and College of Engineering. It would have been a difficult journey without having such good people around me to reach to in this course of time.

TABLE OF CONTENTS

LIST OF TABLES	iv
LIST OF FIGURES	v
ABSTRACT	vii
CHAPTER 1: INTRODUCTION	1
1.1 Significance and Impact	2
1.2 State of Measurement and Requirements	3
1.2.1 Household Travel Choices	5
1.2.2 Resource and Budget Constraints	6
1.2.3 Spatiotemporal Methodological Framework	7
1.2.3.1 Destination Choices	7
1.2.3.2 Choice of Purpose	8
1.2.3.3 Mode Choice and Trip Frequencies	8
1.3 Research Objectives and Dissertation Structure	10
CHAPTER 2: ISSUES IN MEASURING AND MODELING LONG DISTANCE PLEASURE TRAVEL	13
2.1 Data-Collection Methods	15
2.1.1 Trip Length Definitions and Purposes	15
2.1.2 Recall Period, Time of Collection, and Measure of Pleasure Travel	16
2.2 ATS and NPTS Data Processing	21
2.2.1 Alternative Trip Expansion Method for Data with Shorter Recall Periods	23
2.2.1.1 Step 1—Monthly Trip Weight Generation	23
2.2.1.2 Step 2—Interviewed Population Trip-Rate Estimation	23
2.2.1.3 Step 3—Total Population Trips per Month	24
2.2.1.4 Annual Trip Estimates by Region of Residence Using NPTS	25
2.2.1.5 Annual Trip Estimates by Household Income Using NPTS	26
2.3 Annual Trip Rate Estimation Using NHTS Daily Travel Data	28
2.4 Motivations for Long Distance Travel	30
2.5 Factors Influencing Long Distance Travel	31
2.6 Components of Travel Expenditures	32
2.6.1 Sociodemographic Preferences	33
2.6.2 Regional Patterns	36
2.7 Summary	38

CHAPTER 3: A DISCRETE CONTINUOUS MODEL WITH PERFECT AND IMPERFECT SUBSTITUTES AND PRICE VARIATION	55
3.1 Motivation	55
3.2 Random Utility-Maximization Models	56
3.3 Model Formulation	61
3.4 Utility Form	62
3.5 Interpretation of Parameters	64
3.6 Perfect and Imperfect Substitutes	65
3.7 Karush Kuhn Tucker Conditions of Optimality	66
3.7.1 KKT Conditions for Chosen Destinations	67
3.7.2 KKT Conditions for Unchosen Destinations	69
3.8 Economic Structure	70
3.8.1 Identification Issues	70
3.9 Consumption Probability Expression	75
3.10 Forecasting with Proposed Model	79
3.11 Accommodation of Multiple Budget Constraints	83
3.12 Empirical Application	88
3.13 Data	91
3.14 Model Estimation	93
3.15 Model Results	95
CHAPTER 4: A NATIONAL ANNUAL HOUSEHOLD LONG-DISTANCE PLEASURE-TRAVEL MODEL SYSTEM	105
4.1 Introduction	105
4.1.1 Long-Distance Travel for VFR	106
4.1.2 Substitution Patterns	107
4.1.3 Timeframe of Analysis	108
4.1.4 Current Research	109
4.2 Difference between VFR and Leisure Travel—A Descriptive Analysis	110
4.3 Modeling Framework	114
4.3.1 Stochastic Frontier Model of Household’s Annual Long Distance Pleasure Time Budget	115
4.4 MDCEV-MNL Model for Households’ Annual Choices of Long-Distance Pleasure-Travel Destinations (LDPT) , Travel Purpose, and Time Allocation	118
4.5 Empirical Analysis	122
4.6 Model Estimation Results	123
4.6.1 Stochastic Frontier Model Estimates for Households’ Annual Long Distance Pleasure Time Budgets	123
4.6.2 Annual LDPT Destination Choice, Purpose Choice, and Time Allocation Model	123
4.6.3 MNL Model of Travel Mode and Trip Frequency	126
4.7 Demonstration of Model Application	127
4.8 Summary	128
CHAPTER 5: CONCLUSION AND FUTURE WORK	138
5.1 Data Collection and Measurement	139

5.1.1 Long and Short Period of Data Collection	139
5.1.2 Travel Motivations and Patterns	141
5.2 Methodological Contribution	143
5.3 Empirical Application	145
5.3.1 Empirical Application of RUM Framework for Discrete Continuous Choices with Perfect and Imperfect Substitutes	146
5.3.2 Annual Household Pleasure-Travel-Demand Model System	147
5.4 Future Work	149
 REFERENCES	 151
 APPENDICES	 161
Appendix A: Estimation of Per Night Cost at Destination	162
Appendix B: National Pleasure Travel Demand Model Estimates	166

LIST OF TABLES

Table 1	Travel survey methods comparison	51
Table 2	Expanded annual trip estimate using nationwide personal transportation survey data collected over 2-week periods	52
Table 3	1995 annual person trips and per capita trip rate estimated using the nationwide personal transportation survey and American travel survey	53
Table 4	Annual person trips by census region estimated using the nationwide personal transportation survey and American travel survey	53
Table 5	Annual person (18 and older) trips estimated by quarter and household income	54
Table 6	Descriptive statistics for estimation data	102
Table 7	Goodness of fit for models estimated in the study	103
Table 8	Parameter estimates of the time- and money-constrained model	104
Table 9	Descriptive statistics of long-distance non-visiting friends and relatives leisure and visiting friends and relatives travel characteristics in the 1995 American travel survey data	132
Table 10	Parameter estimates for annual out-of-home time frontier for pleasure	134
Table 11	Multiple discrete–continuous extreme-value multinomial logit model of long-distance pleasure-travel destination, purpose, and time allocation	135
Table 12	Trip and mode choice model specification by purpose	137
Table A1	Regression estimates for lodging costs per day on vacation (CEX data)	165
Table A2	Regression estimates for nonlodging costs per day on vacation (CEX data)	165
Table A3	Ordered logit model for no. of days spent at destination (ATS data)	165
Table B1	Visiting model estimates with and without migration volumes	166
Table B2	Leisure model estimates with and without migration volumes	168

LIST OF FIGURES

Figure 1	Weighted annual person trip length (greatest circle distance miles) distribution	41
Figure 2	Weighted annual person trip length distribution by survey	41
Figure 3	Weighted total person trip share distributed by purpose	41
Figure 4	Weighted annual person trip distribution by specific pleasure purpose	42
Figure 5	Share of total households making no, one, or more trips per quarter	42
Figure 6	Share of total households making no, one, or more trips per semi-annum and annum	42
Figure 7	Interhousehold and intrahousehold variation in reporting trips across time periods	43
Figure 8	Trip-length distribution in greatest circle distance miles of trips made by persons 5 years and over in an annum	43
Figure 9	Expanded NPTS trip length distribution	43
Figure 10	Percent difference in NPTS vs. ATS trip estimates by quarter and region	44
Figure 11	Sample (persons) interviewed by region and survey vs. population	44
Figure 12	Per capita trip rate by household income of person	44
Figure 13	2009 Nationwide Personal Transportation Survey—Daily long distance travel distribution	45
Figure 14	American Travel Survey/ Nationwide Personal Transportation Survey- Long distance trip length distribution (all person trips over 100 miles)	47
Figure 15	Long distance trip length distribution (person trips over 100 miles)	47
Figure 16	Average consumer expenditure on annual out-of-town trips	47
Figure 17	Average annual expenditure on vacation by consumer unit	48

Figure 18	Average annual Consumer Unit expenditure on out-of-town travel by expenditure category	48
Figure 19	Average trip length on and off vacation by pleasure purpose	48
Figure 20	American Travel Survey person pleasure-trip distribution by age	49
Figure 21	Average CU expenditure on out-of-town trips by Consumer Unit respondent age group	49
Figure 22	Average Consumer Unit expenditure on out-of-town trips by annual income	49
Figure 23	Average Consumer Unit transportation expenditure on out-of-town trips by annual income	50
Figure 24	American Travel Survey average trip length by purpose and region	50
Figure 25	Average annual Consumer Unit expenditure on out-of-town trips by region	50
Figure 26	Average Annual CU Expenditure on out-of-town trips by region	51
Figure 27	Total utility as a function of inside good consumption for different values of \mathbb{E}_0 , w_0 , r , and ... (keeping all else same)	101
Figure 28	Structure of the household-level annual long-distance pleasure-travel-demand model system	130
Figure 29	Prediction exercise results (annual trips) using the proposed model system	131
Figure 30	Prediction exercise results (annual time allocation) using the proposed model system	131

ABSTRACT

The United States (US) witnessed remarkable growth in annual long distance travel over the past few decades. Over half of the long distance travel in the US is made for pleasure, including visiting friends and relatives (VFR) and leisure activities. This trend could continue with increased use of information and communication technologies for socialization, and enhanced mobility being achieved using fuel-efficient (electric/hybrid) and technology enhanced vehicles. Despite these developments, and recent interest to implement alternate mass transit options to serve this market, not much exists on the measurement, analysis and modeling of long distance pleasure travel in the U.S.

Statewide and national models are used to estimate long distance travel, but these are predominantly trip-based models, making it difficult to understand long distance trips as collection of household-level travel behavior. This form of travel behavior has been studied a lot in tourism, but in a piecemeal manner, such as to (from) a specific destination. Further, most of these studies are confined to analyzing leisure market, with VFR market gaining recognition only recently. In essence, annual household long distance pleasure travel behavior needs to be studied in a comprehensive manner rather than as isolated trips. This is because, most of these household travel decisions are undertaken considering their annual time and monetary budget, and their perceived cost to travel to one (or more) destination for given pleasure purpose on one (or more) occasion using a given mode of travel. Thus, the main objective of this dissertation is to develop a comprehensive behavioral model framework to analyze the above-discussed annual household long distance pleasure travel choices.

To start the above effort, it is first required to collect detailed annual household travel data, last collected over two decades ago (e.g.: ATS, 1995). No such recent effort has been pursued due to the significant labor and economic resource required to undertake it. There exist recent surveys (NHTS, 2001), but collected over a shorter (four week) period, and require significant processing even to arrive at aggregate annual travel estimates. Second, besides surveys, there is a need for additional data to estimate households' annual pleasure travel budget, and their cost to travel and stay at each of their potential destination choices, which are not readily available.

Thus, as the first goal, this dissertation analyzes long distance travel reported across historical surveys (NPTS; ATS; NHTS), to understand the differences in their definition, enumeration of purpose and collection methods. The intent here is twofold, first to conceive a method to estimate annual travel from surveys with shorter collection period. Further, the second intent is to gather travel patterns from these historical datasets such that it informs the second goal of this dissertation, i.e. development of a behavioral framework to analyze annual household pleasure travel. To this effect, this research also analyzes pleasure expenditures using Consumer Expenditure Survey (CEX, BLS) data. Interestingly, the analysis reveals CEX pleasure travel expenditure pattern to be similar to the travel pattern reported for the same market segments in travel survey (ATS).

Importantly, the above analysis informs the development of behavioral models, pursued as two distinct tasks to achieve the second goal. As the first task, a novel econometric model and forecasting procedure is developed to analyze a household's annual long distance leisure travel decisions. Specifically, a households' time spent across one (or more) destination and travel mode to such destination for leisure is modeled subject to time and money budget constraints. In

this methodological framework, the destination choice is modeled as a continuous variable (time at destination) using Multiple-Discrete Continuous Extreme Value model (MDCEV). While, travel mode choice to these destination(s) are modeled as a discrete choice, through a nested Multinomial Logit Model (MNL), with price variation introduced across the above choice of destination(s) and travel modes (air/ground). This required estimating annual monetary budgets, travel cost and per night lodging cost for each sample household, with each of them having 210 potential destinations and 2 travel mode choices respectively.

The second task, involved the development of a broader national model system to analyze households' annual pleasure travel decisions such as: choice (duration) at destination(s), travel purpose (VFR or leisure), mode (airplane or auto) choice and trip frequencies to these destination(s) using the same dataset. It was modeled in two stages, with the first stage estimating households' annual pleasure time budget using a stochastic frontier model. This budget was then used as constraint to analyze households' annual choice of destination and purpose using a nested MDCEV-MNL model in the second stage. A log sum variable from a nested joint multinomial logit model of trip frequency and mode choice for each purpose (VFR or leisure) is also introduced as input at this stage. This model was then validated using a prediction procedure, and further applied to test a policy scenario (increase in travel cost). The above national pleasure travel demand model could be further enhanced by including monetary constraints and price variation as in the first task. Overall, the model system proposed in this dissertation forms the foundation for a national comprehensive long distance travel model. This could be achieved through inclusion of other prominent travel purpose such as business and commuting to the national travel demand model presented in this research.

CHAPTER 1: INTRODUCTION

The United States witnessed a significant rise in long distance travel over the past few decades exceeding 2.5 billion trips in 2001 (NPTS Brief, 2006). The above reported growth could be attributed to economic development during this time period complemented with lower transportation costs (fuel costs and auto ownership). This facilitated households to travel farther and more often, not just for work and maintenance, but to pursue pleasure activities as well. In this research, long distance travel for pleasure is defined as trips made to visit friends and relatives (VFR) or participate in leisure activities.

Long distance travel for pleasure, as defined above, accounted for more than half of all long distance trips per annum, and is observed to sustain this share over time as well (American Travel Survey [ATS], 1995; National Household Travel Survey [NHTS], 2001; Nationwide Personal Transportation Survey [NPTS], 1983, 1987, 1990). This trend could continue in the future with people migrating to seek better opportunities (or retirement) resulting in new social ties across geographies.

The above-discussed growth in long distance travel, more so for pleasure, could pick up with use of information and communication technologies for socialization, and efficiencies in travel being achieved with introduction of hybrid and electric (and automated) automobiles. However, there are numerous other factors, which could dampen the demand for pleasure travel, such as adverse economic conditions, or households combining travel purposes together (i.e. pursuing leisure on visiting friends or relatives or business). Besides the above, shifts in socioeconomic state of household (e.g., unemployment, retirement, childbirth) could also

influence household decisions and their consequent demand for long distance pleasure travel as discussed in the following section.

1.1 Significance and Impact

According to the Bureau of Labor Statistics (BLS, 2010), U.S. households, on average, spent about \$1400 per annum on their annual vacation, i.e. for above reported pleasure trips. Among them, a major share (about 44%) was spent on travel (transportation) to the vacation destination, with rest of it being incurred at destination for lodging, food and entertainment. As is apparent from above, long distance pleasure travel is bound to have significant economic impact not just on tourism but in the transportation sector as well.

The above reported expenditures, could be distinct across households, and could change for a household over time as well. It could alter with rise in cost for travel (fuel), lodging, relative cost at destination or change in the household's socioeconomic condition (employment, birth of child etc.). At an aggregate level, such household-level changes could affect the overall demand at (or between) destinations. In some scenarios, it could influence travel demand to all destinations(s), such as during a recession. This could have significant implications on the demand for transportation to various destinations along with the demand for tourist services at the destinations (lodging, car rental services etc.).

The above influences on aggregate travel demand could be better understood through analysis of disaggregate, household-level long-distance pleasure travel behavior. Specifically, it would require the analysis of household-level annual pleasure travel choices such as the destinations to travel to, the motivation for travel to these destinations (e.g., VFR or leisure), travel mode to these destinations (e.g.: air or auto), and the number of trips made to the destinations. This would not only assist in arriving at more precise annual travel estimates, but

could also allow to examine the implications of relevant policies for transportation and tourism due to its disaggregate nature.

For example, it could be used to assess the influence of travel demand management strategies at seasonal attractions (e.g., Key West, Pigeon Forge, and Yellowstone), which often face severe bottlenecks along the main access/egress arterial roads to these destinations. It could save travelers valuable time invested in such travel, and enhance the attractiveness of such destination, with travel to access them becoming more pleasant. Further, a households' decision to undertake a pleasure trip to a destination is not just influenced by their socioeconomic state, but also the cost to pursue them. To this effect, a household with a diverse set of destination(s) might choose the most attractive destination in terms of opportunities as well as the expected cost. To this effect, this model could inform Destination Marketing Organizations (DMO) about the preferences of its potential consumers and their ability to afford such choices, which could support them in devising promotional strategies to attract them.

1.2 State of Measurement and Requirements

Understanding the above discussed household-level pleasure travel choices requires collecting such travel information from households over longer periods (e.g.: year), a resource intensive effort. Besides data collection, development of such a comprehensive behavioral model requires additional household information such as their annual travel budget, their perceived cost to travel and expenditures at destination. This data, as important as it is, is typically not collected in travel surveys, and are not readily available. This section further discusses the household choices that need to be modeled within this comprehensive framework to understand annual household long distance pleasure travel behavior, and the supporting data required to model them, including travel expenditures.

Long distance travel, often defined based on a distance threshold (50, 75 or 100 mile) in the U.S., could be made to neighboring town/cities to distant out of state or international destinations. A major share (62%) of these trips is made to destinations within their state of origin (Bureau of Transportation Statistics, 2001), and are likely to be accounted for in the statewide models using advanced estimation methods such as household/person level discrete choice models (Outwater et al., 2010). However from a household perspective, not all of their annual long distance trips tend to be accounted for in the above models. This is because, long distance trips, particularly those for pleasure are made far and few in between months, and most of the statewide models analyze long distance trips as daily travel.

Further, the abovementioned trips tend to be interrelated, in that the choices made on such trips (destination, duration, travel mode etc.) influence their travel decisions on other trips, which could be made to places within or outside the state of residence. The exclusive treatment of within state travel, and above all its examination as daily trips, restricts the analysis of annual household long distance travel behavior. In particular pleasure travel, which makes up for more than 50% of all long distance trips, could be made to different travel extents (within and out of state). This would depend on the specific motivations of pleasure travel, and location of opportunities to pursue them. There are transportation models (e.g., Ashiabor, Baik, & Trani, 2007; Baik et al., 2008; Moeckel & Donnelly, 2010) that estimate national level travel demand within the U.S. However, as mentioned earlier, most of these models adopt the traditional four-step approach at an aggregate trip level, making it difficult to represent annual household travel behavior.

A households' long distance travel needs to be studied over a longer period and across the U.S because household pleasure travel decisions tends to be diverse in their choice of

destination, duration at destination, trip frequencies and mode of travel. These choices often depend on the characteristics of the household, purpose (motivation) to undertake such travel, and location providing facilities to pursue such purpose and ability to access them. Broadly speaking, the two major motivating factors (purpose) within pleasure travel tend to be VFR and leisure. Of the above, the choice of destination for the former (VFR) purpose relies on social ties, while leisure travel to a destination are made based on the households' leisure motivation and attractions available at such destination to pursue them.

Further, a household with intent to visit might as well participate in leisure with their friends/relatives at destination, without requiring additional trip to satiate their leisure needs. In addition, a household could make multiple visiting trips to a destination, while make one or fewer trips to a destination for leisure. This chapter further assesses the needs and objectives of this research from the perspective of developing such a national level behavioral model for long distance pleasure travel that allows analyses of the above choices within one framework.

1.2.1 Household Travel Choices

The motivations to undertake long distance travel for pleasure could be broadly categorized as either to visit friends and relatives (VFR) or to pursue leisure. Here, leisure encompasses activities such as rest/relaxation, sightseeing, outdoor recreation or entertainment. Households could travel to different geographic extents to pursue each of these activities based on the spatial distribution of the opportunities (locations) to pursue them. For example, the opportunities for visiting, entertainment or outdoor recreation might more often involve travel to neighboring out of town destination; whereas, trips to pursue activities such as sightseeing and rest/relaxation might require distant travel. Thus, Moscardo et al. (2000) proposed a VFR tripology to segment such trips based on geographic (travel) extent.

From an annual perspective, a households' annual pleasure travel on being categorized based on travel extent would exhibit distinct trip frequencies, mode choice and time spent (duration) at such destinations (s). For example, a household on making an occasional trip to a distant destination(s) would want to make the most out of such trip, and thus might spend more time pursuing leisure activities in and around such destination(s). While the same household might spend less time on a trip to a neighboring town to visit friends or pursue other leisure activities (watch a game at a stadium or go hiking). However, they might make more such trips over a year compared to their travel to a distant destination. These decisions, as would be expected, are influenced by resource (cost/time) required to access such destination(s) and participate in activities at these destinations.

1.2.2 Resource and Budget Constraints

Importantly, the above discussed household choices that ultimately shape a trip to a particular destination, and makeup a households' total annual trips are made under several constraints, main among them is household's time and monetary budget constraint. Brons, Pels, Nijkamp, and Rietveld (2001) state that households' aim to maximize their utility (satisfaction) – derived from the cost of travel and from consumption of holiday experiences subject to these constraints. Each household has to adhere to these constraints in deciding their travel choices such as: purpose, destination, duration at destination, and trip frequencies to these destination(s).

The aforementioned choices, subject to income and time budget in tourism is analyzed through three major groups of variables: household sociodemographic, destination, and travel mode characteristics. These factors are referred to as push and pull factors (Armario, 2008; Dann, 1977; Uysal & Hagan, 1993; Yoon & Uysal, 2005) in tourism marketing literature. Push factors include household sociodemographic characteristics and their travel motivations. The pull

factors include the attractiveness of the destination and ability to access these destinations. The destination pull factors could also include factors such as: attractions, climate, culture, architecture, transportation, entertainment or cost (Kozak, 2002). Heung, Hailin, and Chu (2001) provide a comprehensive review of these factors as motivations for vacation travel.

1.2.3 Spatiotemporal Methodological Framework

The household travel choices and constraints that influence these choices, discussed in previous section, suggest the decision made across trips to be inter-related. Hwang, Gretzel, and Fesenmaier et al. (2006), state that people (households) tend to visit more than one destination for several reasons, and at the same time, also take into consideration the constraints and opportunities associated with visiting these multiple destinations. Importantly, they suggest that households cannot satiate all their pleasure travel motivations from traveling to one destination. Thus, the decision made on a particular trip is bound to influence a household's choice of destination, duration at destination, trip frequencies and mode choice on another trip. This makes it imperative to collect and examine these travel choices over longer period rather than examining a households' trip to a destination as an isolated event.

1.2.3.1 Destination Choices

In essence, households' could travel to one or more destination(s) to satiate their annual pleasure demand arising from diverse pleasure motives such as visiting, entertainment, outdoor recreation etc. This is intuitive, and is more representative of annual household pleasure travel behavior compared to considering household to decide their pleasure travel to a particular destination from set of choices, as an isolated decision making process. Thus, examination of a households' annual choice of one or more destination(s) requires it to be treated as an imperfect choice substitute rather than its traditional assessment as a perfect substitute.

1.2.3.2 Choice of Purpose

In contrast, a household pleasure travel purpose to a destination tends to be a perfect choice. Primarily, the decision to travel to a particular destination from several destinations arises from the primary motivation (VFR or leisure), and the opportunities available to pursue this motivation at the destination. Thus, this research considers, VFR and leisure travel to a particular destination as exclusive choices. A household traveling to a destination(s) for VFR might not travel to the same destination(s) for leisure and vice versa. This is because, the households might accomplish either of the purposes as secondary purposes on given trip with either of them as the prime motive. This, in turn, makes these choices to be perfect substitute, i.e. a household might travel to a destination for either of the above purposes not both.

1.2.3.3 Mode Choice and Trip Frequencies

The above choice of purpose along with the spatial distribution of opportunities (destinations) to pursue them would influence the travel mode choice, and trip frequencies as well. Leisure motivation(s) for travel such as rest/relaxation and sightseeing, might lead household to travel to distant out of state destination choices resulting in fewer such trips, but with longer time being spent at such destinations. Alternatively, leisure motivations such as outdoor recreation or entertainment might lead to travel to proximate in-state destinations, often involving more frequent trips, but each of short duration.

While within VFR, households' could either make frequent trips (short durations) to visit friends out of town, and at the same time make an occasional trip to distant out of state destination to visit family but for longer duration. This heterogeneity within VFR travel led tourism researchers (Moscardo et al., 2001) to recommend further categorization of VFR travel into visiting friends and visit relatives respectively.

The above discussed spatial distribution of opportunities to pursue motivations could also lead to distinct modal preferences as identified by McGuckin (2013). Hence, researchers consider having an independent model for travel mode choice as neglecting the package nature of vacation travel (Decrop & Snelders, 2004; Eugenio-Martin, 2003; Hackney, 2004; LaMondia, Snell, & Bhat, 2009; Nicolau & Mas, 2005). Furthermore, besides mode choice, it might influence annual trip frequencies to such destination as well, due to cost and time associated with making such trips across different travel extents.

As is apparent, a comprehensive assessment of long distance travel for pleasure could not be made considering shorter data collection period or smaller geographies (spatial scales). This is because of the fact that the choice of destination (imperfect substitutes) for pleasure travel could span from out of town destination(s) to distant out of state locations. In addition, the time of occurrence of trips to these destinations might as well be far and few in between months.

This research, thus pursues the development of an annual household long distance pleasure travel behavior model to comprehensively analyze the above discussed travel choices using ATS (1995) survey. The ATS (1995) survey, although two decades old, is considered because no recent data collection effort of this kind was pursued due to the significant labor and economic resource required to undertake it. Before undertaking the model development effort, historical surveys and supplemental data sources were analyzed to inform the development of this model and construction of relevant input variables. In the due course, this research also constructed a method to expand data collected over shorter period to arrive at aggregate annual estimates of long distance travel. The following section outlines the objectives and goals pursued in this dissertation to accomplish the above effort.

1.3 Research Objectives and Dissertation Structure

The overall objective of this dissertation is twofold: first to develop theoretical knowledge on long distance pleasure travel, and second to conceive a model structure to formulate a national level pleasure travel demand model. These two objectives are accomplished through chapter 2 through chapter 5 in this dissertation.

Chapter 2 evaluates the nature of long distance travel from the perspective of data collection, examining the influence of survey design and collection methods. The two main goals pursued as part of this chapter are as follows. The first goal is to understand the influence of the data collection period on reporting of long distance trips. Specifically, the influence of shorter and longer recall periods on reporting of long distance trips is examined to evaluate the potential to use the former, with latter being highly resource intensive. Subsequently, an alternate approach to expand data collected over shorter recall period to arrive at annual estimates is proposed. The second goal pursued in this chapter is to examine differences in long distance travel across socio-demographic groups, and its pattern across time and geographies for different pleasure travel purposes. These aspects are explored from the perspective of informing the development of a behavioral model of a household's annual choice of destination(s), travel mode, trip frequencies, and duration at destination subject to time and monetary budget constraints.

It becomes apparent from chapter 2 that annual long distance pleasure travel behavior could be distinct across households., i.e. in terms of their choice of destination (imperfect substitutes), and duration at such destination and their mode of travel (perfect substitutes). These choices predominately depend on a households' ability to spend their time and money budget to

travel to destination (using air or auto as travel mode) that offers the desired pleasure opportunities.

Thus, Chapter 3 presents the development of model formulation to analyze such imperfect and perfect choice substitutes accounting for the price variation across these choice sets. It commences with the formulation of a single time budget constrained model leading into the formulation for multiple constraints to analyze pleasure travel choices accounting for time and monetary budget constraints of a household. This chapter also presents a forecasting procedure for single budget constraint model, with potential to be extended to accommodate multiple budget constraints. The proposed forecasting procedure for single budget constraint is subsequently implemented to validate a model and analyze different scenarios.

Specifically, the model is estimated to analyze a household's annual leisure destinations (imperfect substitute) and mode of travel (perfect substitute) to these destinations accounting for the price variation across mode and destination. This is accomplished through first estimating travel cost across two (air and ground) mode choices to 210 potential destinations, and per night cost at each of these destinations within the U.S. for each sample household. This model is estimated accounting for households' socio-demographic, destination and modal characteristics recognized in chapter 2 to be the key factors in assessing long distance leisure travel decisions.

Subsequently, chapter 4, further broadens the model presented in chapter 3 in terms of the choices being analyzed. It presents a comprehensive national pleasure travel demand model system estimating a household's leisure and VFR travel decisions along with a forecasting procedure to analyze scenarios using the estimated models. The model system in this chapter analyzes each household's annual choice of destination(s) for specific pleasure purpose (visiting or leisure) and their choice of mode and trip frequencies to such destinations. Prior to modeling

the above household travel choices, the model system first estimates each household's annual pleasure time budget, which is used as a input budget constraint in estimating the abovementioned annual household pleasure choices. The input to this model system is the same ATS (1995) data with each household in the sample considered to have a choice of 210 destinations, as the leisure only model estimated in Chapter 3. However, it is estimated accounting for just time budget, and does not include price variation and money budget. The model system is subsequently validated, and used in a prediction exercise to understand the influence of rise in fuel cost using a forecasting procedure developed as part of this research.

Chapter 5 concludes the dissertation with discussion on the three main contributions from this research. First section in this chapter summarizes the theoretical knowledge gained through the assessment of historical long distance travel surveys, and proposes an alternate approach to estimate annual long distance travel using data collected with shorter recall periods. Second section discusses the methodological development pursued as part of this dissertation in formulating a novel approach to analyze perfect and imperfect choices with price variation for single as well multiple budget constraints. The third section presents the outcomes from the development of a comprehensive national long distance pleasure travel demand model, and its limitations. The chapter concludes this dissertation with a brief discussion on areas for future research to improve the methodological and empirical specification of this model.

CHAPTER 2: ISSUES IN MEASURING AND MODELING LONG DISTANCE PLEASURE TRAVEL

Long distance passenger travel in the U.S. exceeded 2.5 billion person trips in 2001 (NPTS Brief, 2006). Despite its remarkable growth, it first started receiving greater attention from the transportation sector at the end of last decade coincident with the onset of the economic recession and the growing interest in pursuing alternate intercity public transportation options. In contrast, it has had sustained interest from the tourism and hospitality stakeholders because of its economic relevance to this sector, where more than ½ of all long distance trips per annum are made for leisure. Long distance travel is bound to have significant economic influence not just on tourism, but on the transportation sector as well.

However, there has been no significant national data on long-distance travel collected in the recent past to understand its current impact since NHTS (2001). Adequate national data is essential to get a better estimate of the more recent travel demand, but a main deterrent to pursue such an effort has been the inherent need to collect long-distance travel over a longer period (year) to get a comprehensive picture. This makes it a resource intensive task to undertake for both respondents as well as surveyors. There have been numerous efforts in the past considering longer (a year) and shorter collection periods (2-4 weeks), in essence implementing different strategies trading off costs and period of data collection.

The shorter period, although economical, has issues of its own, prominent among them being inadequate sample size, coverage bias and under reporting of trips (Giesbrecht and Bose, 2005). Thus, this chapter first examines the potential issues associated with data collected over

different recall periods - short (NPTS, 1995 / NHTS 2001) and long period (ATS, 1995), to arrive at annual estimates. Subsequently, this research also investigates the influence of other aspects of data collection such as trip length definition, purpose enumeration, and period of collection for reporting such trips. Following the above investigation, this research proposes an alternate weighting (expansion) methodology to arrive at annual estimates using single wave of data collected with a shorter recall period.

A significant share of trips (travel) that are made for leisure can be expected to be undertaken by households following some amount of advanced planning taking into account their budget constraints (Brons et al., 2001). The choices made on such trips could be distinct across households based on their budget, and could differ across regions, based on the available leisure opportunities and the cost to access them. Given the consumption oriented nature of such travel, the levels of leisure travel could change over time based on the overall economic condition of the nation and (or) changes in economic conditions for a given household(s).

However, it has been infeasible to collect data to assess such changes over time, due to the significant resource required to collect such data even for a year. Thus, this research explores other supplemental data sources that could provide information on such trends. The intent of this effort is twofold, first to gain insights on leisure travel trend using this dataset. Second intent is to explore the potential to synthesize this data with travel surveys to construct a household level model that analyzes annual leisure travel within the U.S. This is pursued because, there exists no national model or behavioral framework to analyze long distance leisure travel behavior, with most of the previous studies analyzing long distance travel adopting traditional 4 step urban model structure (Horowitz, 2008; Miller, 2003).

Thus, this chapter through its analysis of the historical data collection methods, and supplemental data sources such as Consumer Expenditure Survey (CEX) conceives a theoretical framework to analyze the leisure component of long distance travel. This theoretical framework is further applied to develop a disaggregate household level national model of annual long distance leisure travel in the United States, which could be extended to include other prominent purposes such as business travel etc.

2.1 Data-Collection Methods

2.1.1 Trip Length Definitions and Purposes

The aforementioned efforts in this research commences with the analysis of historical national long distance travel data collected in the U.S. Prior to examining these datasets (surveys), the elements defining long distance travel across surveys are considered to enable consistent comparison across them. Long distance travel survey methods have evolved significantly over the past few decades; with the definition itself modified over time from an overnight trip (NPTS, 1977) to a trip length based definition (NPTS: 1987, 1990, and 1995; ATS 1995; NHTS 2001). The NPTS defined trips to places 75 or more miles apart as a long distance trip, while ATS (1995) and NHTS (2001) report such trips to be made to places over 100 and 50 miles respectively (Table 1).

Due to the above differences, this research only analyzed long distance trips made to destinations 100 or more miles apart in straight-line distance (SLD) to make consistent comparisons. A comparison of these trips (Figure 1), as mentioned in the introduction, shows remarkable growth in long distance trips from 1990 to 2001, with NHTS (2001) reporting over 2 billion trips in 2001. A major share of the observed growth in trips could be natural, but a number of other factors could also have led to this growth. A few of the prominent factors that

could have influenced the trip estimates include the number of enumerated purposes; recall period, time of recall and collection timeframe (1 or more waves). A preliminary analysis of trips over 100 miles in SLD (Figure 2) reveals NHTS reports almost twice the number of trips to short haul destinations/distance (100 – 200 miles) compared to other surveys. A non-trivial share of trips causing this spike in short haul (100 to 200 miles) travel could be attributed to the enumeration of additional purposes in NHTS. Specifically, NHTS reported commute as one of the long distance trip purposes. Commuting makes up a significant share of short haul trips and is excluded as a long-distance trip purpose from ATS by design. In contrast, NPTS (1990, 1995), which did include commute as a to/from trip purpose reported far less trips compared to NHTS (2001) in the 100-200 miles category. This could be partially due to the limited number of long-distance trip purposes enumerated in NPTS.

2.1.2 Recall Period, Time of Collection, and Measure of Pleasure Travel

Besides enumeration of purposes, ATS asked its respondents to recall their travel over the past three(3) month period. This could have resulted in fewer short haul trips to be reported compared to NHTS in this trip length segment. It is because, respondents with a longer recall period might not in effect be able to recall all their short haul trips. This recall bias could be minimized by undertaking surveys with shorter recall periods. However a draw back with shorter recall period, is that the respondent is confined to report their travel that occurred within their assigned 2 (or 4) week period for the year, i.e. the collection period. This restriction affects the reporting of infrequent long haul trips, which also tends to be few in number. Further, most of these long haul trips are made to pursue leisure or business activities rather than commute or shopping. The ATS in this regard, with its longer recall period, and more importantly an annual

collection period provides a more comprehensive measure of annual long haul trips across households.

The annual time frame for collection also makes it straightforward to arrive at annual estimates of long distance travel per capita as well as that for the nation from ATS. In contrast, NPTS (NHTS) with a shorter two (four) week recall period to an extent minimizes the recall bias (in reporting short haul trips), but does not precisely yield national annual travel estimates. This is primarily due to current method of expansion applied to the single wave of data collected with a shorter recall period (NPTS handbook, 1998). It does not mean that surveys with shorter recall periods cannot be used to arrive at annual estimates. More importantly, the weighting method implemented to expand the single wave of sample data collected with shorter recall period needs to be revised, demonstrated latter in this section.

Besides minimizing resource requirement, the motivation to tradeoff between longer recall period and annual collection time frame (ATS) against shorter recall period (NPTS/NHTS) could also depend on whether the objective is to collect annual short haul trips (200 miles or less) or long haul (more than 200 miles) travel estimates. An ideal survey that could provide both of the above estimates would require one to collect data from each respondent with shorter recall periods (2/4 week) as in NHTS, but within an annual data collection time frame as ATS. The proposed approach although ideal, would be more expensive and burdensome for the respondents, having to report their travel over about 26 (2 week) or 13 (4 week) recall periods in a yearlong collection time frame.

Thus the two main aspects of data collection that affects reporting of long distance trips are i) the recall period and ii) time frame of data collection, with former predominately affecting short haul trip estimates, and the time frame (and time of collection) affecting long haul trip

estimates. With this perspective, long distance travel surveys could be divided into two segments, the first segment involving frequent trips to short haul destinations such as for commute, shopping, weekend entertainment. While the second segment involving infrequent travel to long haul destination, such as those made to go on vacation or business. The effort to collect these two distinct forms of long distance travel within a single survey leads to the need to tradeoff between recall period and time frame of collection. Giesbrecht and Bose (2005) suggests having different recall periods to collect travel for different extents. For example, enquiring respondents to recall trips that were made between 50 and 100 miles in the past month, and recall trips made to destination over 100 miles in the past three months.

The above proposed approach to data collection could lead to gathering appropriate long distance trips across all purposes including leisure, which makes up more than 50% of long distance travel. Leisure travel although observed to make up a major share of long distance trips (Figure 3) across surveys, could be distinctly different in terms of their miles traveled i.e. share of short and long haul trips. For example, the leisure trips being reported in ATS (longer recall period and collection time frame) could be more biased towards reporting trips to distant destinations (e.g.: vacation). In contrast, the trips reported for leisure in NPTS/NHTS (with shorter recall period) might be more biased towards reporting short haul trips (e.g.: shopping, weekend recreational trip). This becomes evident from the subsequent analyses presented in this chapter, with leisure as its central theme.

In addition to examining the above mentioned difference in the trips and miles traveled, leisure trips are also analyzed based on specific motivations (Figure 4). The first segment consisting of trips to visit friends and relatives, makes up about 50% of all leisure travel reported annually, with rest being made for non-visiting leisure motivations such as vacation,

entertainment, shopping etc. Each of the above motivations (visiting and non-visiting leisure) could be pursued at short haul (100 to 200 miles) as well as long haul (over 200 miles) distances. Surveys with shorter recall period – single wave of data collection (e.g.: NPTS) and longer recall – multiple waves (e.g.: ATS) are likely to predominately report short or long haul trips respectively. Thus, from the perspective of developing long distance leisure travel model, survey with longer recall period might be more appropriate, as household would often these trips after advance planning and considering their annual time and money budget. Besides the above, the trips reported in surveys with longer recall period and annual collection frame (e.g.: ATS) could be interpreted in straightforward manner to analyze total annual travel as well as across specific purpose.

In contrast, the estimates from shorter recall surveys (NPTS, NHTS) calculated using daily person trip weights could not readily lead to such annual estimates of leisure travel for visiting or non-visiting purposes. This is because, currently, sample trips from these surveys are weighed using daily person trip weights scaled by the ratio of number of days in an annum to the period of the data collection. In essence, it factors up the individuals travel reported during a pre-assigned two-week period to be representative of the rest of the 50 weeks in the year in NPTS (1995). The above-described approach, would not lead to appropriate estimates of annual trips and miles per capita (NPTS 1995 handbook, 1998). This is because an individual could exhibit different trip rates and travel extent across season, in particular for leisure travel. As a consequence, present estimates from shorter recall period (ex: NPTS) are bound to lead to imprecise annual estimates of long distance leisure travel, unless an appropriate weighting procedure is implemented to account for seasonal variations. This research further presents an alternate approach to expand data collected from shorter recall periods, particularly with single

wave data collected with a time frame of less than one year (NPTS, NHTS) to arrive at annual estimates.

Prior to pursuing the above analysis, the influence of reducing the collection time frame and time of collection on reporting of long distance leisure trips is demonstrated using ATS (1995) data. Specifically, this analysis reveals the effect of shortening the collection time frame from one year to six months to three months. Assuming, each household in ATS is confined to reporting their travel to one time frame for a year. An examination of the number of leisure trips with a collection time frame of 3 months (Figure 5) reveals about 30 - 60% of the households to report no long distance leisure trips, with the least in the third quarter (summer) and the most during first quarter (winter) as expected. The share of households reporting no long distance leisure trips drops significantly on increasing the reporting time frame to six month (Figure 6), with more households reporting during the 2nd half than 1st half of the year. The analysis shows two features of the data collection that influences the reporting of long distance leisure trips. First being the time frame of collection, and second the time of collection (reporting) for each respondent. The latter is more relevant from the perspective of collecting leisure travel, due to its seasonal nature.

It is evident from above that the infrequent nature of long distance travel causes a significant amount of variation in the number of trips being reported based on the time of data collection and time frame of collection. Specifically, about 50% of the variation in trips reported by households are observed to occur due to the 3-month collection time frame, which is found to drop by a fifth when expanding the collection time frame to six months (Figure 7). The variability could be expected to be far higher with shorter period of collection such as two (2) or four (4) weeks implemented in NPTS (1995) and NHTS (2001).

The above discussion suggests that just a month of data used to represent the travel for other months, as undertaken at this time in NPTS to estimate annual travel might not be appropriate. To mitigate this issue, Giesbrecht and Bose (2005) suggested a panel design (more than one wave) to collect such data for a household over a full year as was discussed earlier in this chapter. However, this could be resource intensive, as it would require undertaking a survey similar ATS (1995), with different recall periods for short and long haul trips. Figure 7, also shows the variation in leisure travel across households due to the differences in socio-demographics (family type, age, income etc.). The influence of these aspects are examined in subsequent sections of this chapter by analyzing these surveys and other supplemental data sources.

The next section proposes a methodology based on above discussion to expand single wave of data collected from respondents with two (2) week recall period (NPTS, 1995) to estimate annual long distance travel. The NPTS is considered in this analysis instead of NHTS (2001), despite it being a more recent dataset because the latter is atypical with the data collection including travel (pre)post 9/11. Further, NPTS (1995), although an older dataset, was collected for about the same time period as ATS (1995).

2.2 ATS and NPTS Data Processing

Prior to implementing the alternate weighting method, NPTS and ATS data were further processed. Overall, NPTS data, as expected, with a 2-week recall period, reports more trips than ATS (Figure1). However, on implementing the 100-mile SLD filter, ATS reports more trips than NPTS (Figure 2), with latter losing a significant share of frequent (routine) short haul trips due to the filter. Further, NPTS just reports trips on individuals aged 5 and over compared to ATS, which reports travel for all age groups. Thus to make a consistent comparison, ATS data is

processed to account for only trips made by individuals aged 5 years and over. It results in ATS (1995) to report about 775 million trips compared to NPTS (1995) reporting about 713 million trips. Overall, NPTS reports fewer long distance trips, but interestingly it reports a higher number of trips in the short haul segment, i.e. to destinations within 200 miles (Figure 8). These trips are most likely frequent (routine) trips such as those made for commute, which was excluded from ATS by design. While, ATS with a longer recall period, reports more long haul trips (200 miles or more) distant destination(s) as expected. Further, with an annual collection time frame , i.e. 4 waves of data collection, it could be expected to report more total trips compared NPTS, with the 100 mile filter and single wave of data collection.

Overall between ATS (1995) and NPTS (1995), the former reports about 60 million trips more than NPTS (1995). It is to be noted in this comparison, that aggregation of the weighted bi-weekly trips from NPTS over the year does not necessarily yield annual estimates (NPTS handbook, 1998). It is because NPTS expands each bi-weekly sample to annual estimate by multiplying the person trip weights by scale of $(365/14)$, i.e. about 26 two week periods in an annum. The above approach assumes population represented by each two-week sample to make same amount of trips in the remaining 25 two weeks in a year, ignoring seasonal variation. In addition, the samples' being non-uniformly distributed across two-week periods over a year does not allow to capture this variation either. Hence, this section aims to mitigate this issue using an alternate method to expand NPTS data to arrive at annual estimates presented in the next section.

2.2.1 Alternative Trip Expansion Method for Data with Shorter Recall Periods

The sample trips reported by individuals (respondents) pre-assigned to different periods (months/2 week periods) in a year is undertaken in three steps for each month of data collection: i) trip weight modification ii) non-interviewed population trip estimation and iii) total population trip estimation. In this method, the total population trips estimated from the last step for each month is aggregated to arrive at an annual estimate of travel for the entire population (Table 2). It is to be noted that the expansion process is undertaken for a month rather than a two-week recall period. It is because the survey only provides the month of interview, and not the specific two-week recall period assigned to each respondent.

2.2.1.1 Step 1—Monthly Trip Weight Generation

This step modifies the NPTS travel period trip weight formulation by replacing (365) with the number of days in the month in which the sample respondent were assigned to report their trip ending in the two-week recall period (Equation 1). It results in a monthly trip weight.

$$\text{Monthly Trip Weight} = \frac{(\text{Person Trip Weight} * \text{Number of Days in the Collection Month})}{\text{Period of Data Collection (i.e 14 days)}} \quad (1)$$

Expanding the monthly samples using the weight in Equation 1 results in an estimate of travel during that month (season) of the year (Table 2, Column 2) for population represented by the sample in that month (not for the entire population).

2.2.1.2 Step 2—Interviewed Population Trip-Rate Estimation

Subsequently, the population travel estimate for each month is estimated by aggregating travel estimates for each month for the population represented (and not represented) by the sample for that month. The monthly trip estimates for the population not represented by the sample interviewed in the given month is estimated by applying the per capita trip rate for

segments of the population represented by each months' interviewed sample to the non-interviewed population in that month. Here, it is assumed that the per capita trip rates estimated for segments of population represented by sample in a given month is applicable to the rest of the population as well. It could be further developed to obtain differential trip rate across socio-demographic groups. The per capita trip rate per month (Equation 2) for the population (column 3) represented by the sample in each month is calculated as the ratio of the estimated trips (column 2) to the population that is represented by the sample.

$$\text{Interviewed Person Trip Rate} = \frac{\text{Total Person Trips Reported Per Month}}{\text{Total Persons Interviewed Per Month}} \quad (2)$$

2.2.1.3 Step 3—Total Population Trips per Month

The product of the interviewed person trip rate (Table 2, Column 4) and the population not represented by the sample interviewed in each month (Table 2, Column 5) provides an estimate of the person trips made by the population not represented by the interviewed sample in each month (Table 2, Column 6) as shown in Equation 3 below.

$$\text{Non Interviewed Person Trips} = \text{Interviewed Person Trip Rate} * \text{Non Interviewed Population} \quad (3)$$

The non-interviewed population in each month is calculated by subtracting the population represented by the sample in each month (column 3) by the total of column 3 (national population estimate), i.e. 241 million people for 1995. An aggregation of estimated person trips from column 2 and 6 in table 2 results in monthly population trip estimates, which on being aggregated across months provides an annual person trip estimate. It results in a total estimated annual trips of over 730 million trips, within 6% of the person trips estimated from ATS (1995) for population aged 5 and over (Table 3) over an annum. A comparison of the annual estimate

arrived at through the above illustrated alternate method of expansion is found to result in 20 million more trips than that estimated using the original expansion procedure in NPTS (1995).

This is a significant improvement but still 43 million less than the ATS (1995) estimate as can be seen in Table 3. Further, a comparison of trip length distribution (Figure 9) reveals the frequent (routine) trips to scale up the most on expanding the NPTS dataset, likely due to the shorter recall period. In comparison, there are insignificant but observable increases in trips to distant destination located between 200 and 1000 miles, and none for trips over 1000 miles. This difference could be more due to single wave of data collected in NPTS compared to ATS, which collects such data over an annum from all its respondents. This issue could be mitigated by implementing different recall periods for different travel extent as suggested by Giesbrecht and Bose (2005). Further, examination of the expanded NPTS estimates across quarters (Table 3) to understand the under reporting of trips showed NPTS estimates to be comparable to ATS estimates across all quarters except for the 4th quarter (~ 20 million trips).

2.2.1.4 Annual Trip Estimates by Region of Residence Using NPTS

The reason for underestimation of NPTS trips during the 4th quarter becomes evident on analyses of these trips across regions (Table 4), with NPTS estimates found to be reasonable for all regions, except the west. Specifically, the NPTS annual travel estimate for the west is 20% less than that reported in ATS. Further analysis of these differences across region and quarters (Figure 10) shows a significant share of this under-estimation to result from underestimation of such trips from all regions, during 4th quarter and particularly more so from the west. Across all regions, NPTS under estimated trips by about 15% in the 4th quarter, while that from the west was almost double (~30%).

The under sampling of respondents from the west could be one of the potential reasons for significant under estimation of trips (Figure 11). It is bound to not only affect the trip estimates but also the trip length distributions, as individuals from the west in particular with a thriving technology job sector and higher incomes in potential could make relatively more long distance trips, particularly to distant destinations. This also becomes evident from analyses of CEX data presented later in this chapter. Furthermore, a large share of population in this region happens to have migrated from other parts of the country and world, and hence may be more inclined to travel to prior household locations at distant destinations.

2.2.1.5 Annual Trip Estimates by Household Income Using NPTS

Beside the above, the socio-demographics of the individuals also have significant influence of reporting of long distance travel, with majority of these trips being made for leisure. This becomes evident from the trip statistics reported across income in Table 5 from ATS and expanded NPTS data. It is to be noted though; that these estimates (Table 5) are for individuals 18 years and older, and is not comparable to the prior estimates that were for individuals 5 years and older. Overall, as expected the number of trips reported (Table 5) increases with increase in income across both ATS as well as NPTS.

However, the annual estimates from NPTS is observed to fall short by about 40 million trips, and continues to show significant differences even on inclusion of trips from individuals not reporting their household income. The NPTS trips estimated for low and middle income groups are found to be within 2% of that estimated in ATS (Table 5), but are found to be significantly lower for the high income group. Assuming the rest of the non-reported trips to be those undertaken by high income individuals leads to high income individuals to make about 310 million trips, still 10% less than the ATS estimate for the same group.

Overall, both datasets (NPTS and ATS) presents a similar distribution pattern with high-income individuals found to be responsible for about ½ of all annual long distance trips, with the remaining ½ split 70/30 between medium and low-income groups. The same is the case in terms of per capita trip rate estimates (Figure 12) , with ATS reporting high income group individuals to make most of trips per annum (just over 5 trips per annum) compared to middle income trip making just below 5 trips per annum. However, NPTS (1995) suggests high income group to make fewer trips per annum compared to middle income, even after accounting for the trips reported from individuals (assumed to belong to high income group) that did not report their income.

Some of the above issues could be the artifact of the under-sampling of the individuals from the west, with this region being known to constitute high income earning individuals. Further, it can also be noticed that the per capita trip rate for low-income group is distinctly lower compared to the middle income and high-income group. This implies that there could be significant drop in annual trips with economic downturn (e.g.: loss of employment, income), such as that witnessed in the recent past (e.g.: 2009 recession). This might have adverse economic influence on both tourism and transportation sector, with about 60% of the population making long distance trip (McGuckin, 2013) in normal economic conditions.

As is evident from this section, NPTS (1995) data collected in single wave of two-week collection period could be a viable alternative to estimate annual long distance travel estimates. It provides reasonable estimates across specific segments such as quarters, regions and socio-demographics, few of the factors that were identified to be highly correlated to analyze long distance travel (Contrino et al, 2000). The proposed method of expansion to arrive at annual estimates could be further refined through segmentation across socio-demographic groups.

However, prior to pursuing this effort, the data collection design would need to be improved. In particular, the annual estimates could be made more precise, through undertaking a sampling procedure that conforms to the population distribution across geographies (spatial sampling), along with uniform sampling distribution over time (ex: quarters) (temporal sampling).

2.3 Annual Trip Rate Estimation Using NHTS Daily Travel Data

Besides long distance surveys, this research also explored the potential to understand annual long distance travel across the U.S. using daily travel surveys. These surveys (e.g.: NHTS, 2009), although primarily designed to collect local (urban) travel, do implicitly collect long distance travel, at least a segment of such travel made on the survey travel day. It might not provide as comprehensive a picture of long distance travel as traditional long distance surveys. However, it could provide a small but representative sample of such travel over a year. This section further explores long distance (i.e. 100 or more miles in length) daily trips from NHTS (2009) from this perspective.

A distribution of daily long distance trips (1st plot from top in Figure 13) shows a seasonal pattern similar to that obtained from long distance surveys, with leisure trips identified as the dominant travel purpose around holidays and long weekends (2nd plot from top in Figure 13). The trips in these plots are reported across weekdays in a month as NHTS only reports the month and day of travel. Overall, leisure trips dominate all long distance trips, as in long distance surveys. However, the shares of long distance leisure trips are relatively lower than expected around prominent holidays (e.g.: Thanksgiving, Memorial Day). This could be due to lack of sample to represent travel around these dates. In contrast, it reveals significant travel around non-traditional holidays such as Valentine's Day, not evident in long distance surveys. This section

further investigates the estimates of long distance travel from daily travel surveys (NHTS, 2009), and compares them against ATS estimates (1995).

A weighted estimation of all daily long distance (determined using `trp_miles` variable) trips from NHTS (2009) results in about 3.4 billion trips, four times more than ATS (1995) annual estimates (~800 million). A share of the above NHTS (2009) estimates could be attributed to natural growth, but a significant share of these trips might not be appropriate for comparison with ATS due to numerous reasons. This becomes apparent from the trip length distribution (Figure 14), which shows over 2/3rd of NHTS trips made to destination within 200 miles compared to 1/2 from ATS. First, a significant share of these trips could be made for frequent purposes such as commute, specifically excluded from ATS. Second, this share could also include a segment of long distance trip, and not the entire trip as the questionnaire asks to report the travel that was undertaken on survey travel day. This becomes apparent from examination of the trip length distribution by mode, which shows over 70% of auto trips reported to be made to destinations within 200 miles, while the majority of air trips are found to be made to destinations over 500 miles. Third, the NHTS (2009) reports distance traveled using `trp_miles` variable, a calculated distance (e.g.: from reported blocks traveled etc.) converted to miles. This could also have led to higher number of circuitous interurban travel being reported (less than 200 mile). The above issues could be addressed in daily travel data collection through asking respondents to provide more detail on long distance travel, including information on their longest trip that ended on the travel day. The current survey does include some variables that provide insights in this context, such as the out of town variable.

The out of town variable in NHTS (2009) asks individuals whether one was out of town the entire travel day, and if so it further enquires on whether he (she) was within or outside the

country. Based on the above question, it only collects trips that were made on being within the country. The NHTS (2009) estimates about 6.4 billion such trips to be made within the U.S. from a sample of over 22,000 person trips. This includes all trips (local as well as long distance) that respondents made on being out of town. From above, about 280 million trips were identified to be long distance trips, making up for about 20% of the trips reported in ATS (1995). This would be expected, as the estimate of long distance trips comes from an extremely small national sample of over 1000 person trips from the overall out of town sample of the 22,000 person trips in NHTS (2009). Despite the sample size, it shows a trip length distribution (Figure 15) similar to ATS (1995). This investigation suggests that incorporating additional questions to collect long distance travel within daily travel surveys along with appropriate weights for such trips could potentially lead to more reasonable estimates.

2.4 Motivations for Long Distance Travel

As is well known, annual long distance trips are predominately made to pursue leisure, and can be broadly categorized into two prime motivations a) visiting, i.e. to visit friends and relatives b) non –visiting , i.e. to pursue leisure (non-visiting) activities such as sightseeing, rest/relaxation, outdoor recreation and entertainment. The above household leisure travel preferences have dominated long distance travel for decades (ATS, 1995; NPTS 1990, 1995 and NHTS 2001), and could be anticipated to continue into the future as well. However, their travel extents could change over time depending on household budgets, national economic conditions, and changing travel options (e.g.: hybrid cars, high speed rail etc.). Hence, it is essential to understand the above trend over time, but is difficult to obtain such insights, as collecting annual long distance survey itself is extremely resource intensive.

Thus, this research explores alternate data sets that could provide a surrogate perspective on this form of travel, which may also be synthesized with surveys to model long distance travel behavior of households over time. The annual Consumer Expenditure Survey (CEX) microdata published by Bureau of Labor Statistics is an excellent source, particularly to analyze annual household leisure travel. This becomes evident from the analysis of the trend in an average consumer unit (households') annual out of town expenditure over the past decade (Figure 16). The above expenditure trend is also available at disaggregate levels across different categories of expenditure (travel, lodging etc.) as well as by socio-demographic characteristics, which could prove beneficial in analyzing annual household leisure travel.

2.5 Factors Influencing Long Distance Travel

A major share of the above reported out of town expenditure for an average household is incurred on vacation travel (Figure 17). However, about 60% of household across the U.S. do not make any long distance trip in a year (McGuckin, 2013). Further, those households that do undertake a long distance trip might not necessarily make such trips to pursue leisure activities, such as going on vacation.

Given the consumption oriented nature, the number of households making none, one or more long distance trips per year could be distinct across income groups, with higher income households having a higher likelihood of making one or more such trips. Besides income, leisure travel preferences could be influenced based on household composition (family size, workers per household etc.), which could significantly alter travel expenditures. Mallet (2001) analyzed long distance travel by segmenting households based on their income and composition. Travel surveys provide detailed information on household and trip characteristics but do not usually include expenditures on such trip, vital to understand long distance travel, particularly leisure

travel. To this effect, the CEX data collects expenditure across socio-demographic group on an annual basis, making it an excellent source of data to analyze such behavior across segments over time.

The CEX data, besides being used to analyze trend in leisure travel over time, could also supplement cross-sectional household level travel data (e.g.: ATS, 1995) to model households' annual leisure travel decisions (e.g.: purpose, destination, mode, trip frequencies). Specifically, the data could be used to estimate the leisure travel budget for households based on their socio-demographic segments (e.g.: income, household type etc.). To undertake the above effort, this expenditure data would need to be disaggregate, i.e. be available for transportation and lodging expenditures, which make up major share of leisure travel expenses. The subsequent sections in this chapter examine CEX dataset from this perspective.

2.6 Components of Travel Expenditures

The major components of a households' annual expenditure that CEX collects on out of town travel are on transportation, lodging, entertainment, food and beverages. On average, households spend the most on transportation followed by that on lodging and food (Figure 18). As a behavior, households could minimize expenditures at destination (e.g.: lodging) based on their travel purpose, i.e. visit instead of pursue non-visiting leisure. However, they more often than not (on being reimbursed) have to incur transportation expenditures. Further, the transportation expenditure could either be a nominal expense on gas (auto travel) or a significant expenditure (airfare), depending on distance to destination and travel party size.

The above expenditures were identified to be major factors influencing long distance travel decisions (Sharp, 2011). To this effect, households could alter their choice of destination or number of trips based on their anticipated expenditure. It could as well change over time with change in economic conditions. In this regard, it is interesting to see (Figure 18) that households (consumer units) significantly curbed their expenditure on transportation but not much on their expenditure on food and lodging during recession. This could be either due to household reducing their recurrent trip to a destination or substituting their annual vacation trip to distant destinations with travel to local attractions (e.g.: staycation phenomenon).

The reduction in leisure travel might not be the same across all households. It could be distinct across households based on their socio-demographic segments such as income, age, and region of residence. Hence, it is vital to understand the extent of travel across aforementioned segments, from the perspective of conceiving a model to estimate annual leisure travel behavior. These socio-demographic variables could have strong explanatory power in estimating annual household leisure travel budget, a key input constraint to model leisure travel behavior.

2.6.1 Sociodemographic Preferences

Thus, with the intent to synthesize CEX with travel survey (ATS), two socio-demographic segments (age and income) were examined across datasets to understand the correlation between them. This was pursued through comparison of the levels of travel from ATS with expenditures for similar segments from CEX data. The ATS data was analyzed by segmenting its trips based on whether it was undertaken on (or off) vacation, and then based on specific leisure trip purpose. It was segmented across two levels because ATS (1995), unlike other surveys, introduced vacation as a separate flag variable rather as an enumerated leisure purpose. Vacation, as per the dictionary means to travel to out of town location for more than a

day, which implicitly indicates whether it was a daytrip or overnight trip. Thus, this variable in combination with enumerated leisure purpose helps in better understanding different extents of travel for a given leisure purpose. It can be seen from analysis of the trip length for trips reported to be made on and off vacation (Figure 19). It shows significant differences between trips made on and off vacation for same leisure purpose, with off-vacation trips made most likely to destinations that are within a day's reach, compared to those made on vacation to distant overnight destination.

Further, the off vacation trips could be more frequent compared to the vacation trips more likely being an occasional trip to distant location. The latter, vacation trips would require advance planning, considering household budget etc., which makes up a major share of annual out of town expenditure. Further, the leisure travel segments could be distinct across socio-demographic groups, based on their stage in life (age, lifecycle) and income.

Figure 20 reveals this through analysis of trips made on and off vacation for visiting and non-visiting leisure purposes across age groups. It identifies distinct levels of travel for two major categories, i.e. younger (35 year old and less) and older adults (65 and older) compared to the middle age working adults. The younger adults show higher propensities to undertake visiting travel, and more so on vacation. This could be representative of children traveling to visit their parents during school breaks. While, lower middle age adults (35-44 year old), and to lesser extent upper middle age adults (45-54 year old) are found to make more non-visiting leisure trips than visiting trips. In comparison, the elder adults (65 and over) are observed to make the least of all trips, but more often make visiting than leisure trips particularly on vacation.

Overall, the younger and older retired adults make more visiting than leisure trips, likely due to their economic condition as well as their health respectively. Further, the older adults might have already satiated their desire for leisure travel during the younger years, and are less inclined to pursue them with loss of mobility. The CEX analysis on annual out of town expenditures across age groups over the past decade (Figure 21) reveals a similar pattern across age groups. It identifies the elder and younger adults to spend the least on travel expenditures, likely due to them making more visiting than other leisure trips, and thus not spending as much on lodging etc. as middle age groups. It also shows an interesting trend with elder adults (65 and over) spending more on leisure post-recession, while the expenditure pattern of lower middle age group of working adults are declining likely a consequence of the impact of recession. Overall, the middle age group spends the most on out of town travel, which correlates with their travel extent reported in ATS (Figure 20).

As can be seen from ATS as well as CEX data, the younger and middle age group makeup a major share of the leisure travel market, with the younger household making more visiting trips and the latter making more leisure trips. Among the middle age group, the 55 to 65 year old are found to spend more on leisure travel, while elder adults (65 years and older) are identified to spend the least in travel due to a combination of factors including dependence on fixed retirement income and perhaps physical and stamina constraints.

The above reported average annual out of town expenditures across age groups could be incurred from them making one or more trips to one or more destinations over a year. Across these trips, the expenditures could be incurred on transportation, lodging, food etc. The above expenditures are the outcome of a household's travel decisions such as choice of destination, travel mode, stay duration, etc. The aforementioned choices tend to be significantly influenced

by their income along with other household characteristics such as their family size, number of workers etc. Contrino et al. (2000) identified that the above variables have strong explanatory power in modeling such travel, which is consistent with analysis of CEX data across these groups (Figure 22).

It shows the highest income group to exhibit a distinctly higher annual spending pattern compared to other income groups, exhibiting their affordability to pursue leisure, an inherently discretionary travel behavior. The influence of recession on such travel is also apparent, with an average consumer unit (household) from highest income groups significantly reducing their expenditure (~ \$500). In contrast, other income groups show a steady decline in their expenditure on leisure travel, and do not exhibit any significant year over year change. Alternatively, it shows relatively lower levels of participation in discretionary leisure travel from income groups other than those from highest income group.

The above identified drop in expenditure by the high income households could either be due to them curbing their vacation travel to distant exotic destinations (e.g.: national attractions or resorts) or substituting it with travel to a more economical/proximate destinations. This could have an adverse effect on tourism, in particular on the revenues at destinations that attract this group of households. It could also lead to reduced revenues for transportation providers, providing access to these destinations (e.g.: air travel). This becomes evident from Figure 23, which shows significant reductions in transportation expenditures, from households belonging to the highest income group (\$70,000 per annum and above).

2.6.2 Regional Patterns

The decision of household to switch their destination choice as mentioned in the last section, not only depends on the household characteristics, but also on the spatial distribution of

opportunities around the household. To this effect, BTS (2006) reports over 60% of long distance trips to be made within a households' state of residence, suggesting a major share of leisure opportunities to be within a household's state of residence. This could be due to frequent trips made to destinations within the state. However, a significant share (40%) of trips is made to destinations outside the state as well. These trips, involve significantly more miles per trip than within state travel. The share of trips made within and outside a households state of residence could be distinct based on their region of residence as well as trip purpose.

An examination of ATS data from this perspective shows the average trip length (Figure 24) to be distinct across the four census regions for each leisure purpose. Overall, business and sightseeing trips have the longest trip length with outdoor recreation and entertainment having the shortest trip length. Sightseeing having the longest trip length within the leisure purpose could be expected, with it being an occasional trip often to national attractions (e.g.: Niagara Falls or Yellow Stone Park).

In contrast, long distance trips for entertainment and outdoor recreation are the shortest, likely due to opportunities to pursue these activities being numerous and far more accessible. The trip length to visit friends/relatives shows the largest amount of variation across regions. Visiting trips, unlike non-visiting leisure trips, could be expected to be more influenced by presence of social ties than other attractions. This could be associated with migration of households for either employment or family reasons. In this context, the western households could be seen to travel the furthest and mid-western household to travel the least.

The above differences in levels of travel across regions based on trip length are also evident from the expenditure trend on out of town travel across regions (Figure 25). CEX shows the western households to spend significantly more compared to other regions. This could be

either due to: a) the leisure opportunities for western households on average are further than in other regions requiring them to travel longer, and spend more on such travel ,b) the inherent interests of individuals to travel to distant destinations (e.g.: higher preference to sightsee or pursue vacation at distant locations), c) higher travel costs (gas and airline prices). Besides the above, on average households from the west are likely to travel further to visit, due to the level of migration to the west from distant out of state destinations in the U.S. Consequently, a significant share of these migrants might be traveling back to rest of the U.S. to maintain their social ties. In this context, western households are more likely to utilize faster and more expensive modes (airplane) of transportation (McGuckin, 2012). This behavior results in higher transportation expenditures (Figure 26) for individuals from this region, and consequent overall higher leisure expenditure (Figure 25).

Further, the CEX analysis over time shows a significant rise in travel cost for western households in the middle of last decade. It shows the average western consumer (household) spent almost twice as much prior to the onset of the recession (2006) compared to the south. These expenditure patterns suggest the relative strength of technology sector with opportunities being far more lucrative in the west. All else being considered, Figure 26 primarily reflects the influence of economic conditions on household travel decisions that could be distinct across census regions as well as across households within each region.

2.7 Summary

This chapter analyzed long distance travel with two objectives, first to identify issues in collecting long distance travel data and second to conceive a theoretical framework to model national leisure travel demand at household level. The first effort commenced with understanding the issues in defining long distance travel followed by assessment of data

collection methods. Further, this research proposed an alternative method to expand data collected from NPTS (1995) with 2-week collection period and compared it against ATS (1995). The annual estimates from NPTS did show significant improvement from its original estimate, but was still lower than ATS and was biased towards reporting higher share of short haul trips.

Further investigation of NPTS sample revealed the under-estimation to primarily result from inappropriate spatial and temporal sampling. The western region was significantly under sampled relative to its share of population, while across all regions the fourth quarter was under sampled as well. Implementing the alternate expansion method proposed in this chapter after addressing the above issues should result in annual trip estimates comparable to ATS. Besides the above, the bias in NPTS, reporting higher share of short haul trips (less than 200 miles) likely resulted due to single wave of data collection with shorter two-week recall period. This economical approach better captures the frequent short haul trips, but tends to discount the reporting of infrequent long haul trips. This research suggests considering multiple recall periods for different travel extent as proposed by Giesbrecht and Bose (2005). In this context, short data collection periods could employ different recall periods for trips that are made on and off vacation to capture both the short haul as well as long haul trips. Incorporating multiple recall periods across travel extents could help mitigate some of the above-discussed bias towards short haul trips from collecting data with shorter recall periods.

In addition to NPTS, this research also analyzed daily travel data (NHTS, 2009) as a potential source to collect long distance travel. Preliminary analysis of long distance trips from the daily survey showed the expected seasonal pattern, but had significant bias towards short haul trips as well. A major share of this bias could have resulted from inclusion of frequent trips purposes such as commute and use of calculated distance variable (`trip_miles`). Besides the

above, respondents might have only reported a segment of their overall long distance travel adding to the bias in short haul trips due to the survey only enquiring on travel undertaken on the survey date. Thus, this research explored some additional variables (out of town), and was able to obtain a more reasonable trip length distribution, but with a very small sample (~ 1000 person trips) for the nation. Overall, the daily travel survey could yield better estimates with a larger long distance travel sample with questions on long distance travel, and separate weights for such trips.

The second effort analyzed long distance trips across different socio-demographic segments comparing the levels of travel from ATS against the expenditures from CEX. This was pursued with an intent to model household annual leisure travel subject to budget constraints. Specifically, this research analyzed household expenditures from CEX across prominent socio-demographic groups, and compared them against ATS estimates for the same groups to discern household travel behavior across these segments. This exploratory analysis of the CEX data revealed consistent levels of expenditure across select socio-demographic groups such as income, age and regions, confirming the usefulness of CEX variables in estimating annual household leisure travel budgets. The household budgets estimated from this process could be very useful in analyzing and modeling their annual leisure travel behavior, i.e. their choice of destination(s), mode of travel, and trip frequencies. Accordingly, these variables are further constructed and implemented in subsequent chapters of this dissertation.

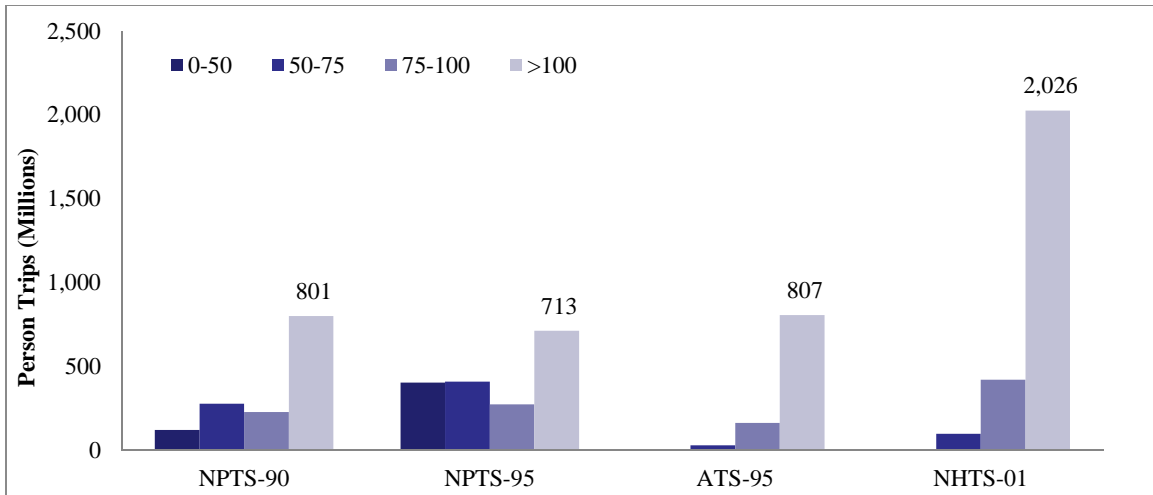


Figure 1 Weighted annual person trip length (greatest circle distance miles) distribution

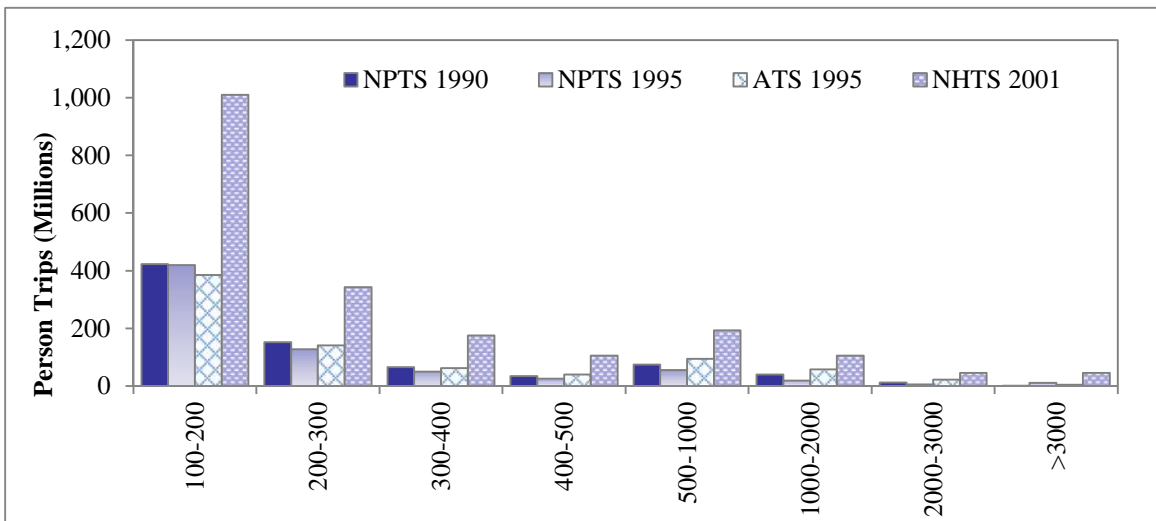


Figure 2 Weighted annual person trip length distribution by survey

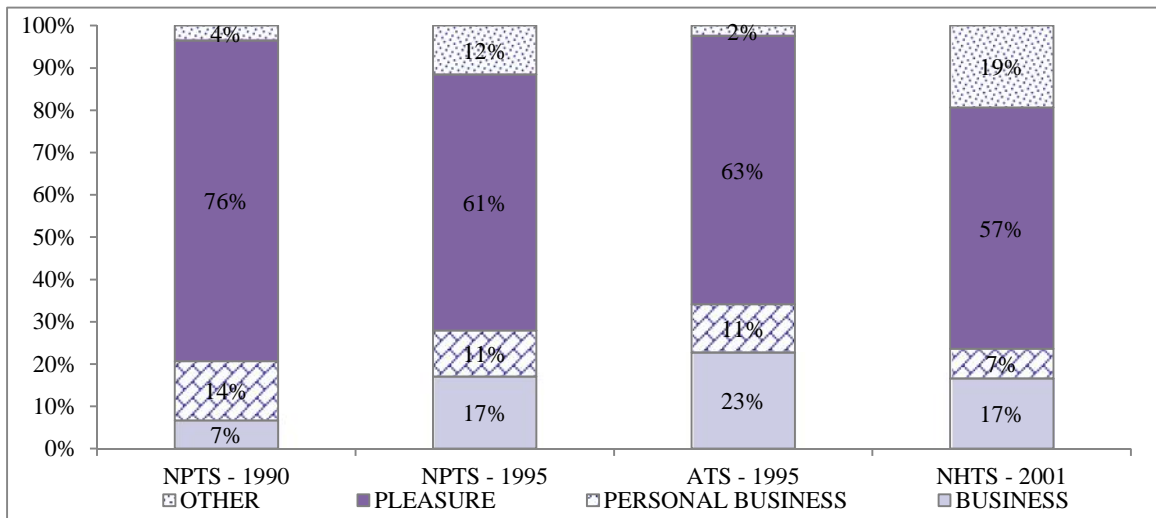


Figure 3 Weighted total person trip share distributed by purpose

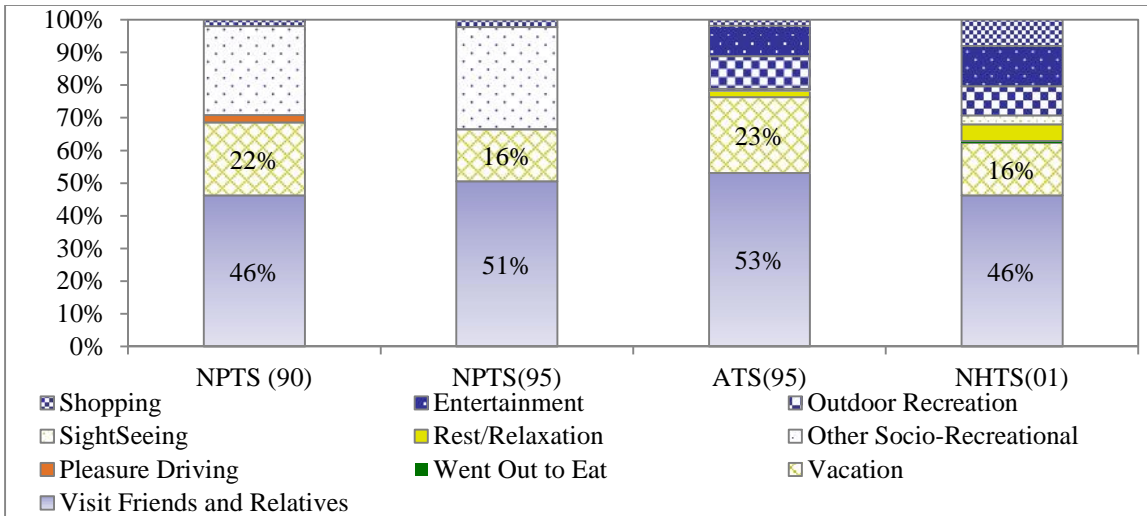


Figure 4 Weighted annual person trip distribution by specific pleasure purpose

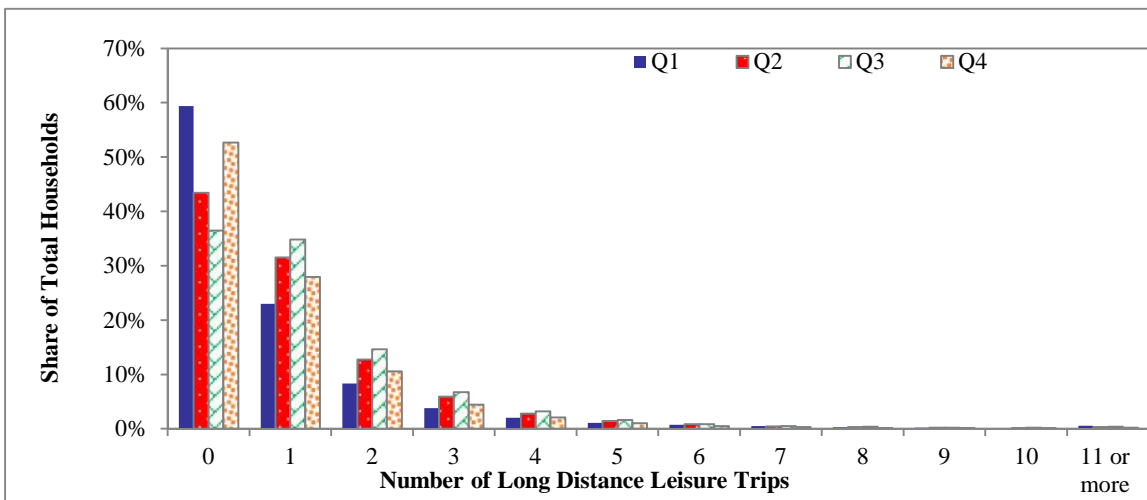


Figure 5 Share of total households making no, one, or more trips per quarter

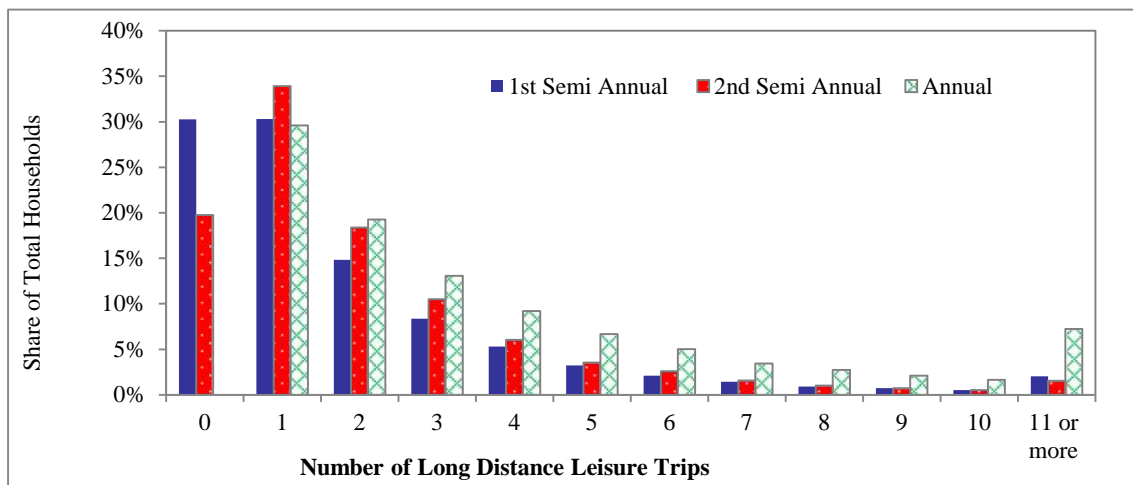


Figure 6 Share of total households making no, one, or more trips per semi-annum and annum

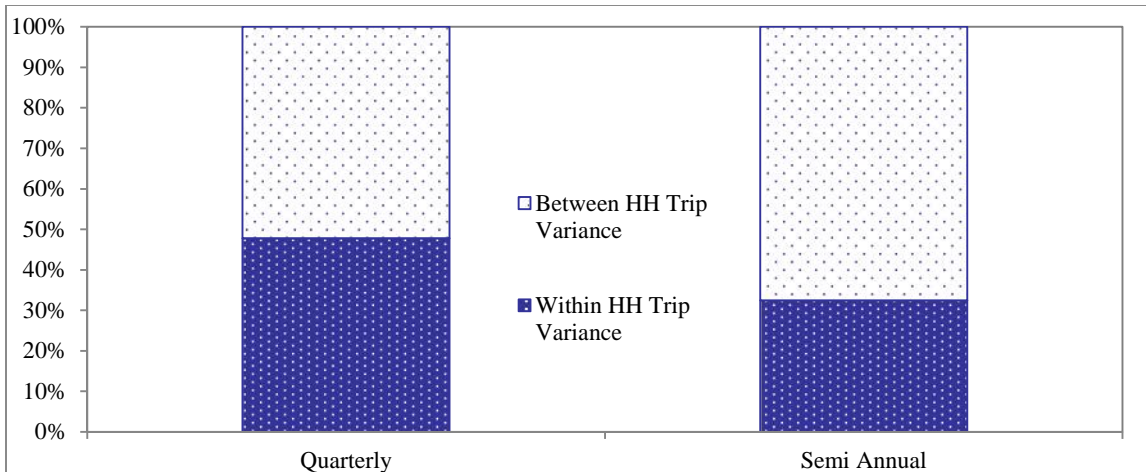


Figure 7 Interhousehold and intrahousehold variation in reporting trips across time periods

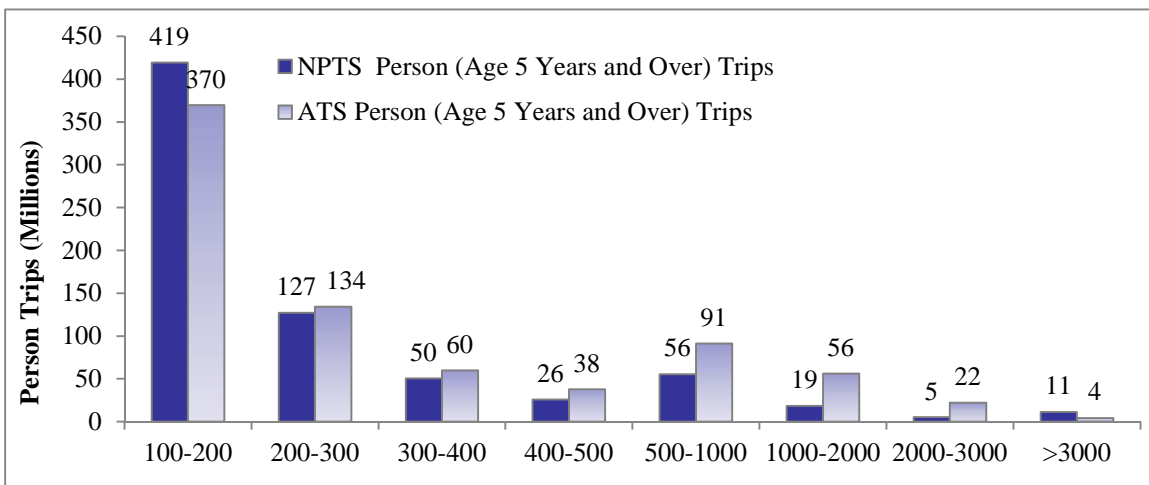


Figure 8 Trip-length distribution in greatest circle distance miles of trips made by persons 5 years and over in an annum

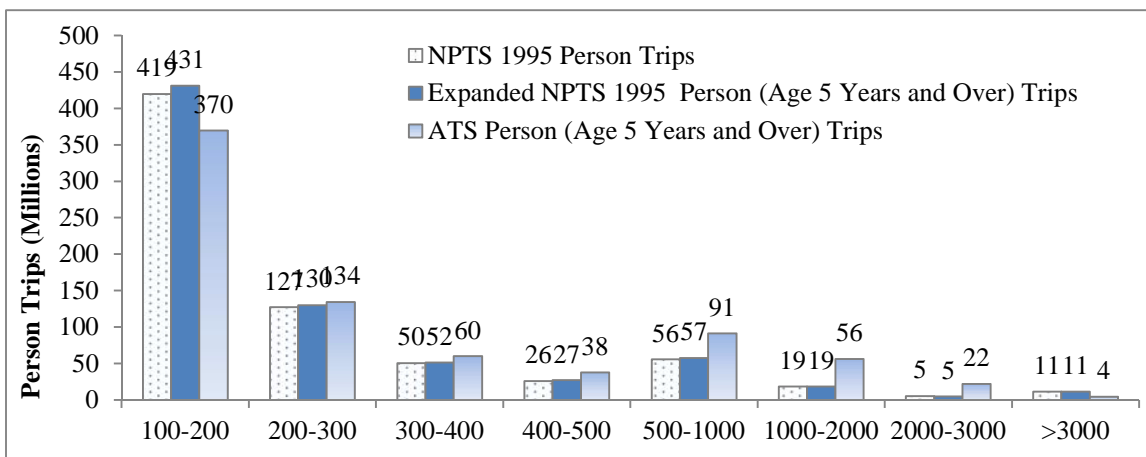


Figure 9 Expanded NPTS trip length distribution

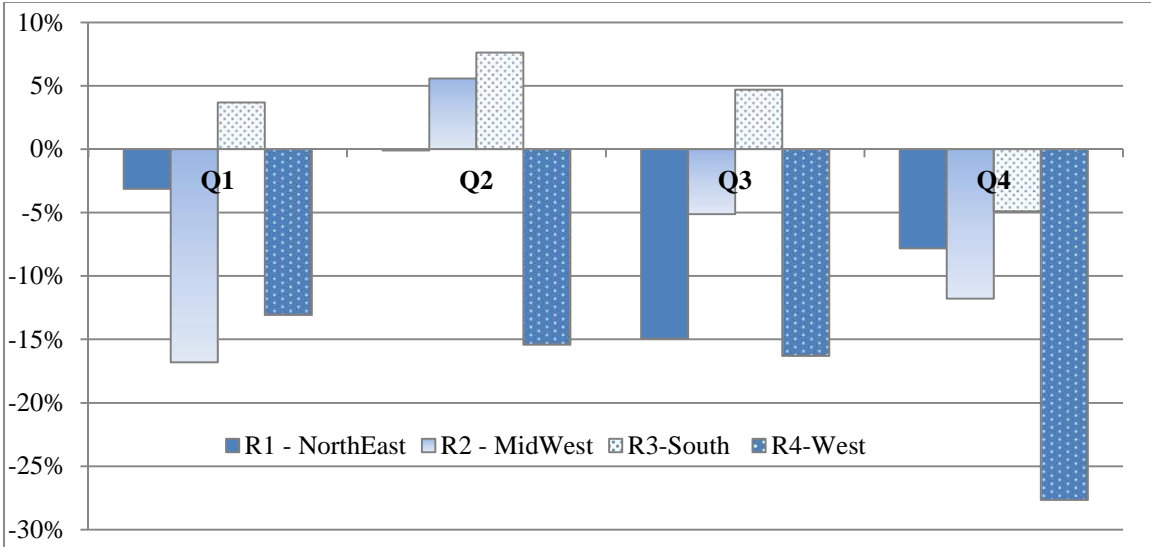


Figure 10 Percent difference in NPTS vs. ATS trip estimates by quarter and region

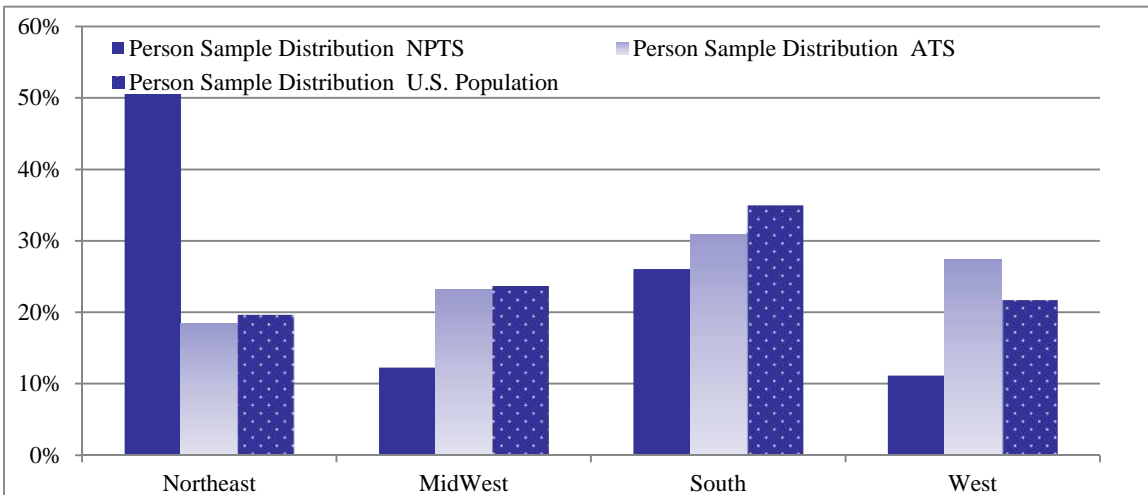


Figure 11 Sample (persons) interviewed by region and survey vs. population

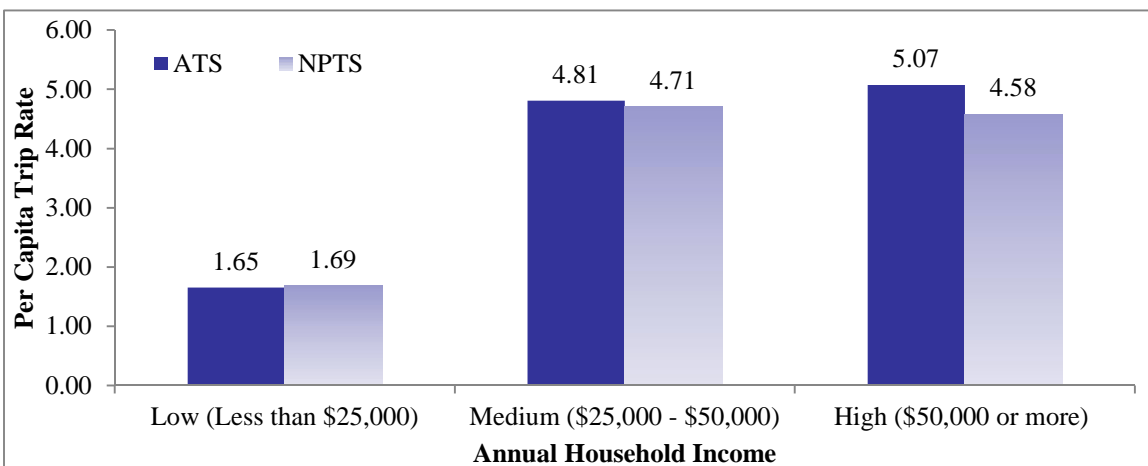


Figure 12 Per capita trip rate by household income of person

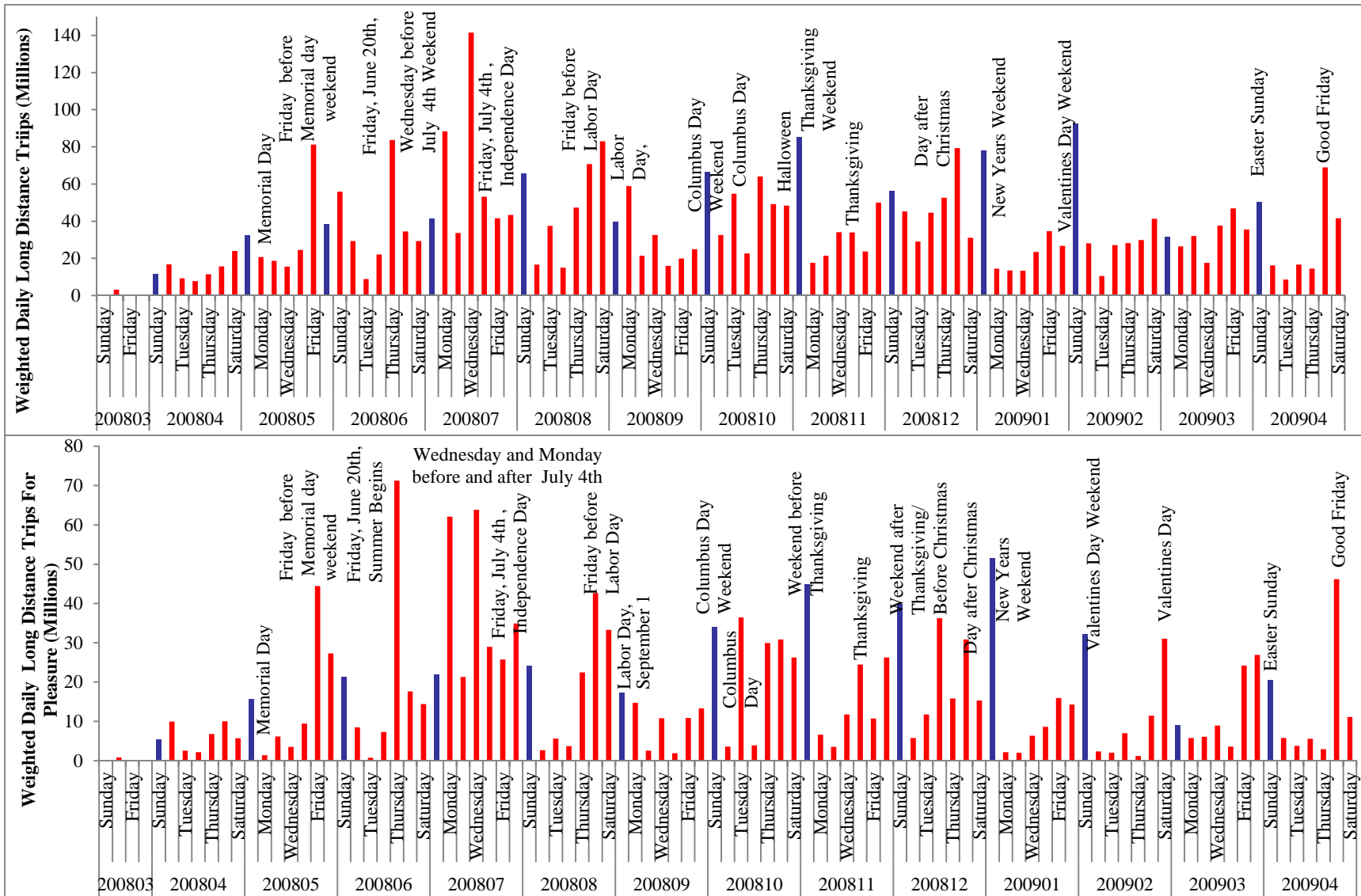


Figure 13 2009 Nationwide Personal Transportation Survey—Daily long distance travel distribution

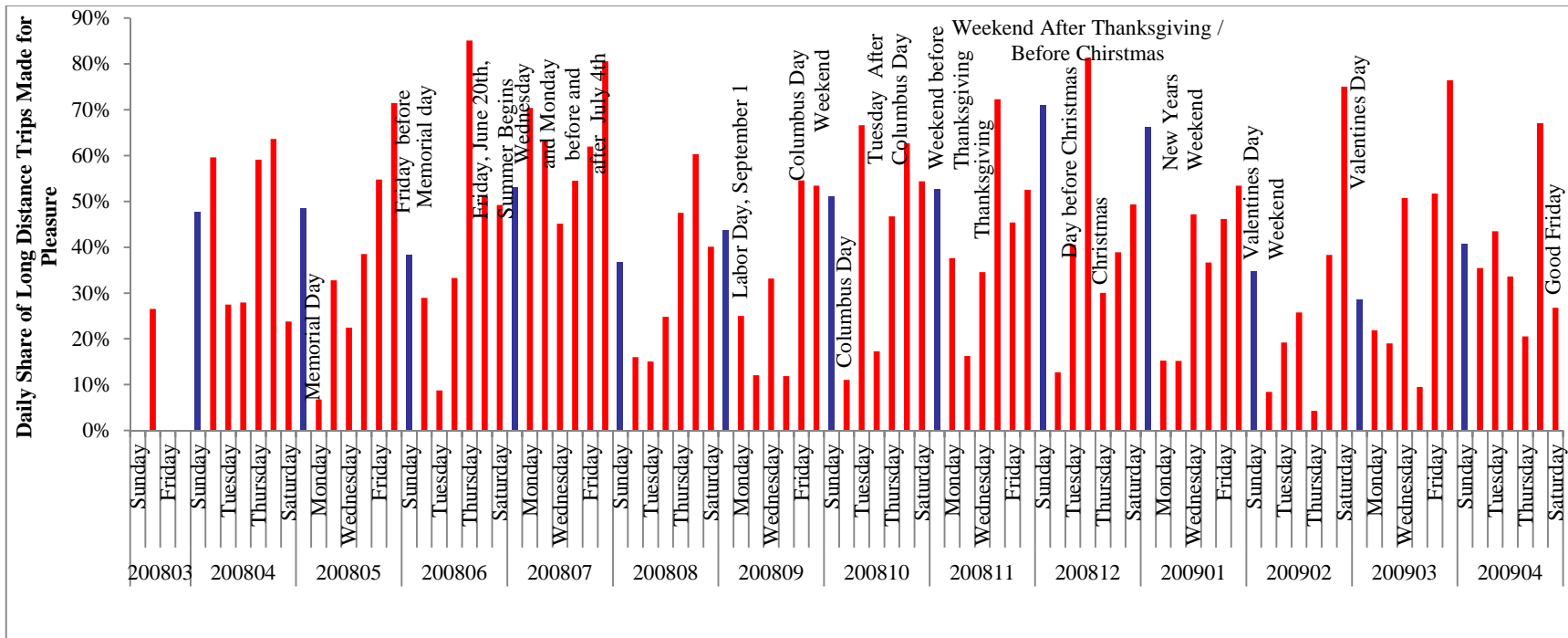


Figure 13 (Continued)

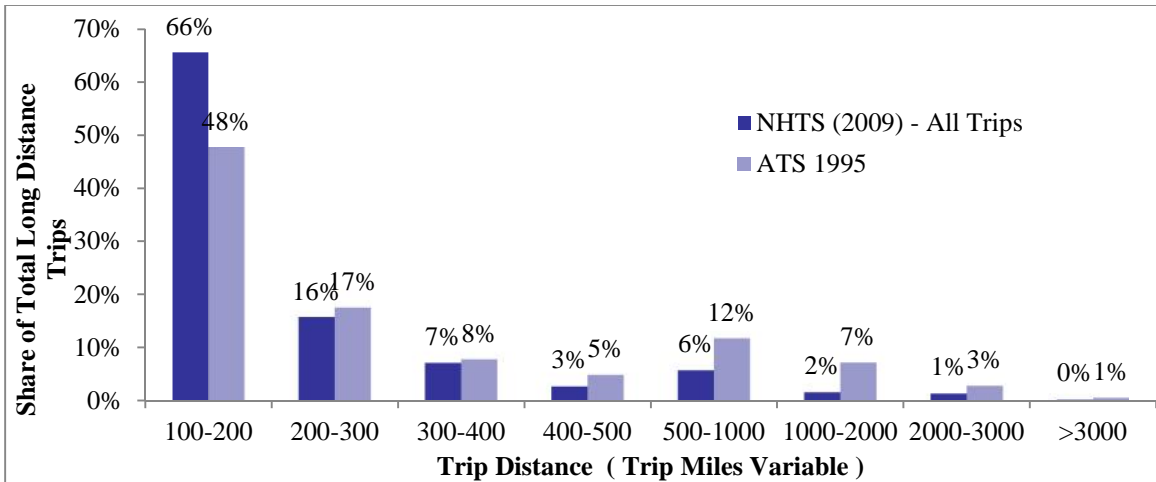


Figure 14 American Travel Survey/ Nationwide Personal Transportation Survey- Long distance trip length distribution (all person trips over 100 miles)

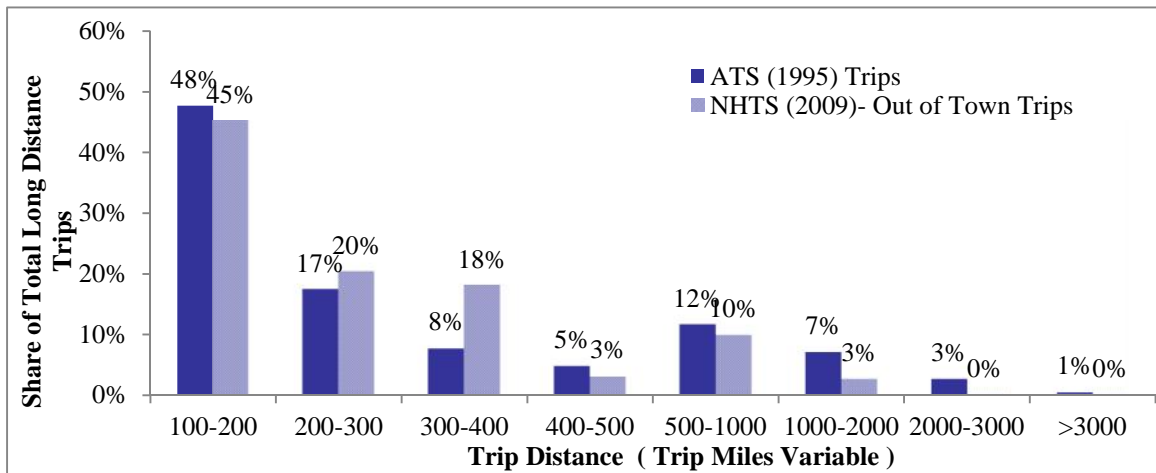


Figure 15 Long distance trip length distribution (person trips over 100 miles)

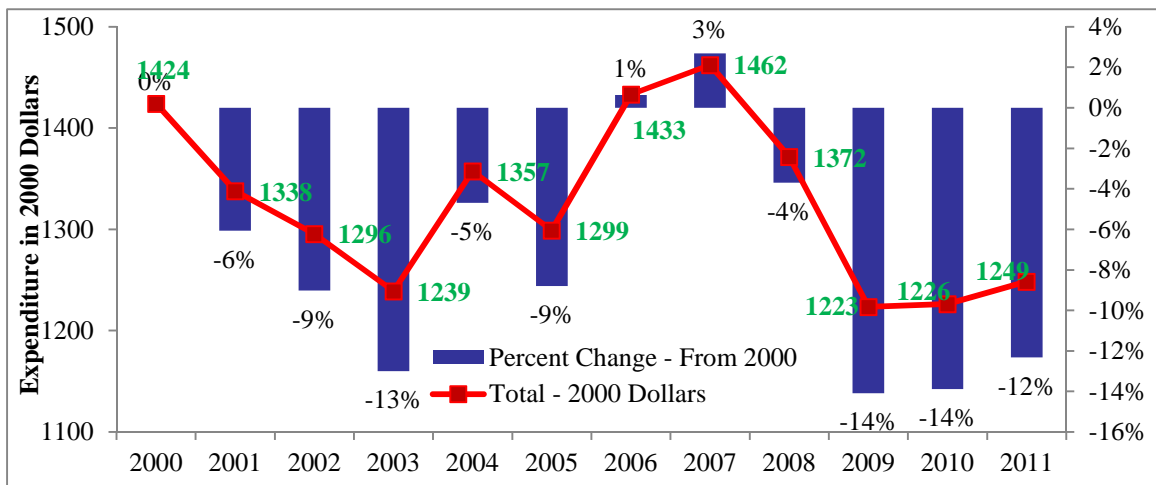


Figure 16 Average consumer expenditure on annual out-of-town trips

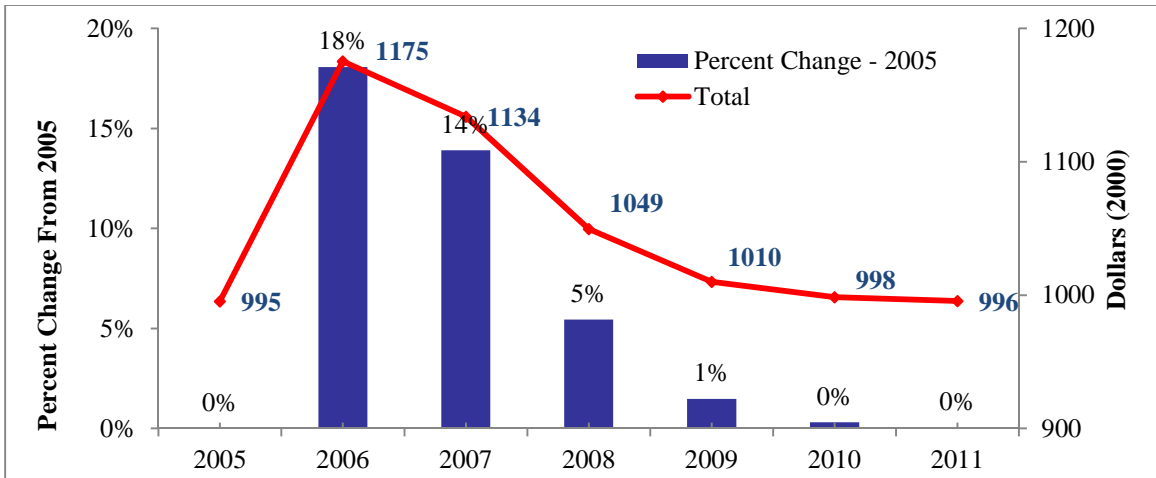


Figure 17 Average annual expenditure on vacation by consumer unit

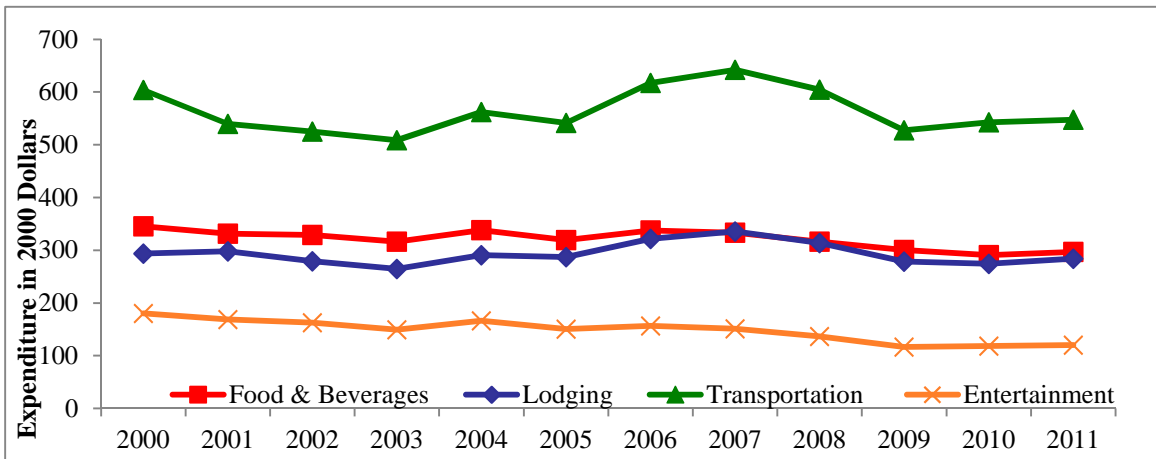


Figure 18 Average annual Consumer Unit expenditure on out-of-town travel by expenditure category

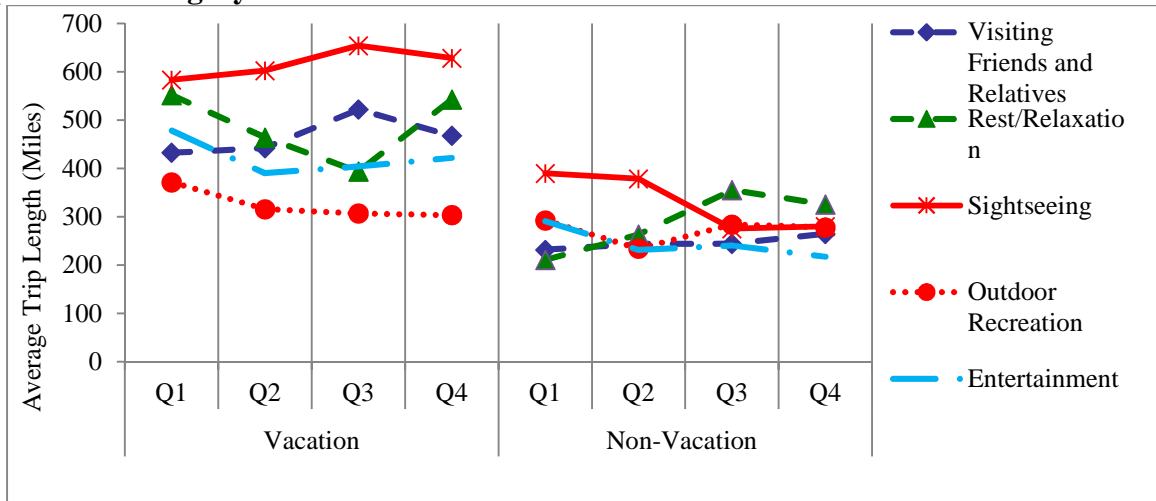


Figure 19 Average trip length on and off vacation by pleasure purpose

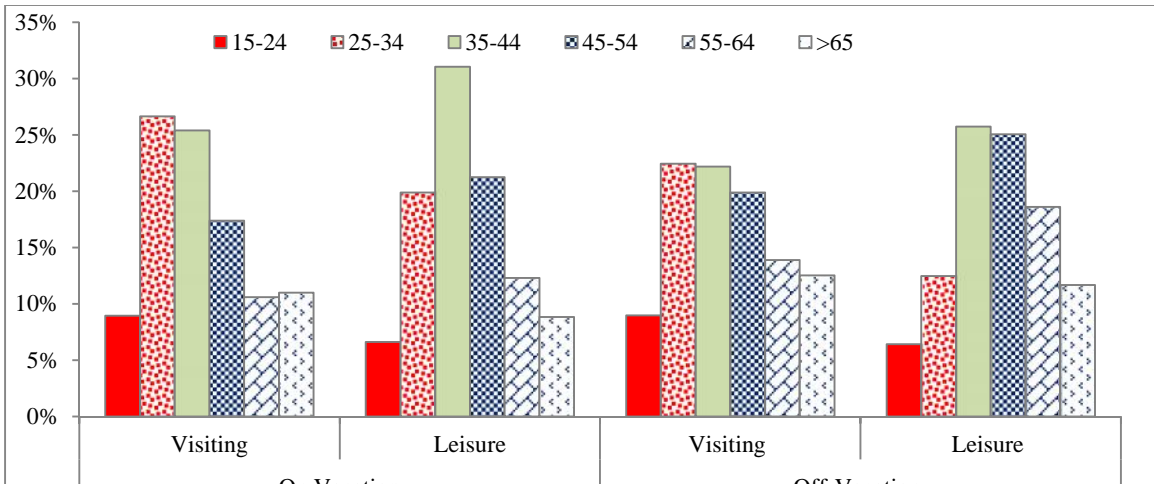


Figure 20 American Travel Survey person pleasure-trip distribution by age

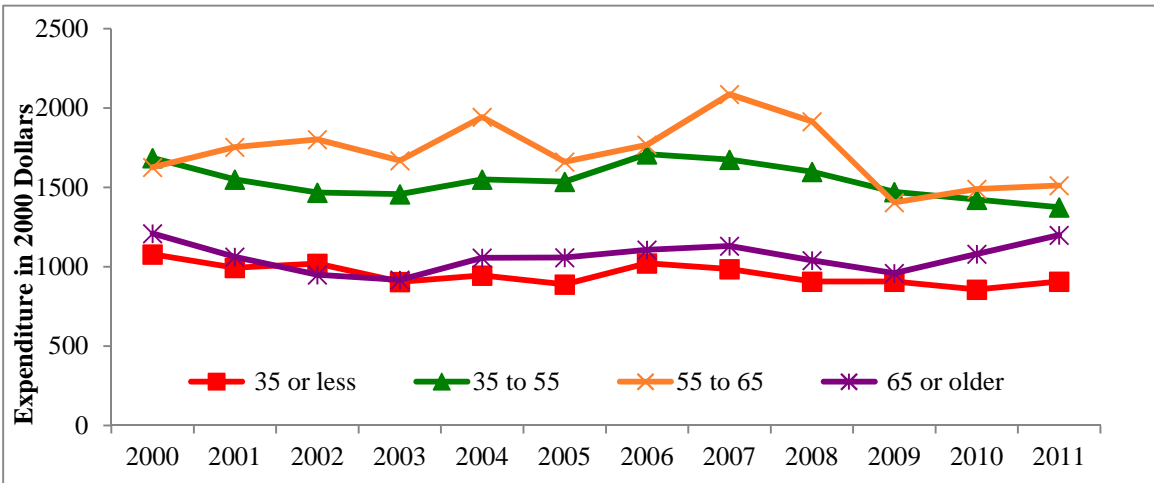


Figure 21 Average CU expenditure on out-of-town trips by Consumer Unit respondent age group

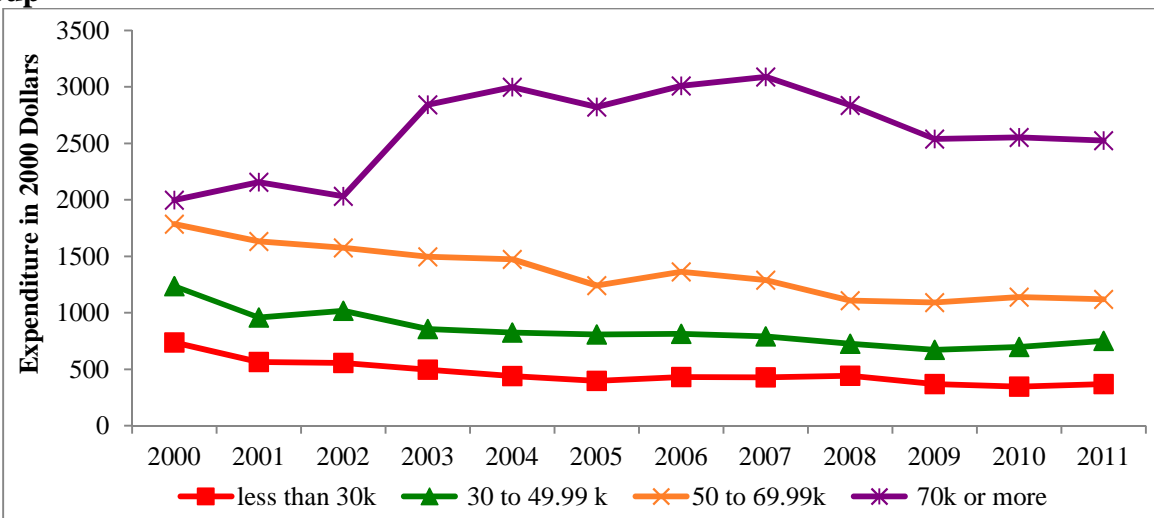


Figure 22 Average Consumer Unit expenditure on out-of-town trips by annual income

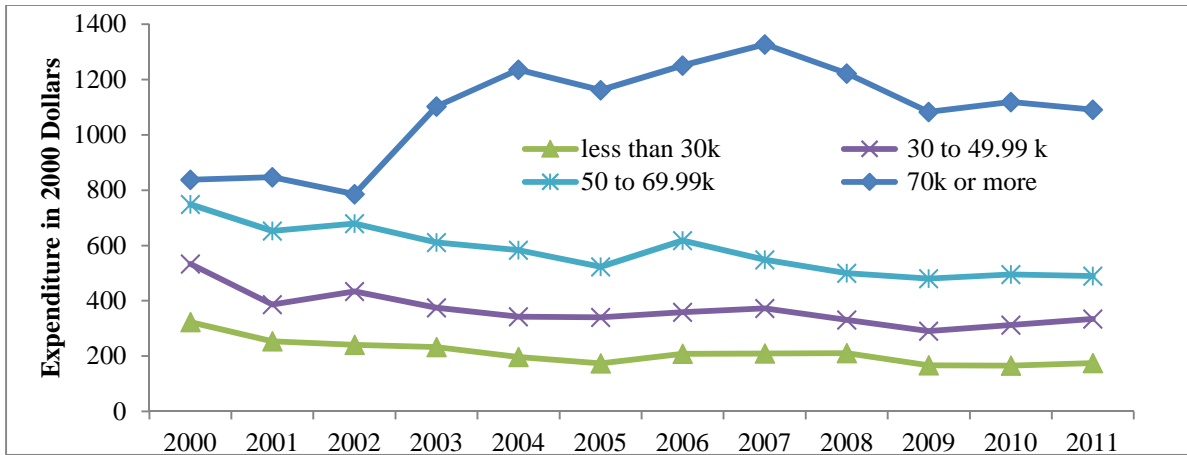


Figure 23 Average Consumer Unit transportation expenditure on out-of-town trips by annual income

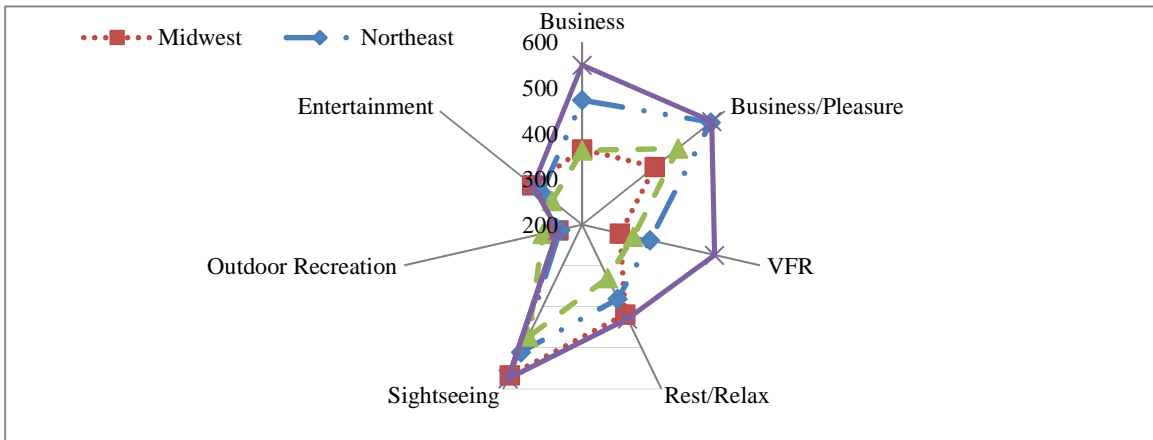


Figure 24 American Travel Survey average trip length by purpose and region

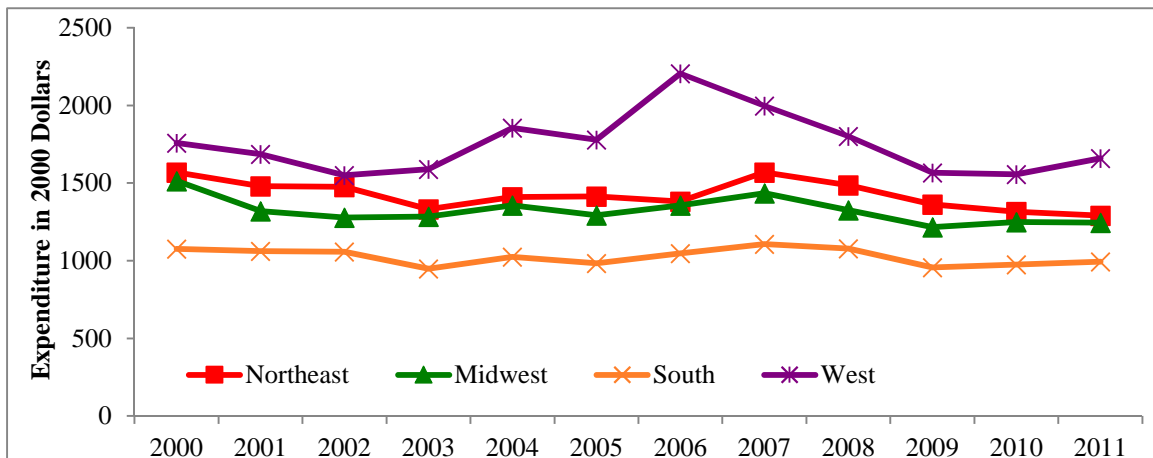


Figure 25 Average annual Consumer Unit expenditure on out-of-town trips by region

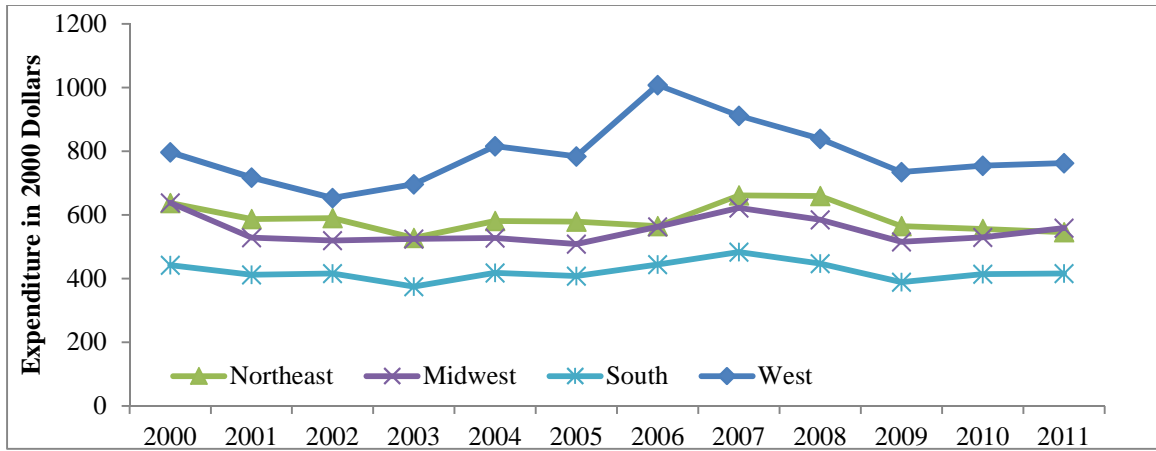


Figure 26 Average Annual CU Expenditure on out-of-town trips by region

Table 1 Travel survey methods comparison

	NPTS (1990,1995)	ATS (1995)	NHTS (2001)
Trip Length	75 mile	100 mile	50 mile
Recall Period	14 days	3 months	28 days, if no trips reported , enquires on trips ending outside this period
Wave Per Annum	1	4 (1 per quarter)	1
Visiting	Visit Friends	Visit Friends and Relatives	Visit Friends and Relatives
Pleasure Enumeration			
Visiting	-Friends	-Friends and Relatives	-Friends and Relatives
Leisure	-Other Socio-Recreational -Vacation	-Sightseeing -Rest/Relaxation -Outdoor Recreation -Entertainment	-Sightseeing -Rest/Relaxation -Outdoor Recreation -Entertainment -Vacation
Vacation Flag	No	Yes	No
Commute	Yes	No	Yes
Data Collected at	Person Level	Household Level	Person Level
Trips reported for	Persons 5 year and over	All Persons	Persons 5 year and over

Note. NPTS = Nationwide Personal Transportation Survey; ATS = American Travel Survey; NHTS = National Household Travel Survey; Source: NPTS (1990, 1995), ATS (1995) and NHTS (2001)

Table 2 Expanded annual trip estimate using nationwide personal transportation survey data collected over 2-week periods

NPTS 1995 : Persons 5 years and older						
Travel period end month and year	Trips estimated for population represented by interviewed sample persons reported	Share of the population represented by the sample persons interviewed	Interviewed population trip rate	Remaining population (persons)	Remaining (non) interviewed populations' estimated trips	Total population trips
May-95	3,409,885	10,821,055	0.32	230,797,239	72,727,855	76,137,741
Jun-95	4,772,008	17,477,947	0.27	224,140,347	61,197,096	65,969,105
Jul-95	6,453,128	19,514,663	0.33	222,103,631	73,445,446	79,898,573
Aug-95	6,532,629	21,761,025	0.30	219,857,269	66,000,846	72,533,475
Sep-95	5,488,027	18,361,599	0.30	223,256,695	66,728,324	72,216,351
Oct-95	5,091,470	21,002,184	0.24	220,616,110	53,483,028	58,574,498
Nov-95	5,470,331	20,123,287	0.27	221,495,007	60,211,386	65,681,717
Dec-95	3,494,161	20,323,502	0.17	221,294,792	38,046,571	41,540,731
Jan-96	4,010,992	20,464,466	0.20	221,153,828	43,345,683	47,356,675
Feb-96	3,072,693	20,764,444	0.15	220,853,850	32,681,644	35,754,338
Mar-96	4,615,724	19,607,372	0.24	222,010,922	52,263,055	56,878,779
Apr-96	5,059,191	20,508,049	0.25	221,110,245	54,546,335	59,605,526
May-96	1,321,880	7,301,288	0.18	234,317,006	42,422,524	43,744,404
Jun-96	869,396	3,587,413	0.24	238,030,881	57,685,878	58,555,274
Annual Total	59,661,517	241,618,294	Annual Total Person Trips (using May and June 1995)			732,147,510

Note. NPTS = Nationwide Personal Transportation Survey. Source: 1995 National Personal Travel Survey, 1995 U.S. Census Population Estimate

Table 3 1995 annual person trips and per capita trip rate estimated using the nationwide personal transportation survey and American travel survey

ATS(1995)—Total trip rate by quarter (5 years and over census population)					
Quarter	Q1	Q2	Q3	Q4	Total
Person trips	149,174,248	199,444,372	239,075,506	187,917,171	775,611,297
Person trip rate (using census population)	0.61	0.82	0.98	0.77	3.19
NPTS (1995)—Person trip rate by quarter (5 years and over census population)					
Quarter	Q1	Q2	Q3	Q4	Total
Person trips	139,989,792	201,712,372	224,648,400	165,796,947	732,147,510
Trip rate (using npts persons)	0.58	0.83	0.93	0.69	3.03
Trip rate (using census population)	0.58	0.83	0.92	0.68	3.01

Note. ATS = American Travel Survey; NPTS = Nationwide Personal Transportation Survey. Source: ATS (1995), NPTS (1995) and U.S. Census (1995)

Table 4 Annual person trips by census region estimated using the nationwide personal transportation survey and American travel survey

ATS (1995)—Total person trips by region					
Region	R1—Northeast	R2—Midwest	R3—South	R4—West	Nation
Trips	116,336,625	202,195,039	282,561,915	174,517,718	775,611,297
Trip rate	2.43	3.51	3.32	3.31	3.19
NPTS (1995)—Total person trips by region					
Region	R1—Northeast	R2—Midwest	R3—South	R4—West	Nation
Annual	107,758,913	189,737,376	290,545,235	142,875,290	730,916,813
Trip rate	2.25	3.29	3.42	2.71	3.00

Note. ATS = American Travel Survey; NPTS = Nationwide Personal Transportation Survey.

Table 5 Annual person (18 and older) trips estimated by quarter and household income

ATS (1995)—Total trips by region and quarter (18 years and over population)					
Quarter	Not Reported	Less than 25 K	25 to Less than 50 K	50 K or more	Total
Q1	—	18,059,821	49,665,760	65,425,870	133,151,451
Q2	—	25,845,769	63,114,474	84,927,572	173,887,815
Q3	—	30,554,256	72,368,020	101,210,173	204,132,449
Q4	—	27,570,935	59,419,577	82,244,869	169,235,381
Annual	—	102,030,781	244,567,831	333,808,484	680,407,096
NPTS (1995) month weights—Total trips by region and quarter (15 years and over population)					
Q1	19,209,710	20,080,640	45,073,873	39,767,863	124,132,086
Q2	24,134,390	31,295,276	61,648,052	62,973,216	180,050,935
Q3	23,860,387	27,238,526	76,404,913	65,256,991	192,760,818
Q4	19,054,017	25,447,673	56,539,258	47,151,084	148,192,031
Annual	86,258,504	104,062,116	239,666,097	215,149,154	645,135,870

Note. ATS = American Travel Survey; NPTS = Nationwide Personal Transportation Survey.

CHAPTER 3: A DISCRETE CONTINUOUS MODEL WITH PERFECT AND IMPERFECT SUBSTITUTES AND PRICE VARIATION

3.1 Motivation

Many consumer choice situations involve decisions of “what to choose” from a set of discrete goods (or alternatives) along with the decisions of “how much to consume” of the chosen good(s). Such discrete-continuous choices are rather common in consumer decisions and of interest in a variety of social sciences, including transportation, economics, and marketing. Examples include households’ choice of vacation destinations and corresponding time allocation, and grocery shoppers’ choice of brand and purchase quantity.

A special case of discrete-continuous choices is the single discrete-continuous (SDC) choice, where consumers choose a single discrete alternative along with the corresponding continuous quantity decision. In such situations, the choice alternatives can be viewed as perfect substitutes where the choice of one alternative precludes the choice of other alternatives. Another case is when the choice alternatives are imperfect substitutes where the choice of one alternative does not necessarily preclude the choice of other alternatives. In the context of this dissertation, a households’ mode choice to travel to a particular destination are perfect substitutes. While a households decision to travel to one or more destination(s) constitute imperfect substitutes.

In such situations, consumers can potentially choose multiple discrete alternatives (destinations), along with the corresponding continuous quantity (duration at destination) as choice decisions. For example, a household might choose to visit multiple vacation destinations

over a given time frame. Similarly, a grocery shopper might choose a variety of brands of a product, as opposed to a single brand. Such multiple discrete-continuous (MDC) choices are being increasingly recognized and modeled in the recent literature.

A more general case of discrete-continuous choices includes both SDC and MDC choices, where consumers choose at most a single alternative from a subset of alternatives and potentially multiple alternatives from the remaining alternatives. Such situations arise when the choice set comprises a mix of both perfect substitutes (from which no more than a single alternative could be consumed) and imperfect substitutes (from which potentially multiple alternatives could be consumed). As reviewed in the next section, the vast majority of choice modeling literature has focused on analyzing SDC choices, while there recently has been growing interest in analyzing MDC choices. Not much exists in the literature on modeling consumer behavior involving both SDC and MDC choices from choice sets that comprise a mix of perfect and imperfect substitutes.

To fill this gap, this dissertation formulates a unified random utility maximization (RUM) framework that can be used as a joint MDC-SDC modeling framework to analyze discrete-continuous choices from a combination of perfect and imperfect substitutable choice alternatives. In addition to the RUM formulation, a procedure to apply the proposed framework for forecasting purposes is presented. As importantly, the formulation is as well extended to accommodate multiple linear budget constraints, as opposed to a single budget constraint.

3.2 Random Utility-Maximization Models

A variety of approaches have been used in the literature to model discrete-continuous choices (see Bhat, 2008 for a review). Among these, a particularly attractive approach is based

on the classical microeconomic consumer theory of RUM.¹ Specifically, consumers are assumed to optimize a direct utility function $U(\mathbf{x})$ over a bundle of nonnegative consumption quantities $\mathbf{x} = (x_1, \dots, x_k, \dots, x_K)$ subject to a budget constraint, as:

$$\text{Max } U(\mathbf{x}) \text{ such that } \mathbf{x} \cdot \mathbf{p} = E \text{ and } x_k \geq 0 \forall k = 1, 2, \dots, K \quad (4)$$

In Equation 4 above, $U(\mathbf{x})$ is a quasiconcave, increasing and continuously differentiable utility function with respect to the consumption quantity vector \mathbf{x} , \mathbf{p} is the vector of unit prices for all goods, and E is a budget for total expenditure. Geometrically, the optimal consumption bundle for this utility maximization problem is the point in consumption space where the budget line meets (tangentially) the indifference curves corresponding to the utility function. The form of the utility function governs the characteristics of the optimal consumption bundle. If $U(\mathbf{x})$ is such that the marginal utility of consumption at the point of zero consumption (i.e., baseline marginal utility) is either infinity or indeterminate for each and every good, it results in interior solution (positive consumption) for all goods.

In this case, the indifference curves are asymptotic to the consumption axes precluding the possibility of corner solutions (zero consumption). On the other hand, goods with a finite value of baseline marginal utility result in indifference curves that meet the consumption axes with a finite slope and allow corner solutions. Within the context of corner solutions, again the functional form of $U(\mathbf{x})$ determines whether the formulation corresponds to an SDC case or an

¹ Another approach is to employ a multivariate statistical system, where separate equations are used to model each component of choice—with statistical correlations between the random components of different equations. The approach has been widely used to model SDC choices by *statistically* joining a discrete choice equation with a continuous regression equation in a bivariate statistical system or in a switching regression formulation (Amemiya, 1974; Heckman, 1974, 1979; L.-F. Lee, 1983; Maddala, 1983). However, such a reduced form multivariate statistical system is typically not based on an underlying economic theory. Besides, although the approach has been used to model multiple discrete/discrete-continuous choices (Edwards & Allenby, 2003; Fang, 2008; Manchanda, Ansari, & Gupta, 1999; Srinivasan and Bhat, 2006), it becomes cumbersome to statistically tie all equations for discrete and continuous components into a multivariate system for more than a modest number of choice alternatives.

MDC case. A linear utility function with respect to consumption results in the SDC case where the choice alternatives are perfect substitutes (Deaton and Muellbauer, 1980; page 262) in that a utility maximizing consumer chooses only one alternative. Linear utility forms result in linear indifference curves that intersect the budget line only on a consumption axis—hence the optimal consumption bundle is a corner solution in which only one alternative is chosen. On the other hand, a nonlinear utility form that allows for diminishing marginal utility and corner solutions results in the MDC case where the choice alternatives are treated as imperfect substitutes allowing for the possibility of “multiple discreteness”.

A notable contribution to the analysis of SDC choices is the work by Hanemann (1984), who proposed a general class of utility functions that assures perfect substitutability among choice alternatives. An example of his utility forms is a bivariate function given by:

$$U(\mathbf{x}) = u\left(x_1, \sum_{k=2}^K \mathbb{E}_k x_k\right)$$
, where the first good x_1 is as an outside good² and the other goods k ($k = 2, \dots, K$) are inside goods, each with its own quality index \mathbb{E}_k ($k = 2, \dots, K$). As can be observed, this functional form is linear with respect to the consumption of different inside goods.

The above functional form ensures that, in addition to the outside good, only one of the inside goods ($k = 2, 3, \dots, K$) is consumed—hence perfect substitutability among inside goods. To derive the demand functions implied by this utility function, Hanemann used an indirect utility approach which involves solving the dual of the optimization problem in equation 4.

Specifically, the analysis starts with the specification of a conditional indirect utility function and makes use of “virtual prices” to determine the discrete choice and Roy’s identity to derive the

² The outside good represents a composite of all goods other than the $K-1$ inside goods of interest to the analyst. The presence of the outside good helps in ensuring that the budget constraint is binding. A typical assumption is that the prices and characteristics of the goods grouped into the outside category do not influence the choice and resource allocation among the inside goods (see Deaton and Muellbauer, 1980).

Marshallian demand functions (also see L.-F. Lee & Pitt, 1986). Among other studies that adopt this approach are Chiang (1991) and Chintagunta (1993), who extended Hanemann's formulation to include the possibility of no inside goods being selected.³

While most literature in the area of discrete-continuous choice analysis has been geared toward SDC choices, the past decade has seen a surge of interest in analyzing MDC choices. The indirect utility approach, however, is difficult to use for analyzing MDC choices. An alternative approach, due to Hanemann (1978) and Wales and Woodland (1983), is to employ the Karush-Kuhn-Tucker (KKT) conditions of optimality to solve the utility maximization problem in Equation 4 (also see Bockstael, Strand, & Hanemann, 1987). Specifically, a randomly distributed (over the population) utility function results in randomly distributed KKT conditions, which in turn form the basis for deriving the likelihood expressions for consumption patterns.

Recent years have witnessed significant developments in the use of this KKT approach for analyzing MDC choices in the fields of environmental economics (Kuriyama, Shoji, & Tsuge, 2011; Phaneuf, Kling, & Herriges, 2000; Phaneuf & Smith, 2005; von Haefen, Phaneuf, & Parsons, 2004), marketing (Kim, Allenby, & Rossi, 2002; Satomura, Kim, & Allenby, 2011) and transportation (Bhat, 2005, 2008). Notable among these is Bhat's (2005, 2008) MDCEV model. The MDCEV model is based on a Box-Cox transformation of a translated constant elasticity of substitution utility function that subsumes many other nonlinear utility forms proposed in the literature as special cases, and enables a clear interpretation of the structural parameters.

Further the specification of Type-1 extreme value distributed random utility components in the MDCEV model leads to simple closed form likelihood expressions making it easy to

³ Also see Dubin and McFadden (1984) for a slightly different approach for SDC choice analysis, which begins with the Marshallian demand functions and utilizes Roy's identity to solve for the implied indirect utility function.

estimate the structural parameters. Thanks to these advances, KKT-based models are being increasingly used to analyze a variety of MDC choices relevant to transportation planning, including individuals' activity participation and time-use (Bhat, 2005; Habib & Miller, 2008) and vehicle ownership and usage (Ahn, Jeong, & Kim, 2008; Bhat, Sen, & Eluru, 2009; Jaggi, Weis, & Axhausen, 2013).

On the methodological front, as reviewed in Pinjari and Sivaraman (2013), recent literature in this area has started to enhance the basic formulation in Equation 4 along different directions, including: (a) toward more flexible, nonadditively separable functional forms for the utility specification (Bhat, Castro, & Pinjari, 2013; S. Lee, Kim, & Allenby, 2010; Vasquez-Lavin & Hanemann, 2009), (b) toward greater flexibility in the specification of the constraints faced by the consumer (such as the consideration of multiple linear budget constraints; Castro, Bhat, Pendyala, & Jara-Diaz, 2012; Parizat & Shachar, 2010; Satomura et al., 2011;) and (c) toward more flexible stochastic specifications for the random utility functions (Bhat, Castro, & Khan, 2013; Pinjari, 2011; Pinjari & Bhat, 2011).

Despite the above advances, it is worth noting that most literature in this area has focused on analyzing choice situations that fall into either the SDC case or the MDC case. However, as discussed earlier, several choice situations can potentially include both SDC and MDC choices, as a result of choice from a combination of perfect and imperfect substitutable alternatives in the choice set. To the authors' knowledge, only a few recent studies by Bhat and colleagues (Bhat et al., 2009; Bhat, Srinivasan, & Sen, 2006; Eluru, Pinjari, Pendyala, & Bhat, 2010) use joint MDC-SDC modeling frameworks for such choice situations. However, the model formulations in these studies are critically hinged on the assumption that prices per unit consumption do not vary

across the perfect substitutes. Specifically, their model formulations remain consistent with utility maximization only when there is no price variation across the perfect substitutes.

In this chapter, we formulate a unified random utility maximization framework that can accommodate perfect and imperfect substitutes regardless of the presence or absence of price variation. As discussed in the next section, the key to this approach is a utility form that is nonlinear with respect to consumption across different imperfectly substitutable alternatives but linear with respect to consumption across perfectly substitutable alternatives. In addition to the RUM formulation, we present a procedure to apply the proposed framework for forecasting purposes. Furthermore, building on recent literature (Satomura et al., 2011) we demonstrate how the formulation can be extended to accommodate multiple linear budget constraints, as opposed to a single budget constraint.

3.3 Model Formulation

Let $j (=1,2,3,\dots,J)$ be the index to represent the vacation destination alternatives available to households, $l (=1,2,\dots,L)$ be the index to represent the travel mode alternatives, and jl be the index to represent a vacation destination and travel mode combination. Let

$\mathbf{t} = (t_1, t_2, \dots, t_j, \dots, t_J)$ be the vector of vacation time allocations by a household to each of the destination alternatives. Considering that one can travel to a destination by any of the available modes, one can expand each element t_j of \mathbf{t} as a subvector $(t_{j1}, t_{j2}, \dots, t_{jL})$ representing the vacation time allocation to destination j reached by each of the available travel modes. In the subsequent discourse, we will assume that $L = 2$ (i.e., only two modes available for traveling to any destination). Thus, \mathbf{t} can be expressed as $\mathbf{t} = ((t_{11}, t_{12}), (t_{21}, t_{22}), \dots, (t_{j1}, t_{j2}), \dots, (t_{J1}, t_{J2}))$ but can also be generalized for $L > 2$.

Over the time frame of a year, a household may choose to visit none, one, or more destinations (although not necessarily all destinations). Thus, one can expect the data to exhibit imperfect substitution (hence, multiple discreteness) among destination choice alternatives. For the chosen destinations, however, households are observed to travel by a single mode of travel regardless of the number of times they visited the destination. Thus, if a destination j is visited, the entire time t_j allocated for the destination would be allocated to only one element in the time-allocation subvector (t_{j1}, t_{j2}) for that destination while the time allocation to the other element would be zero, exhibiting perfect substitution (hence, single discreteness) among mode choice alternatives. In this formulation, for ease in notation without losing generality, the first modal alternative is assumed to be chosen for any chosen destination (i.e., $t_{j1} = t_j$ and $t_{j2} = 0$, if $t_j > 0$).

3.4 Utility Form

To model a household's vacation destination and mode choices over an annum, consider the following utility function (the subscript for the household is suppressed for simplicity):

$$U(\mathbf{t}, e_0) = \sum_{j=1}^J x_j \ln \left\{ \frac{(\mathbb{E}_{j1} t_{j1} + \mathbb{E}_{j2} t_{j2})}{x_j} + 1 \right\} + \mathbb{E}_0 \ln e_0; \\ \mathbb{E}_{jl} > 0, \mathbb{E}_0 > 0, x_j > 0 \quad \forall j = 1, 2, \dots, J \text{ and } l = 1, 2 \quad (5)$$

In the above utility function (Equation 5), the first term represents the utility accrued due to vacation. Specifically, the term $x_j \ln \left\{ (\mathbb{E}_{j1} t_{j1} + \mathbb{E}_{j2} t_{j2}) / x_j + 1 \right\}$ is a subutility function representing the utility U_j accrued due to consuming $t_j (= t_{j1} + t_{j2})$ amount of time at a vacation destination j . Note that utility is assumed to be additively separable across destinations in that the total utility from vacation over the time frame of a year is the sum of utility accrued from the

time spent at all the vacation destinations $j (= 1, 2, \dots, J)$ over the year. For each destination j , however, the functional form is not separable in the time allocations, t_{j1} and t_{j2} , by the different modes of travel. Further, the functional form exhibits the property of weak complementarity (Herriges, Kling, & Phaneuf, 2004; Mäler, 1974; von Haefen, 2007) that the utility contribution from a destination alternative is zero if that destination is not visited (i.e., if the time allocated to the destination is zero).

The second term in (5), $\mathbb{E}_0 \ln e_0$ completes the utility function to form an incomplete demand system. Specifically, e_0 is the Hicksian composite outside good representing annual expenditure for all purposes other than vacation (i.e., income—annual expenditure on vacation). This outside good is assumed to be “essential” with some positive consumption by all households. The presence of this term recognizes that only a part of the available annual budget (e.g., annual income) is spent on vacation.

Households are assumed to allocate the annual income (E) available to them to maximize the utility in Equation 5 subject to the following constraint:

$$\sum_{j=1}^J (p_{j1}t_{j1} + p_{j2}t_{j2}) + e_0 = E \quad (6)$$

In the above equation, p_{jl} represents the money-price of consuming unit time of a vacation destination j traveled by mode l ($l = 1, 2$). These prices accommodate two components—(a) destination costs that do not depend on the mode of travel (e.g., lodging, dining, and entertainment costs) and (b) travel costs that depend on travel mode—with the latter leading to the difference between the prices by different modes of travel (l) for a same destination j . The Hicksian composite outside good is assumed to be a numeraire with unit money-price.

3.5 Interpretation of Parameters

To interpret the role of \mathbb{E}_{jl} ($j=1,2,\dots,J; l=1,2$), consider the marginal utility of the

utility function in (5) with respect to t_{jl} , $\frac{\partial U(\mathbf{t}, e_0)}{\partial t_{jl}} = \frac{\mathbb{E}_{jl}}{\left\{ \frac{(\mathbb{E}_{j1}t_{j1} + \mathbb{E}_{j2}t_{j2})}{\chi_j} + 1 \right\}}$. It can be observed

that $\frac{\partial U(\mathbf{t}, e_0)}{\partial t_{jl}} = \mathbb{E}_{jl}$ when $t_{j1} = 0$ and $t_{j2} = 0$. Thus, \mathbb{E}_{jl} can be interpreted as the marginal

utility for destination-mode combination alternative jl at the point of zero time allocation to all destination-mode combination alternatives that share the same destination j (in short, baseline marginal utility for destination-mode combination alternative jl).⁴

This chapter further shows that an alternative with a greater price-normalized baseline marginal utility (*i.e.*, \mathbb{E}_{jl}/p_{jl}) is more likely to be chosen than other alternatives sharing the same destination. In other words, a greater value of \mathbb{E}_{jl}/p_{jl} implies a lower likelihood of corner solution for the alternative. In addition, \mathbb{E}_{jl} also plays a role in determining the continuous consumption quantity t_{jl} for alternative jl , if that alternative is chosen. Specifically, between two different chosen alternatives with equal values of χ_j , the alternative with a greater value of baseline marginal utility will have a greater amount of consumption.

The primary role of χ_j , as discussed in Kim et al. (2002) and Bhat (2008), is to translate the utility function so that the indifference curves become asymptotic to the consumption axes at $(-\chi_1, -\chi_2, \dots, -\chi_J)$. Consequently, the indifference curves strike the consumption axes (in the

⁴ Note, however, that \mathbb{E}_0 cannot be interpreted as the baseline marginal utility for the outside good, since $\left[\frac{\partial U(\mathbf{t}, e_0)}{\partial e_0} \right]_{e_0=0}$ is not \mathbb{E}_0 .

positive orthant as long as $x_j > 0$) with a finite slope and result in a possibility of corner solutions (i.e., zero consumption). In addition to allowing corner solutions, as explained in Bhat (2008), differences in the x_j terms allow differential rates of satiation (i.e., diminishing marginal utility) across different destinations. Specifically, all else being the same, a destination alternative with a greater x_j value exhibits a slower rate of satiation (hence greater amount of consumption) than those with smaller values of x_j .

3.6 Perfect and Imperfect Substitutes

The utility form in Equation 5 is nonlinear with respect to time allocations (t_{jl}, t_{kl}) across different destinations (j, k) , regardless of the mode (l) used to travel to the destinations. As discussed in many previous studies (e.g., Kim et al., 2002; Bhat, 2005), the nonlinear form allows diminishing marginal utility to accommodate the possibility of multiple destinations being chosen (i.e., imperfect substitution across destination choice alternatives). However, the utility form is linear with respect to time allocation across different modes of travel available for a same destination (t_{j1}, t_{j2}) . The linear form ensures perfect substitution across mode choice alternatives for a destination in that a utility maximizing consumer chooses only one mode of travel to any chosen destination. To verify this, one can examine the marginal rate of substitution (MRS) between time allocations to different choice alternatives. The MRS between time allocations for destination-mode alternatives across different destinations $(t_{jl}, t_{kl}; j \neq k)$ is given by:

$$MRS(t_{jl}, t_{kl}) = \frac{\partial U(\mathbf{t}, e_0)}{\partial t_{jl}} \bigg/ \frac{\partial U(\mathbf{t}, e_0)}{\partial t_{kl}} = \frac{\frac{\mathbb{E}_{jl}}{\left\{ \frac{(\mathbb{E}_{j1}t_{j1} + \mathbb{E}_{j2}t_{j2})}{x_j} + 1 \right\}}}{\frac{\mathbb{E}_{kl}}{\left\{ \frac{(\mathbb{E}_{k1}t_{k1} + \mathbb{E}_{k2}t_{k2})}{x_k} + 1 \right\}}} \quad (7)$$

As can be observed, the MRS between destination-mode alternatives that do not share a same destination is not a constant and dependent on the time allocations to the two destination alternatives. This leads to nonlinear indifference curves, implying imperfect substitution (hence multiple discreteness) across different destination choice alternatives.

On the other hand, The MRS between time allocations for destination-mode alternatives that share the same destination (t_{j1}, t_{j2}) is a constant and independent of the time allocations to the two alternatives, as given below:

$$MRS(t_{j1}, t_{j2}) = \frac{\partial U(\mathbf{t}, e_0)}{\partial t_{j1}} \bigg/ \frac{\partial U(\mathbf{t}, e_0)}{\partial t_{j2}} = \mathbb{E}_{j1} / \mathbb{E}_{j2} \quad (8)$$

Therefore, the indifference curves between such alternatives are straight lines with a slope of $-\mathbb{E}_{j1} / \mathbb{E}_{j2}$. Such goods whose indifference curves are straight lines are called perfect substitutes in the microeconomic literature (see Deaton & Muellbauer, 1980, p. 262) since the choice of one good by a utility maximizing consumer precludes the choice of others by the same consumer.

In summary, while the utility form allows imperfect substitution across destination choice alternatives (a household can potentially choose multiple destinations over an annum), it also ensures perfect substitution across different mode choice alternatives for a destination (only one mode is chosen to travel to a destination). This feature arises automatically from the utility form regardless of the presence or absence of price variation, eliminating the need for an explicit constraint to impose perfect substitutability among mode choice alternatives.

3.7 Karush Kuhn Tucker Conditions of Optimality

To solve the household's utility maximization problem subject to the budget constraint, one can form the Lagrangian as below, and derive KKT conditions of optimality:

$$L = \mathbb{E}_0 \ln e_0 + \sum_{j=1}^J x_j \ln \left\{ \frac{(\mathbb{E}_{j1} t_{j1} + \mathbb{E}_{j2} t_{j2})}{x_j} + 1 \right\} - \left[e_0 + \sum_{j=1}^J (p_{j1} t_{j1} + p_{j2} t_{j2}) - E \right] \quad (9)$$

The KKT condition for the numeraire outside good is below:

$$\frac{\partial L}{\partial e_0} = 0 \text{ since } e_0 > 0, \text{ or } \lambda = \frac{\mathbb{E}_0}{e_0}. \quad (10)$$

The KKT conditions of optimality for the inside goods are given next:

$$\begin{aligned} \frac{\partial L}{\partial t_{jl}} &= \lambda p_{jl}, \text{ if } t_{jl} > 0; \forall j = 1, 2, \dots, J; l = 1, 2 \\ \frac{\partial L}{\partial t_{jl}} &< \lambda p_{jl}, \text{ if } t_{jl} = 0; \forall j = 1, 2, \dots, J; l = 1, 2 \end{aligned}$$

or,

$$\begin{aligned} \frac{(\mathbb{E}_{jl} / p_{jl})}{\left\{ \frac{(\mathbb{E}_{j1} t_{j1} + \mathbb{E}_{j2} t_{j2})}{x_j} + 1 \right\}} &= \lambda \text{ if } t_{jl} > 0; \forall j = 1, 2, \dots, J; l = 1, 2 \\ \frac{(\mathbb{E}_{jl} / p_{jl})}{\left\{ \frac{(\mathbb{E}_{j1} t_{j1} + \mathbb{E}_{j2} t_{j2})}{x_j} + 1 \right\}} &< \lambda \text{ if } t_{jl} = 0; \forall j = 1, 2, \dots, J; l = 1, 2 \end{aligned} \quad (11)$$

The expression in the left side of the above two conditions represents the price-normalized marginal utility of time consumption for destination-mode combination jl ($l=1,2$).

We now expand on these KKT conditions for two different cases: (a) for chosen destinations, and (b) for unchosen destinations.

3.7.1 KKT Conditions for Chosen Destinations

If a destination j is chosen, as discussed before, only one of the modal alternatives will be chosen to travel to that destination. Without loss of generality, assume that the first mode is chosen (i.e., $t_{j1} > 0$ and $t_{j2} = 0$). The resulting KKT conditions are as below:

$$\frac{\left(\frac{\mathbb{E}_{j1}}{p_{j1}}\right)}{\left(\frac{\mathbb{E}_{j1}t_{j1}}{x_j} + 1\right)} = \} \text{ since } t_{j1} > 0 \text{ and } t_{j2} = 0; \forall j \in \text{chosen destination alternatives}$$

and

$$\frac{\left(\frac{\mathbb{E}_{j2}}{p_{j2}}\right)}{\left(\frac{\mathbb{E}_{j1}t_{j1}}{x_j} + 1\right)} < \} \text{ since } t_{j1} > 0 \text{ and } t_{j2} = 0; \forall j \in \text{chosen destination alternatives} \quad (12)$$

The first of the above KKT conditions can be substituted into the second condition to result in the following condition:

$$\frac{\mathbb{E}_{j2}}{p_{j2}} < \frac{\mathbb{E}_{j1}}{p_{j1}} \text{ since } t_{j1} > 0 \text{ and } t_{j2} = 0; \forall j \in \text{chosen destination alternatives} \quad (13)$$

The above condition implies that the price-normalized baseline utility of an unchosen mode to a chosen destination is always less than the price-normalized baseline utility of the chosen mode. Therefore, for a chosen destination, the mode choice alternative with the greatest price-normalized baseline utility would be the chosen alternative. Given this, the KKT conditions in Equation 12 can be rewritten as:

$$\frac{\left(\frac{\mathbb{E}_{j1}}{p_{j1}}\right)}{\left(\frac{\mathbb{E}_{j1}t_{j1}}{x_j} + 1\right)} = \} \text{ since } t_{j1} > 0 \text{ and } t_{j2} = 0; \forall j \in \text{chosen destination alternatives}$$

and

$$\frac{\mathbb{E}_{j2}}{p_{j2}} < \frac{\mathbb{E}_{j1}}{p_{j1}} \text{ since } t_{j1} > 0 \text{ and } t_{j2} = 0; \forall j \in \text{chosen destination alternatives} \quad (14)$$

Substituting \mathbb{E}_0/e_0 for λ , making algebraic rearrangements, and taking logarithms on both sides, one can further rewrite the KKT conditions for chosen destination alternatives as:

$$\ln(\mathbb{E}_{j1}) = -\ln\left(\frac{e_0}{\mathbb{E}_0 P_{j1}} - \frac{t_{j1}}{x_j}\right); \forall j \in \text{chosen destination alternatives}$$

and

$$\ln(\mathbb{E}_{j2}) - \ln(p_{j2}) < \ln(\mathbb{E}_{j1}) - \ln(p_{j1}) \forall j \in \text{chosen destination alternatives} \quad (15)$$

Note from the first of the above two conditions that the term $\left(\frac{e_0}{\mathbb{E}_0 P_{j1}} - \frac{t_{j1}}{x_j}\right)$ is inside a logarithmic function. During model estimation, the analyst must ensure that this term is always positive, lest it should lead to estimation breakdowns.

3.7.2 KKT Conditions for Unchosen Destinations

For a destination k that is not chosen, the time allocation would be zero for both the destination-mode alternatives corresponding to that destination (i.e., $t_{k1} = 0$ and $t_{k2} = 0$), resulting in the following KKT conditions for the destination:

$$\mathbb{E}_{k1} - \lambda_{k1} < 0, \text{ and}$$

$$\mathbb{E}_{k2} - \lambda_{k2} < 0; \text{ since } t_{k1} = 0 \text{ and } t_{k2} = 0; \forall k \in \text{non-chosen destination alternatives} \quad (16)$$

Substituting \mathbb{E}_0/e_0 for λ and taking logarithms on both sides, one can rewrite the KKT conditions for unchosen destination alternatives as:

$$\ln(\mathbb{E}_{k1}) < \ln(\mathbb{E}_0) + \ln(p_{k1}) - \ln(e_0), \text{ and}$$

$$\ln(\mathbb{E}_{k1}) < \ln(\mathbb{E}_0) + \ln(p_{k1}) - \ln(e_0); \forall k \in \text{non-chosen destination alternatives}$$

$$(17)$$

3.8 Economic Structure

To complete the model specification, we define the baseline marginal utility for destination-mode combination alternatives, \mathbb{E}_{jl} as a function of observed and unobserved household characteristics, destination (j) characteristics, and modal (jl) characteristics as: $\mathbb{E}_{jl} = \exp(\mathbf{z}_{jl}'\mathbf{w}_0 + v_{jl})$, where \mathbf{z}_{jl} is a vector of destination and modal characteristics influencing the household's destination and mode choices (and their interactions with household characteristics); \mathbf{w}_0 is a corresponding vector of parameters; and v_{jl} is a destination-mode specific random term to accommodate the unobserved factors influencing the choice of destination-mode alternative jl. For \mathbf{w}_0 , two different specifications have been used in the literature, arising from identification issues with use of linear budget constraint.

3.8.1 Identification Issues

This is explained in this section in the context of multiple budget constraints for four different utility scenarios as illustrated in Figure 27. First, as discussed in Castro et al. (2012), it is difficult to estimate alternative-specific parameters of explanatory variables separately on the baseline utility parameters for inside goods (i.e., \mathbb{E}_{jl}) as well as those for outside goods (i.e., α_0 and (or) \mathbf{w}_0). Thus, for each explanatory variable, the corresponding coefficients in the baseline utility parameters for at least as many goods as the number of budget constraints need to be normalized (to zero). This is because the consumption amounts of any (one) two goods can be derived from the consumptions of the remaining $(K-1) K-2$ (a good(s) and (single) two budget constraints. In the current empirical context, given the inside goods (vacation destinations) are of greater interest, the baseline utility parameters of the (single) two outside goods are the natural candidates for such normalization.

Second, for the same reasons just discussed, it is difficult to identify the baseline utility parameters \mathbb{E}_0 and(or) w_0 separately. Thus, another important normalization is to set the baseline utility parameters of all the constraint-specific outside goods as equal. In the current empirical context for multiple budget constraints $\mathbb{E}_0 = w_0$. While the former normalization has been discussed in Castro et al. (2012) and is applicable regardless of the presence/absence and the nature of the outside goods, the latter normalization is equally important in situations with constraint-specific Hicksian composite outside goods.⁵ Neglecting this normalization can potentially lead to severe estimation problems.

To better explain this, Figure 27 illustrates the identification issues arising in absence of such normalization for the following utility expression with one inside good t_{in} and (multiple) two constraint-specific outside goods t_0 and e_0 . This in concept is applicable to single constraint models as well:

$$U(t_{in}, t_0, e_0) = \mathbb{E}_{in} x_{in} \ln((t_{in} / x_{in}) + 1) + (\mathbb{E}_0 / r) t_0^r + (w_0 / \dots) e_0^w \quad (18)$$

subject to the constraints: $t_{in} + t_0 = T$ and $p_{in} t_{in} + e_0 = E$, where T is the time budget set to 365, E is the money budget set to 100, p_{in} is the price for the inside good set to 0.2, \mathbb{E}_{in} is the baseline utility parameter for the inside good set to 5, x_{in} is the translation parameter for the inside good set to 0.64.

The Figure 27 shows the profiles for the total utility $U(t_{in}, t_0, e_0)$ as a function of the consumption of inside good (t_{in}) for four different cases with distinct sets of values for the

⁵ These two normalizations are overlapping but not equivalent. The former normalization ensures that the deterministic components of \mathbb{E}_0 and w_0 are equal but the random terms V_0 and $<_0$ are still different, which leads to a double integral for the likelihood function as in Castro et al. (2012). On the other hand, the latter normalization does not ensure that the coefficients of explanatory variables on the baseline utilities of the outside goods have to be normalized to zero.

baseline utility parameters of the outside goods, \mathbb{E}_0 and w_0 . The r and \dots parameters are both fixed to zero in the first three cases, while they take the values of 0.1 and 0.4, respectively in the fourth case. All other parameters ($T, E, p_{in}, \mathbb{E}_{in}, x_{in}$) are the same in all four cases.

As can be observed from the first three cases, the utility curves for all cases follow each other closely, reflecting that different sets of baseline utility parameters for the outside goods can result in a similar utility profile. In addition, the optimal consumption values for the inside good (i.e., the t_{in} values where the utility curves peak) are very close to each other, if not exactly the same, in all three cases.

Note that $\mathbb{E}_0 = w_0$ in case 3, suggesting that, keeping all else same, a set of different \mathbb{E}_0 and w_0 parameters can be replaced with a single value while retaining a similar utility profile and optimal consumptions. These results point to the difficulty of identifying \mathbb{E}_0 and w_0 separately and suggest the need for setting $\mathbb{E}_0 = w_0$.

Third, the satiation parameters for the outside goods, r and \dots are also difficult to estimate, because the quantities of the Hicksian composite outside goods t_0 and e_0 can be determined based on the consumption quantities of the inside goods and the time and money budget constraint identities. Thus, the analyst may have to impose normalizations on the satiation parameters of the constraint-specific outside goods, such as $r = 0$ and $\dots = 0$, for parameter identification and stability in estimation.

This can be observed in Figure 27 from the utility profile in case 4 (with $r = 0.1$ and $\dots = 0.4$), which is empirically indistinguishable from the other utility profiles (with r and \dots set to zero). Further, the utility profile with the most parsimonious specification is in case 3, with $\mathbb{E}_0 = w_0$, $r = 0$ and $\dots = 0$. This suggests that, keeping all else same, a utility profile with

different values for \mathbb{E}_0 and w_0 and nonzero values for r and ... can be replaced with another utility profile with a single value for the \mathbb{E}_0 and w_0 parameters and zero for r and ...⁶ Note that all these normalizations (based on the parsimonious specification in case 3 of Figure 27 are specific to utility formulations with constraint-specific Hicksian outside goods.

Further, the approach adopted by a large number of studies in environmental economics (e.g., von Haefen et al., 2004) and recently in the marketing literature (Satomura et al., 2011), is to normalize \mathbb{E}_0 to a value of 1. The reason for this normalization is that the KKT conditions are sufficient for estimating the expenditures on all but one good in the specification (e.g., the expenditures on all inside goods but the outside good). Given the expenditures on the inside goods, the outside good expenditure can be determined from the budget constraint identity. The second approach is to normalize only the deterministic component of \mathbb{E}_0 to 1 and to specify a random component, as: $\mathbb{E}_0 = \exp(v_0)$, where v_0 is a random error term capturing the unobserved factors influencing the total expenditure allocation for all inside goods (annual vacation, in this case).

Bhat (2008) discusses several reasons why the latter specification should be preferred. A particular advantage with the latter specification is that including the stochastic term v_1 on the outside good helps in capturing correlation among the random utilities of the inside goods. Such correlation helps in inducing greater competition among the consumptions of the inside goods, when compared to the competition between the inside goods and the outside good. However, the former specification is not theoretically inappropriate, especially on a Hicksian composite

⁶ Additional exercises with different sets of parameters in the above-discussed 3-good utility function suggested the same identification issues. Further experiments with 4-good utility functions (with two-dimensional utility surfaces as a function of the consumptions of the two inside goods) also suggested the need for similar normalizations.

outside good; except that it does not recognize greater correlations among the utility terms of inside goods. Further, recall from the discussion after Equation 15 that, to avoid estimation

breakdowns, the analyst must ensure that the term $\left(\frac{e_0}{\mathbb{E}_0 p_{j1}} - \frac{t_{j1}}{\chi_j} \right)$ is always positive.

Specifying \mathbb{E}_0 to be a random variable with an unbounded distribution makes it difficult to impose the above constraint while maintaining stability in estimation. Therefore, for convenience in model estimation, we specify \mathbb{E}_0 as 1. As a result, it becomes relatively easier to

ensure that $\left(\frac{e_0}{p_{j1}} - \frac{t_{j1}}{\chi_j} \right)$ is positive for all individuals in the estimation data by appropriately adjusting the χ_j parameter.

Further, to accommodate differences in satiation rates across destinations, the translation parameter χ_j can be specified as a function of destination (j) characteristics as: $\chi_j = \exp(v'v_j)$, where v_j is a vector of destination characteristics and v is a corresponding vector of parameters.

Specifying the joint cumulative distribution F of the random error terms $((v_{11}, v_{12}), (v_{21}, v_{22}), \dots, (v_{j1}, v_{j2}), \dots, (v_{J1}, v_{J2}))$ completes the random utility specification. In this paper, we assume that the random error terms have a nested extreme value structure with the following joint cumulative distribution:

$$F((v_{11}, v_{12}), \dots, (v_{j1}, v_{j2}), \dots, (v_{J1}, v_{J2})) = \prod_{j=1}^J \exp \left[- \left\{ e^{-\left(\frac{v_{j1}}{\tau_*} \right)} + e^{-\left(\frac{v_{j2}}{\tau_*} \right)} \right\}^\tau \right] \quad (19)$$

In the above cumulative distribution function, the error terms of the modal alternatives for a specific destination j , (v_{j1}, v_{j2}) are grouped into a nest, with a (dis)similarity parameter τ_*

introduced to capture correlations among the random utility contributions of all destination-mode combination alternatives jl sharing a destination j . \dagger is a scale parameter that can be estimated due to variation in prices across the choice alternatives.

3.9 Consumption Probability Expression

Given the above-discussed econometric structure, the KKT conditions in Equations 14 and 15 can be expressed as:

$$v_{j1} = -\Delta'z_{j1} - \ln\left(\frac{e_0}{\mathbb{E}_0} - \frac{t_{j1}}{x_j}\right); \forall_j \in \text{chosen destination alternatives}$$

and

$$v_{j2} - v_{j1} < (\Delta'z_{j1} - \ln p_{j1}) - (\Delta'z_{j2} - \ln p_{j2}); \forall_j \in \text{chosen destination alternatives} \quad (20)$$

and

$$v_{k1} < -\Delta'z_{j1} + \ln\mathbb{E}_0 + \ln p_{k1} - \ln e_0, \forall_j \in \text{chosen destination alternatives}$$

and

$$v_{k2} < -\Delta'z_{j2} + \ln\mathbb{E}_0 + \ln p_{k2} - \ln e_0, \forall_j \in \text{chosen destination alternatives} \quad (21)$$

The above stochastic KKT conditions can be used to derive the probability expression for the household's annual destination and mode choices and corresponding time allocations.

Specifically, conditional on \mathbb{E}_0 , the probability that a household allocates e_0 amount of money

to the outside good and the rest for vacation such that the first M of J vacation destination

alternatives are chosen and the first available mode is chosen to travel to each of these

destinations, with a time allocation pattern $\mathbf{t} = ((t_{11}, 0), (t_{21}, 0), \dots, (t_{M1}, 0), (0, 0), \dots, (0, 0))$ is given

by:

$$\begin{aligned}
& P\{e_0, (t_{11}, 0), (t_{21}, 0), \dots, (t_{M1}, 0), (0, 0), \dots, (0, 0) | \mathbb{E}_0\} \\
& = |J / \mathbb{E}_0| \times \prod_{j=1}^M P\left\{v_{j1} = -\mathbf{z}_{j1} - \ln\left(\frac{e_0}{\mathbb{E}_0 p_{j1}} - \frac{t_{j1}}{x_j}\right)\right\} \\
& \quad \times \prod_{j=1}^M P\{(v_{j2} - v_{j1}) < (\mathbf{z}_{j1} - \ln p_{j1}) - (\mathbf{z}_{j2} - \ln p_{j2})\} \\
& \quad \times \prod_{k=M+1}^K P\{(v_{k1} < -\mathbf{z}_{k1} + \ln \mathbb{E}_0 + \ln p_{k1} - \ln e_0), (v_{k2} < -\mathbf{z}_{k2} + \ln \mathbb{E}_0 + \ln p_{k2} - \ln e_0)\} \\
& = |J / \mathbb{E}_0| \times \prod_{j=1}^M g_{v_{j1}} \left\{ -\mathbf{z}_{j1} - \ln\left(\frac{e_0}{\mathbb{E}_0 p_{j1}} - \frac{t_{j1}}{x_j}\right) \right\} \\
& \quad \times \prod_{j=1}^M G_{v_{j2}-v_{j1}} \{(\mathbf{z}_{j1} - \ln p_{j1}) - (\mathbf{z}_{j2} - \ln p_{j2})\} \\
& \quad \times \prod_{k=M+1}^K G_{v_{k1}, v_{k2}} \{(-\mathbf{z}_{k1} + \ln \mathbb{E}_0 + \ln p_{k1} - \ln e_0), (-\mathbf{z}_{k2} + \ln \mathbb{E}_0 + \ln p_{k2} - \ln e_0)\}
\end{aligned} \tag{22}$$

In the above expression, the Jacobian $|J / \mathbb{E}_0|$ and the next two terms correspond to all chosen destinations, while the last term corresponds to all unchosen destinations. Specifically, the terms $g_{v_{j1}}\{\cdot\}$ and $G_{v_{j2}-v_{j1}}\{\cdot\}$ together correspond to the probability for the choice of destination j ($j=1, 2, 3, \dots, M$) along with the corresponding time allocation and mode of travel.

The term $G_{v_{k1}, v_{k2}}\{\cdot, \cdot\}$ corresponds to the probability that the remaining alternatives are not chosen. $g_{v_{j1}}\{\cdot\}$ represents the probability density function of the random term v_{j1} ; $G_{v_{j2}-v_{j1}}\{\cdot\}$ represents the cumulative density function of the difference $(v_{j2} - v_{j1})$ between the random terms v_{j2} and v_{j1} ; and $G_{v_{k1}, v_{k2}}\{\cdot, \cdot\}$ represents the joint cumulative density function of the random terms (v_{k1}, v_{k2}) .

$$g_{v_{j1}} \left\{ - \mathbf{z}_{j1} - \ln \left(\frac{e_0}{\mathbb{E}_0 p_{j1}} - \frac{t_{j1}}{x_j} \right) \right\} = \frac{1}{\dagger} \exp \left(\frac{\mathbf{z}_{j1}}{\dagger} \right) \cdot \left(\frac{e_0}{\mathbb{E}_0 p_{j1}} - \frac{t_{j1}}{x_j} \right)^{\frac{1}{\dagger}} \cdot \exp \left\{ - \exp \left(\frac{\mathbf{z}_{j1}}{\dagger} \right) \cdot \left(\frac{e_0}{\mathbb{E}_0 p_{j1}} - \frac{t_{j1}}{x_j} \right)^{\frac{1}{\dagger}} \right\} \quad (23)$$

$$G_{v_{j2}-v_{j1}} \left\{ (\mathbf{z}_{j1} - \ln p_{j1}) - (\mathbf{z}_{j2} - \ln p_{j2}) \right\} = \frac{\exp \left(\frac{\mathbf{z}_{j1} - \ln p_{j1}}{\dagger_{\#}} \right)}{\exp \left(\frac{\mathbf{z}_{j1} - \ln p_{j1}}{\dagger_{\#}} \right) + \exp \left(\frac{\mathbf{z}_{j2} - \ln p_{j2}}{\dagger_{\#}} \right)} \quad (24)$$

$$G_{v_{k1}, v_{k2}} \left\{ (-\mathbf{z}_{k1} + \ln \mathbb{E}_0 + \ln p_{k1} - \ln e_0), (-\mathbf{z}_{k2} + \ln \mathbb{E}_0 + \ln p_{k2} - \ln e_0) \right\} = \exp - \left\{ \left(\frac{e_0}{\mathbb{E}_0} \right)^{1/\dagger} \left\langle \exp \left(\frac{\mathbf{z}_{k1} - \ln p_{k1}}{\dagger_{\#}} \right) + \exp \left(\frac{\mathbf{z}_{k2} - \ln p_{k2}}{\dagger_{\#}} \right) \right\rangle^{\dagger} \right\} \quad (25)$$

Note that the discourse so far is based on the assumption that only two modal alternatives are available for each destination and that the first mode is chosen for each visited destination. However, the probability expression in Equation 25 can be easily extended to the general case where $L (>2)$ number of modes are available to travel to each destination.

Let the first M destination alternatives be the chosen destinations ($j = 1, 2, \dots, M$), the chosen mode for each alternative be indexed by l_j , and that the time allocated to a chosen destination-mode combination alternative jl_j be t_{jl_j} . Then the probability expression for the household's annual destination-mode choices and time allocations is given below:

$$\begin{aligned}
& P\left\{e_0, (0, \dots, t_{1l_1}, \dots, 0), (0, \dots, t_{2l_2}, \dots, 0), \dots, (0, \dots, t_{Ml_M}, \dots, 0), (0, 0, \dots, 0), \dots, (0, 0, \dots, 0) \mid \mathbb{E}_0\right\} \\
& = |J / \mathbb{E}_0| \times \prod_{j=1}^M \frac{1}{\dagger} \exp\left(\frac{\mathbf{z}_{jl_j}}{\dagger}\right) \cdot \left(\frac{e_0}{\mathbb{E}_0 p_{jl_j}} - \frac{t_{jl_j}}{x_j}\right)^{\frac{1}{\dagger}} \cdot \exp\left\{-\exp\left(\frac{\mathbf{z}_{jl_j}}{\dagger}\right) \cdot \left(\frac{e_0}{\mathbb{E}_0 p_{jl_j}} - \frac{t_{jl_j}}{x_j}\right)^{\frac{1}{\dagger}}\right\} \\
& \quad \times \prod_{j=1}^M \frac{\exp\left(\frac{\mathbf{z}_{jl_j} - \ln p_{jl_j}}{\dagger''}\right)}{\sum_{l=1}^L \exp\left(\frac{\mathbf{z}_{jl} - \ln p_{jl}}{\dagger''}\right)} \\
& \quad \times \prod_{k=M+1}^K \exp\left\{-\left(\frac{e_0}{\mathbb{E}_0}\right)^{1/\dagger} \left\langle \sum_{l=1}^L \exp\left(\frac{\mathbf{z}_{kl} - \ln p_{kl}}{\dagger''}\right) \right\rangle^{\dagger''}\right\}
\end{aligned} \tag{26}$$

The term $|J / \mathbb{E}_0|$ in the above expression is the determinant of the Jacobian matrix (Conditional on \mathbb{E}_0) obtained from applying change of variables calculus between the vector of stochastic terms $(v_{1l_1}, v_{2l_2}, \dots, v_{jl_j}, \dots, v_{Ml_M})$ for all chosen destination-mode combination alternatives and the corresponding vector of time allocation variables $(t_{1l_1}, t_{2l_2}, \dots, t_{jl_j}, \dots, t_{Ml_M})$. This determinant does not have a compact form but the ih^{th} element of the matrix can be computed as:

$$J_{ih} / \mathbb{E}_0 = \frac{\partial}{\partial t_{hl_h}} \left[-\mathbf{z}_{il_i} - \ln \left(\frac{e_0}{\mathbb{E}_0 p_{il_i}} - \frac{t_{il_i}}{x_i} \right) \right] = \frac{\left(\frac{p_{hl_h}}{\mathbb{E}_0 p_{il_i}} + \frac{u_{ih}}{x_i} \right)}{\left(\frac{e_0}{\mathbb{E}_0 p_{il_i}} - \frac{t_{il_i}}{x_i} \right)}; (i, h = 1, 2, \dots, M) \tag{27}$$

where, u_{ih} is an indicator that takes a value of 1 if $i = h$, or zero otherwise. Note that all Equations (19) through (26), including the probability expression in Equation 26 are conditional on \mathbb{E}_0 . Recall from the earlier discussion that we assume \mathbb{E}_0 to be equal to 1. Thus, simply substituting 1 for \mathbb{E}_0 in Equations 26 and Equation 27 will provide the unconditional probability expression and the corresponding Jacobian expression.

3.10 Forecasting with Proposed Model

Until recent past, forecasting with KKT demand models was pursued to be very difficult, especially in the presence of imperfect substitutes and corner solutions. von Haefen et al. (2004) and Pinjari and Bhat (2011) proposed computationally efficient algorithms for forecasting with KKT demand models with only imperfect substitutes in the choice set. In this paper, we expand on the Pinjari and Bhat algorithm to the case with a mix of perfect and imperfect substitutes in the choice set.

Recall from Equation 10 that, for any chosen destination alternative $j (=1, 2, \dots, J)$, the price-normalized baseline marginal utility of the chosen destination-mode alternative is always greater than that of other alternatives sharing that same destination. One can write the same condition as:

$$\frac{\mathbb{E}_{jl_j}}{p_{jl_j}} = \text{Max}_{l=1,2,\dots,L} \left(\frac{\mathbb{E}_{jl}}{p_{jl}} \right), \quad (28)$$

where $\frac{\mathbb{E}_{jl_j}}{p_{jl_j}}$ is the price-normalized baseline utility of chosen destination-mode alternative jl_j .

Similarly, one can show using the KKT conditions derived in Section 3.5 that the price-normalized baseline utility of any chosen destination-mode alternative is greater than that of destination-mode alternatives for any unchosen destination alternative. Specifically:

$$\text{Max}_{l=1,2,\dots,L} \left(\frac{\mathbb{E}_{kl}}{p_{kl}} \right) < \} < \text{Max}_{l=1,2,\dots,L} \left(\frac{\mathbb{E}_{jl}}{p_{jl}} \right); \text{ if } k \text{ is unchosen and } j \text{ is a chosen destination.} \quad (29)$$

This is because $\left(\frac{\mathbb{E}_{kl}}{P_{kl}}\right) < \lambda \forall l = (1, 2, \dots, L)$ or $\text{Max}_{l=1,2,\dots,L} \left(\frac{\mathbb{E}_{kl}}{P_{kl}}\right) < \lambda$ for unchosen destinations and

$$\frac{\left(\frac{\mathbb{E}_{jl_j}}{P_{jl_j}}\right)}{\left(\frac{\mathbb{E}_{jl_j} t_{jl_j}}{x_j} + 1\right)} = \lambda \text{ or } \text{Max}_{l=1,2,\dots,L} \left(\frac{\mathbb{E}_{jl}}{P_{jl}}\right) > \lambda \text{ for chosen destinations.}$$

Further, one can rearrange the KKT condition for a chosen destination-mode alternative in Equation 15 to express the corresponding time allocation t_{jl_j} as:

$$t_{jl_j} = \left(\frac{\mathbb{E}_{jl_j}}{P_{jl_j}} - \lambda\right) \frac{x_j}{\mathbb{E}_{jl_j}} \quad (30)$$

Substituting this expression for t_{jl_j} into the budget constraint and solving for λ gives:

$$\lambda = \frac{\left(\mathbb{E}_0 + \sum_{j=1 \text{ to } M} x_j\right)}{\left(E + \sum_{j=1 \text{ to } M} \frac{x_j P_{jl_j}}{\mathbb{E}_{jl_j}}\right)} \quad (31)$$

In the above expression, $j = (1, 2, \dots, M)$ are the chosen alternatives and l_j is the chosen mode for a chosen destination j . It is important to note that the summation terms in the above expression include chosen destination-mode alternatives only. Given these insights, the forecasting algorithm for the model proposed in this paper comprises four basic steps as outlined next.

- **Step 0:** Assume that only the outside good is chosen. Initialize the total number of chosen destinations as: $M = 0$, and Lagrangian multiplier as: $\lambda = \frac{\mathbb{E}_0}{E}$.

- Given the input data (z_{jl}, p_{jl}) , model parameters (α_j, x_j) , and the simulated error term (v_{jl}) draws, compute the price-normalized baseline utility values (\mathbb{E}_{jl}/p_{jl}) for all destination-mode alternatives.
- Step 1: For each destination $j (=1,2,\dots,J)$, pick the modal alternative with the maximum price-normalized baseline utility value, $\text{Max}_{l=1,2,\dots,L} (\mathbb{E}_{jl}/p_{jl})$. Label the corresponding destination-mode alternative as jl_j . Then $\text{Max}_{l=1,2,\dots,L} (\mathbb{E}_{jl}/p_{jl})$ can be denoted by $(\mathbb{E}_{jl_j}/p_{jl_j})$.
 - Arrange all the J destination alternatives available to the consumer in the descending order of $\text{Max}_{l=1,2,\dots,L} (\mathbb{E}_{jl}/p_{jl})$.
- Step 2: If $M = 0$ and $\lambda > \text{Max}_{l=1,2,\dots,L} \left(\frac{\mathbb{E}_{1l}}{p_{1l}} \right)$ (i.e., the maximum price-normalized baseline utility among all modes available to the first destination alternative in the above arrangement),
 - Set the optimal consumption of the outside good as $e_0^* = E$ and that of all destination-mode alternatives as zero (i.e., no destinations are visited in this case) and stop.
 - Else, if $M = 0$ and $\lambda < \text{Max}_{l=1,2,\dots,L} \left(\frac{\mathbb{E}_{1l}}{p_{1l}} \right)$,
 - Set $M = M+1$ (because the next destination alternative must be a chosen alternative).
 - Update the Lagrangian multiplier as $\lambda = \left(\mathbb{E}_0 + \sum_{j=1 \text{ to } M} x_j \right) / \left(E + \sum_{j=1 \text{ to } M} \frac{x_j p_{jl_j}}{\mathbb{E}_{jl_j}} \right)$.

- Step 3: If $M = J$,
 - Compute the optimal consumption of the outside good as $e_0^* = (E_0 / \lambda)$
 - Compute the optimal consumptions of the first M destination alternatives (and corresponding chosen mode alternatives) as $t_{jl_j}^* = \left(\frac{E_{jl_j}}{P_{jl_j}} - \lambda \right) \frac{x_j}{E_{jl_j}}$, and stop.
 - Else, if $M < J$, go to step 4.

- Step 4:

- If $\lambda < \text{Max}_{l=1,2,\dots,L} \left(\frac{E_{(M+1)l}}{P_{(M+1)l}} \right)$,

- Set $M = M + 1$ and update the Lagrangian multiplier as:

$$\lambda = \left(E_0 + \sum_{j=1 \text{ to } M} x_j \right) / \left(E + \sum_{j=1 \text{ to } M} \frac{x_j P_{jl_j}}{E_{jl_j}} \right).$$

- Go to step 3.

- Else, if $\lambda > \text{Max}_{l=1,2,\dots,L} \left(\frac{E_{(M+1)l}}{P_{(M+1)l}} \right)$ (i.e., the maximum price-normalized baseline utility

among all modes available to the destination alternative in position $M+1$),

- Compute the optimal consumption of the outside good as $e_0^* = (E_0 / \lambda)$
- Compute the optimal consumptions of the first M destination alternatives in the above arrangement (and corresponding chosen mode alternatives) as

$$t_{jl_j}^* = \left(\frac{E_{jl_j}}{P_{jl_j}} - \lambda \right) \frac{x_j}{E_{jl_j}}.$$

- Set the consumptions of all other destination-mode alternatives as zero and stop.

The above outlined forecasting procedure can be applied repeatedly over a large number of simulated error term (v_{jt}) draws to obtain distributions for the consumer choice predictions. The procedure can be used to apply empirical models for forecasting and policy prediction purposes as well as to simulate data that exhibits a mix of both perfect and imperfect substitution among choice alternatives.

3.11 Accommodation of Multiple Budget Constraints

The discourse so far is based on the assumption that a single linear budget constraint governs consumer choices. However, several consumer choices involve the use of multiple resources such as time and money and therefore governed by multiple constraints. In the current empirical context, households' leisure travel decisions are potentially influenced by both time and money constraints.

For example, some households may have the time to travel to exotic and far away destinations but not enough money to do so. On the other hand, some households may simply not have the time for long vacations even if they are able to afford the expenses. In most cases, both time and money constraints are likely to influence the choices. To accommodate both these constraints, the households' utility maximization problem discussed in Section 3.2 can be extended as below (assuming only two modes of travel are available for each destination):

$$\underset{(t, e_0, t_0)}{\text{Max}} U(t, e_0, t_0) = \sum_{j=1}^J x_j \ln \left\{ \frac{(\mathbb{E}_{j1} t_{j1} + \mathbb{E}_{j2} t_{j2})}{x_j} + 1 \right\} + \ln e_0 + \ln t_0 \quad (32)$$

Subject to two linear budget constraints—one for time and the other for money—as below:

$$\sum_{j=1}^J (p_{j1}t_{j1} + p_{j2}t_{j2}) + e_0 = E,$$

and

$$\sum_{j=1}^J (q_{j1}t_{j1} + q_{j2}t_{j2}) + t_0 = T. \quad (33)$$

In the above utility function (Equation 32), all terms in the subutility function $x_j \ln \{ (\mathbb{E}_{j1}t_{j1} + \mathbb{E}_{j2}t_{j2}) / x_j + 1 \}$ are as described in Section 3.2. The second and third terms, $\ln e_0$ and $\ln t_0$, are specified following Satomura et al.'s (2011) specification for outside goods in the presence of multiple budget constraints. Specifically, e_0 is the Hicksian composite outside good for money (i.e., income - annual expenditure on vacation) and t_0 is the Hicksian composite outside good for time representing all the nonvacation time in a year (i.e., 365 days - annual number of days spent on vacation). These terms complete the utility function to form an incomplete demand system. Both the outside goods are assumed to be “essential” with some positive consumption by all households since neither the entire year nor the whole income is typically spent completely on vacation.

In the two constraints identified in Equation 33, q_{jl} and p_{jl} ($l = 1, 2$) represent the time-prices and money-prices, respectively, of consuming unit time of a vacation destination j traveled by mode l .⁷ As can be observed from the two constraints, following Satomura et al. (2011), the Hicksian composite outside good for time (t_0) is assumed to have unit time-price and zero

⁷ The time-price (q_j), in the current empirical context, is the amount of time that needs to be expended to consume a unit amount of vacation time. If one assumes that the time spent traveling is a part of vacation time (i.e., people derive utility from traveling for vacation), then the time-price is unity. On the other hand, if the time spent traveling is viewed only as a cost without any contribution to utility, the time-price is more than unity; since a day of vacation at a destination incurs one day of vacation plus some amount of travel. In the current empirical application, we consider the time-price to be unity assuming that travel time is a part of vacation and it contributes to utility. The money-price is simply the amount of money expended to consume a unit amount of vacation time.

money-price (i.e., it doesn't appear in the time constraint), while the money-specific outside good (e_0) has unit money-price and zero time-price (i.e., it doesn't appear in the money constraint).

To setup the optimality conditions for the time- and money-constrained utility maximization problem without considering the third constraint, one can form a Lagrangian function as below:

$$L = U(t, e_0, t_0) - \lambda \left[e_0 + \sum_{j=1}^J (p_{j1}t_{j1} + p_{j2}t_{j2}) - E \right] - \mu \left[t_0 + \sum_{j=1}^J (q_{j1}t_{j1} + q_{j2}t_{j2}) - T \right] \quad (34)$$

As described in Section 3.5, applying the KKT conditions for the outside goods result in the following Lagrangian multipliers: $\lambda = 1/e_0$ and $\mu = 1/t_0$, representing the marginal utility of time and money, respectively. Further, the following KKT conditions can be derived for the inside goods (i.e., destination-mode combinations).

$$\begin{aligned} \frac{\mathbb{E}_{jl}}{\left\{ \frac{(\mathbb{E}_{j1}t_{j1} + \mathbb{E}_{j2}t_{j2})}{x_j} + 1 \right\}} &= \dagger_{jl} \text{ if } t_{jl} > 0; \forall j = 1, 2, \dots, J; l = 1, 2 \\ \frac{\mathbb{E}_{jl}}{\left\{ \frac{(\mathbb{E}_{j1}t_{j1} + \mathbb{E}_{j2}t_{j2})}{x_j} + 1 \right\}} &< \dagger_{jl} \text{ if } t_{jl} = 0; \forall j = 1, 2, \dots, J; l = 1, 2 \end{aligned} \quad (35)$$

where, $\dagger_{jl} = \mu q_{jl} + \lambda p_{jl} = \frac{q_{jl}}{t_0} + \frac{p_{jl}}{e_0}$

Again following the discourse in Section 3.2, assuming that the first M destination alternatives are chosen and that the first mode of travel is chosen to travel to each of these destinations, one can derive the following separate KKT conditions for all chosen destinations:

$$\frac{\mathbb{E}_{j1}}{\left\{ \frac{\mathbb{E}_{j1} t_{j1}}{x_j} + 1 \right\}} = \dagger_{j1} \text{ since } t_{j1} = 0; \forall j \in \text{chosen destinations}$$

and

$$\frac{\mathbb{E}_{j2}}{\mathbb{E}_{j1}} < \frac{\mathbb{E}_{j1}}{\mathbb{E}_{j1}}, \text{ since, } t_{j1} > 0 \text{ and } t_{j1} = 0; \forall j \in \text{chosen destinations.} \quad (36)$$

Similarly, one can derive the following KKT conditions for unchosen destinations.

$$\mathbb{E}_{k1} < \dagger_{j1} \text{ and } \mathbb{E}_{k2} < \dagger_{j2}, \text{ since } t_{k1} = 0 \text{ and } t_{j2} = 0; \forall k \in \text{chosen destinations} \quad (37)$$

Assuming the same stochastic distributions as in Section 3.6, the following stochastic KKT conditions can be derived:

$$v_{j1} = -\Delta'z_{j1} - \ln\left(\frac{1}{\dagger_{j1}} - \frac{t_{j1}}{x_j}\right); \forall j \in \text{chosen destinations},$$

and

$$v_{j2} - v_{j1} < (\Delta'z_{j1} - \ln \dagger_{j1}) - (\Delta'z_{j2} - \ln \dagger_{j2}); \forall j \in \text{chosen destinations} \quad (38)$$

and

$$v_{k1} < -\Delta'z_{j1} + \ln \dagger_{k1}, \forall k \in \text{non-chosen destination alternatives}$$

and

$$v_{k2} < -\Delta'z_{j2} + \ln \dagger_{k2}, \forall k \in \text{non-chosen destination alternatives} \quad (39)$$

These stochastic KKT conditions can be used to derive the following probability expression for households' annual destination and mode choices and corresponding time allocations:

$$\begin{aligned}
& P\{t_0, e_0, (t_{11}, 0), (t_{21}, 0), \dots, (t_{M1}, 0), (0, 0), \dots, (0, 0)\} \\
& = |J| \times \prod_{j=1}^M \frac{1}{\dagger} \exp\left(\frac{\mathbf{z}_{j1}}{\dagger}\right) \cdot \left(\frac{1}{\dagger_{j1}} - \frac{t_{j1}}{x_j}\right)^{\frac{1}{\dagger}} \cdot \exp\left\{-\exp\left(\frac{\mathbf{z}_{j1}}{\dagger}\right) \cdot \left(\frac{1}{\dagger_{j1}} - \frac{t_{j1}}{x_j}\right)^{\frac{1}{\dagger}}\right\} \\
& \quad \times \prod_{j=1}^M \frac{\exp\left(\frac{\mathbf{z}_{j1} - \ln \dagger_{j1}}{\dagger''}\right)}{\exp\left(\frac{\mathbf{z}_{j1} - \ln \dagger_{j1}}{\dagger''}\right) + \exp\left(\frac{\mathbf{z}_{j2} - \ln \dagger_{j2}}{\dagger''}\right)} \\
& \quad \times \prod_{k=M+1}^J \exp\left\{-\left\langle \exp\left(\frac{\mathbf{z}_{k1} - \ln \dagger_{k1}}{\dagger''}\right) + \exp\left(\frac{\mathbf{z}_{k2} - \ln \dagger_{k2}}{\dagger''}\right) \right\rangle\right\}
\end{aligned} \tag{40}$$

In the above expression, $|J|$ is the determinant of the Jacobian matrix due to change of variables from the vector of time allocation variables to the corresponding stochastic terms for all chosen destination-mode combination alternatives. The ih^{th} element of the matrix is given below:

$$J_{ih} = \frac{\partial}{\partial t_{h1}} \left[-\mathbf{z}_{i1} - \ln\left(\frac{1}{\dagger_{i1}} - \frac{t_{i1}}{x_i}\right) \right] = \frac{\left(\frac{((q_{i1}q_{h1})/t_o^2) + ((p_{i1}p_{h1})/e_o^2)}{\dagger_{i1}^2} \right) + \frac{u_{ih}}{x_i}}{\left(\frac{1}{\dagger_{i1}} - \frac{t_{i1}}{x_i} \right)};$$

$(i, h = 1, 2, \dots, M)$ (41)

where, $\dagger_{i1} = \frac{q_{i1}}{t_o} + \frac{p_{i1}}{e_o}$, and u_{ih} is an indicator that takes a value of 1 if $i = h$, or zero otherwise.

Note from the KKT conditions in Equation 38 and the probability expression in Equation 40 that the term $\left(\frac{1}{\dagger_{j1}} - \frac{t_{j1}}{x_j}\right)$ is inside a logarithmic function. During model estimation, the analyst must ensure that this term is always positive, lest it should lead to estimation breakdowns.

The reader will note here that the above formulation simplifies to a formulation with a single linear budget constraint when the outside good quantity corresponding to one of the

constraints is infinity (i.e., when one of the constraints is relaxed). For example, when the time constraint is relaxed (i.e., when $t_0 \rightarrow \infty$), the above formulation collapses to the model in Section 3.7 with a money budget constraint. Similarly, when the money constraint is relaxed, the formulation collapses to a model with a time budget constraint. The next section applies this RUM model framework to analyze annual household leisure travel choices utilizing data from the ATS (1995).

3.12 Empirical Application

A significant portion of passenger travel miles in the United States comes from long-distance travel, especially for leisure purposes such as vacation. Statistics from national travel surveys (BTS, 2001) indicate that more than one half of all long-distance travel is for leisure. Besides, long-distance leisure travel garners particular attention due to its impact on the tourism and recreation industry. Due to all these reasons, long-distance leisure travel has been studied extensively in the tourism and recreation demand literatures, and is steadily gaining importance in the transport planning/modeling arena.

Most work on this topic in the transport modelling arena can be categorized into (a) Statewide travel models in the United States (e.g., Horowitz, 2008), (b) National travel models in Europe (e.g., Fosgerau, 2001; Rich et al., 2009), and (c) Intercity travel demand analysis between specific city pairs (e.g., Koppelman & Sethi, 2005). An important end-goal of all these efforts is to estimate travel flows between different regions by different modes of travel for informing various policy and investment initiatives. A drawback of most literature in both the travel demand and tourism fields is that the analysis is typically limited to smaller time frames such as a day or a few weeks. However, analysis of the 1995 American Travel Survey data indicates that on average, a household makes less than four vacation trips over a year.

Given the infrequent nature of long-distance leisure travel as evident from above, a smaller time-frame of analysis (e.g., a day) is likely to provide a distorted picture of leisure travel flows in the nation. Intuitively, vacations are planned over longer time frames, as opposed to daily travel decisions for which shorter time frames may suffice. In this context, Eugenio-Martin's (2003) theoretical framework for tourism demand analysis suggests 1 year as appropriate for vacation travel analysis (also see Morley, 1992). The need for considering a longer time frame of analysis is also well-recognized in the recreational demand literature, where numerous studies analyze recreational choices over longer time horizons such as seasons than over a single choice instance (see Bockstael et al., 1987; Phaneuf et al., 2000; Phaneuf & Smith 2005; von Haefen & Phaneuf, 2003).

To be sure, a few studies in the transport planning arena do consider longer time-frames for analyzing leisure travel. These include the annual leisure travel framework of van Middelkoop, Borgers, and Timmermans (2004), the holiday travel module in the pan-European long-distance travel model (Rich et al., 2009), and annual vacation time allocation studies by LaMondia, Bhat, and Hensher (2008) and Van Nostrand, Sivaraman, and Pinjari (2013). Among these studies, the paper by Van Nostrand et al. is the most relevant to the current empirical research, as discussed next.

Van Nostrand et al. (2013) applied Bhat's MDCEV framework for modeling households' annual vacation destination choices and the corresponding time allocations, recognizing that households could potentially visit multiple destinations over a year (i.e., destination choices are imperfect substitutes). Specifically, they formulated an annual time allocation model assuming that an annual vacation time budget exists for each household.

Building on their work, this chapter estimates an integrated model of annual vacation destination choice and mode choices to recognize that vacation destination and mode choices are made in a joint fashion (Van Nostrand et al., 2013, do not consider mode choice jointly with destination choice). The framework recognizes that, while vacation destinations are imperfect substitutes, the travel mode alternatives to a destination are perfect substitutes in that only one primary mode is chosen to travel to a destination. In addition, the current empirical work explicitly considers the role of both time and money constraints (as opposed to a single constraint) in households' long-distance leisure travel destination and mode choice decisions, while considering the price variation across destination and mode choice alternatives.

The empirical model in this paper is estimated using household-level long-distance travel data from the 1995 ATS, as assembled by Van Nostrand et al. (2013). To define the destination choice alternatives, first each of the metropolitan statistical areas (MSAs) from each of the 48 contiguous states in the United States was identified as a potential vacation destination. Subsequently, the remaining non-MSA area in each state was counted as a single destination (one non-MSA area for each state. All together, the U.S. was divided into 210 destinations comprising 162 MSA destinations and 48 non-MSAs. For each destination, auto and air were considered as the two primary modes of travel.

The primary intent of the current empirical application is only a demonstration of the proposed methodological framework to jointly model SDC and MDC choices in the presence of a mix of imperfect and perfect substitutes in the choice set. Further empirical investigations such as policy simulations with the proposed framework are left out for further research. To this end, one must first develop forecasting algorithms to apply the proposed framework with multiple budget constraints (the algorithm in Section 3.8 is applicable only for single budget constraint).

3.13 Data

The 1995 ATS is the source of household-level vacation destination and mode choice data used in this analysis. The survey collected information from over 60,000 American households on all long-distance trips each household made over an entire year to destinations farther than 100 miles (BTS, 1995a). For each trip, the information on the purpose, mode, and destination of travel and other travel attributes such as the time spent (no. of days) on the trip and travel party size were collected. From this sample, 22,215 households reported making at least one trip over the year for one of the four leisure purposes—relaxation, sightseeing, entertainment, or outdoor recreation—by either the car mode or the air mode of travel.

Out of the above-mentioned households, a random sample of 966 households was selected for model estimation while another random sample of 500 households was selected for model validation. In addition to the 1995 ATS, several secondary data sources were utilized to compile other required information such as (a) the transportation level of service variables, including the travel times and costs between each origin–destination pair via air and auto modes, (b) lodging prices, and nonlodging (dining, entertainment/recreation, and other) prices at each of the 210 destinations, (c) the destination size and attraction variables for the year 1995, including land area, number of employees in different sectors (leisure and hospitality, retail, etc.), total population, and gross domestic product, and (d) the 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.

Details on the data sources and procedures used to create the specific variables of interest can be found in Van Nostrand et al. (2013). Since the focus of this research is on accommodating imperfect substitutability across destination choice alternatives and perfect substitutability across

mode choice alternatives in the presence of price variation across the different vacation destinations and travel modes, the procedures and assumptions used to construct the price variables using the 1995 CEX data are described in Appendix A.

Table 6 presents the descriptive statistics from the estimation sample used in this analysis. The households in the estimation sample have an average size of 2.86 persons/household, householder age of 45.6 years, and an average income of \$50,694 per annum. 38% of them have at least one child of age less than 16 years, and 13% of them have householders who are retired. About 48 % of the households in the estimation sample made multiple long-distance leisure trips over a year.

Further, a significant proportion (close to 40%) of the households visited multiple destinations. Furthermore, a large percentage (81%) of households visited a destination (if they did so) only once. This suggests multiple discreteness (or variety-seeking) in households' annual destination choices. Even if households visited a destination more than once a year, a vast majority of the times (99.5% of the time in the data; not shown in the table) the same mode was used to travel across all the different trips made by a household to that same destination. This suggests perfect substitution (or single discreteness) among the mode choice alternatives to a destination. The total annual household time spent on vacation ranged from a day to about 44 days, with an average of 6 days. The estimated total annual vacation expenditure ranged from \$30 to \$5,660 (not shown in table), with an average of \$693 per household.

The trip level characteristics suggest 84.7% of the trips in the estimation sample were by auto mode and the remaining 15.3% were by the air mode. The average round trip distance was about 990 miles. On average, the households in the sample spent about 4 days and about \$437 on

each trip. The last set of descriptive statistics is for the characteristics of the 210 destinations and level of service variables between the 210 x 210 possible Origin-Destination pairs.

3.14 Model Estimation

The parameters of the proposed model formulations were estimated using the maximum likelihood method. Three different models were estimated: (a) A model that considers only the time budget constraint, (b) A model that considers only the money budget constraint, and (c) A model that considers both time and money budget constraints. In all three models, the destination choice alternatives are considered to be imperfect substitutes while the mode choice alternatives for each destination are considered to be perfect substitutes. The log-likelihood functions of all the three models were coded in the GAUSS matrix programming language.

As discussed earlier, the term $\left(\frac{t_0}{q_{j1}} - \frac{t_{j1}}{x_j} \right)$ in the time-constrained model must be always positive to avoid estimation breakdowns. Similarly, the term $\left(\frac{e_0}{p_{j1}} - \frac{t_{j1}}{x_j} \right)$ must be positive during the estimation of the money-constrained model, while the term $\left(\frac{1}{\dagger_{j1}} - \frac{t_{j1}}{x_j} \right)$ must be positive during the estimation of the time- and money-constrained model.⁸

To consider these constraints, we first attempted to use the constrained maximum likelihood module of GAUSS. But we encountered estimation instability and convergence issues with the constrained maximum likelihood module. Therefore, all the models were estimated

⁸ Recall that $\dagger_{j1} = \frac{q_{j1}}{t_0} + \frac{p_{j1}}{e_0}$. Therefore, $\frac{1}{\dagger_{j1}}$ becomes $\frac{t_0}{q_{j1}}$ in the single(time)-constrained model and $\frac{e_0}{p_{j1}}$ in the single(money)-constrained model.

using the maximum likelihood module of GAUSS, while ensuring positivity of the above mentioned terms at each of the iterations during estimation.

To do so, if the positivity condition was not met at any iteration, the iteration-search parameters of the χ_j function were updated to ensure positivity without deviating too much from their original values. Admittedly, the above-described approach of ensuring the positivity constraints is somewhat ad hoc. To verify whether the approach results in appropriate parameter estimates, we conducted simulation experiments with the single constrained model formulation. Specifically, we used the forecasting procedure described in Section 3.8 to simulate discrete-continuous choice data that exhibits both perfect and imperfect substitution patterns among choice alternatives, assuming a single linear budget constraint (i.e., the money budget constraint).

Two different types of data were simulated based on two different assumptions: (a) with the satiation parameters (χ_j) assumed to be constants, and (b) with the satiation parameters (χ_j) specified as a function of alternative attributes. In both cases, the baseline utility parameters were specified as a function of alternative attributes and decision-maker characteristics. Next, maximum likelihood estimation was performed on the simulated data to retrieve the parameters used to simulate the data. As discussed earlier, at each iteration, the constraint that the term

$\left(\frac{e_0}{p_{j1}} - \frac{t_{j1}}{\chi_j} \right)$ is positive was ensured in a heuristic fashion by updating the iteration-search parameters in the χ_j function (if the constraint was not satisfied).

The overall findings from this exercise are presented in this section. The structural parameters of both the baseline utility and satiation functions used to simulate the data could be

easily retrieved through maximum likelihood estimation when the satiation parameters χ_j were assumed to be constants.

Using different sets of starting values for the parameters did not generally influence either the estimability of parameters or the estimates at convergence. However, the estimation process was relatively slow with its stability and convergence depending on the starting values when the satiation parameters were specified as a function of alternative attributes. Also, the need for the above-described adjustment to the iteration-search parameters depended on the starting values. Using starting values that were closer to the true parameter values resulted in an easier estimation without having to make adjustments to the iteration-search parameters.

Based on the above insights from the simulation experiments, the empirical model estimation was carried out in a step-by-step manner, beginning with the estimation of specifications with only a constant in the satiation (χ_j) functions and using those parameter estimates as starting values for richer specifications with destination-specific variables in the χ_j functions. Different sets of starting values were explored as well. Limited explorations suggest that while not all starting values necessarily lead to convergence, the same set of parameter estimates were obtained whenever the model converged.

3.15 Model Results

For each of the three empirical models estimated—a time-constrained model, a money-constrained model, and a time- and money-constrained model—a variety of model evaluation measures, namely, log-likelihood (LL) at convergence, Rho-square (\dots^2), Bayesian information criterion, and predictive log-likelihood (PLL) on a validation sample of 500 households, are presented in Table 7.

As can be observed, the log-likelihood value of the time- and money-constrained model is better than that of the single, time-constrained (money-constrained) model by 572 (991) points. All other goodness of fit measures in the table (Rho-square, Bayesian information criterion, and PLL on a sample of 500 households) also suggests that the time- and money-constrained model performs better than the two single-constrained models.

An unnested likelihood ratio test was also conducted to compare the model fit of the time- and money-constrained model with that of the time-constrained model (which has a better log-likelihood value than the money-constrained model). To do so, a naïve time constrained model with only constants in it (with a log likelihood value of -12,483.) was considered as the base. The rho-square values for time-constrained model and the time- and money-constrained models are 0.1492 and 0.1951, respectively with respect to the naïve, time constrained model. The difference between the above adjusted rho-squared values is 0.0459. The probability that this difference could have occurred by chance is less than $\Phi(-\sqrt{-2 \times 0.0459 \times -12,483})$. This value is almost zero, suggesting that the time- and money-constrained model has a better data fit compared to the time-constrained model. All these results suggest the need to consider both the constraints. Table 8 presents the parameter estimates from the time- and money-constrained model (or multiple-constrained model). The parameter estimates of the single constrained models are not reported here but are available from the authors. Since the time- and money-constrained model performs better than the two single-constrained models in its goodness of fit to estimation data as well as a validation sample, we use the former model to discuss the influence of different factors on households' annual destination and mode choices.

While most substantive interpretations of the parameter estimates in the multiple-constrained models are not different from the single-constrained models, wherever appropriate,

we discuss the differences in the interpretations from the single-constrained models. The specification of the baseline utility function (β_{jl}) is discussed first, followed by the specification of the translation function (α_j).

The first set of explanatory variables in the baseline utility function have common coefficients across all destination-mode combinations (i.e., inside goods) with the outside goods as the base category (for normalization). Specifically, the constant for all destination-mode combinations is negative suggesting that households spend a smaller proportion of the year on vacation compared to the time spent on all other purposes captured in the outside goods (such as work, sleep, leisure activities pursued closer to the household). This is reasonable because the amount of annual time that a household typically spends on vacation is much less compared to the other time investments to be made in the year. The next variable, leisure employment per capita at the household location captures the influence of opportunities for leisure activities within a closer vicinity of the household (as opposed to long-distance destinations).

As expected, the negative coefficient suggests that households living in places with greater leisure opportunities are likely to spend less time on long-distance vacation. This result points to higher substitution between the leisure time spent locally and the time spent on long-distance vacation for households in locations with greater leisure opportunities. While the result is intuitive with a statistically significant coefficient in the time- and money-constrained model, the corresponding coefficient was not statistically significant when only the time constraint was considered.

The second set of variables in the baseline utility function comprises destination-specific characteristics. The interpretations of these variables have reasonable substantive interpretations similar to those discussed in Van Nostrand et al. (2013) who considered only the time constraint.

Specially, the logarithm of land-area variable that controls for size differences across the destination choice alternatives has a positive coefficient less than one. This can be explained based on spatial aggregation of several smaller destination alternatives into larger destinations for modeling purposes. As explained by Daly (1982), a smaller than unit coefficient suggests significant heterogeneity across the elemental destination alternatives that comprise the destination alternatives in the model. The positive coefficient on the leisure employment per capita variable reflects a greater attractiveness of destinations with higher leisure opportunities. The dummy variables for the destinations being in the same or adjacent states have positive coefficients reflecting that households are more likely to visit familiar destinations that tend to be within or adjacent to their residential state. The coefficients on the temperature variables during winter and summer suggest that destinations with moderate temperatures (65–75 degree Fahrenheit) are generally more attractive for vacation purposes.

The third set of variables is specific to the travel modes under consideration. The alternative specific constant reflects that households have a general preference to travel by car even after considering the time- and money-constraints and other mode-specific variables in the model. MSA origins and destinations are more attractive for the air mode of travel than the non-MSA origins or destinations, perhaps because of a greater access to the air travel mode in the MSAs.

The last variable in this category is the round trip travel time by the alternative modes of travel, whose negative coefficient suggests that households prefer to travel by faster modes of travel. In addition to its influence on mode choice, this variable helps in accommodating that farther destinations are less attractive for vacation compared to closer destinations. Note that mode-specific travel costs are not included as explanatory variables in the time- and money-

constrained model, while the travel times are included as explanatory variables.⁹ This is because the travel costs are already incorporated into the money-budget constraint through the money-prices (p_{jt}) of travel to the destinations. Such money-prices help in incorporating that farther destinations are more pricy to travel to and hence less likely to be chosen because of the monetary constraint. On the other hand, as discussed earlier, the travel times were not incorporated into the time-prices (q_{jt}). This is because the time-price (q_{jt}) of allocating unit time for a destination has been set to unity assuming that traveling also contributes to the utility derived from vacation (in addition to the utility due to the time spent at the destination).

The next parameter is the scale (τ) of the error terms (v_{jt}) in the baseline utility functions (U_{jt}). This parameter provides a measure of variation in the household preferences due to unobserved factors. The parameter was fixed to 1 in the time-constrained model as it could not be estimated due to the absence of price variation (Bhat, 2008). In the other two models, the parameter could very well be estimated and is significantly different from 1. Specifically the estimate is 0.748 in the money-constrained model (not reported in the table) and 0.576 in the time- and money-constrained model.

These estimates suggest that the magnitude of variation in the household preferences due to unobserved factors is lower in the time- and money-constrained model than that in the two single-constrained models. This may be because accounting for both the time and money constraints together helped in capturing a greater proportion of the variation in household preferences. The next parameter is the dissimilarity parameter (σ). The estimate for this parameter is significantly different from 1 (in all three models) suggesting the significant

⁹ The single, time-constrained model, on the other hand, includes mode-specific travel costs as explanatory variable with a negative coefficient. This is because the time-constrained model does not explicitly consider the money constraint.

presence of destination-specific unobserved factors inducing correlations between the baseline utility parameters of the destination-mode combination alternatives that share the same destination. Neglecting such correlations and estimating the destination and mode choice models separately would result in significantly inferior model fit.

The last set of variables correspond to the translation parameters (x_j), which allow for corner solutions as well as differential satiation effects across different vacation destinations. The positive coefficient on the distance variable in the x_j function suggests that households tend to allocate greater amount of time for vacation destinations that are farther (than those that are closer). This may be because households might want to spend more time at a destination that is farther from home (if they chose to visit the destination).

Besides, it generally takes greater amount of time to travel to father destinations. Overall, the model estimation results are all reasonable and shed light on the various factors influencing households' annual vacation destination and mode choices and related time and money allocations. The results demonstrate the applicability of the proposed framework for modeling discrete-continuous choices in the presence of a mix of perfect and imperfect substitutes in the choice set. Empirically, the results highlight the need for accommodating both time and money constraints in modeling households' vacation travel choices.

The model estimated in this chapter is more to demonstrate the functional abilities of the RUM model formulation developed in Chapter 3. It is encouraging to have been able to estimate a stable model with reasonable and intuitive parameters for variables of household sociodemographics, destination and mode choices. One of the main objectives of this dissertation was to develop a comprehensive model of a household pleasure travel demand at a national level with intent to estimate their annual long distance pleasure travel demand. The next Chapter 5

progresses towards this objective, constructing such a model accounting for their motivations (leisure or visiting) and their consequent choice of destination. It also accounts for the likelihood of households' propensities to make one or more trips for each of these motivations to a destination, along with their principal mode of travel across such trips over an annum.

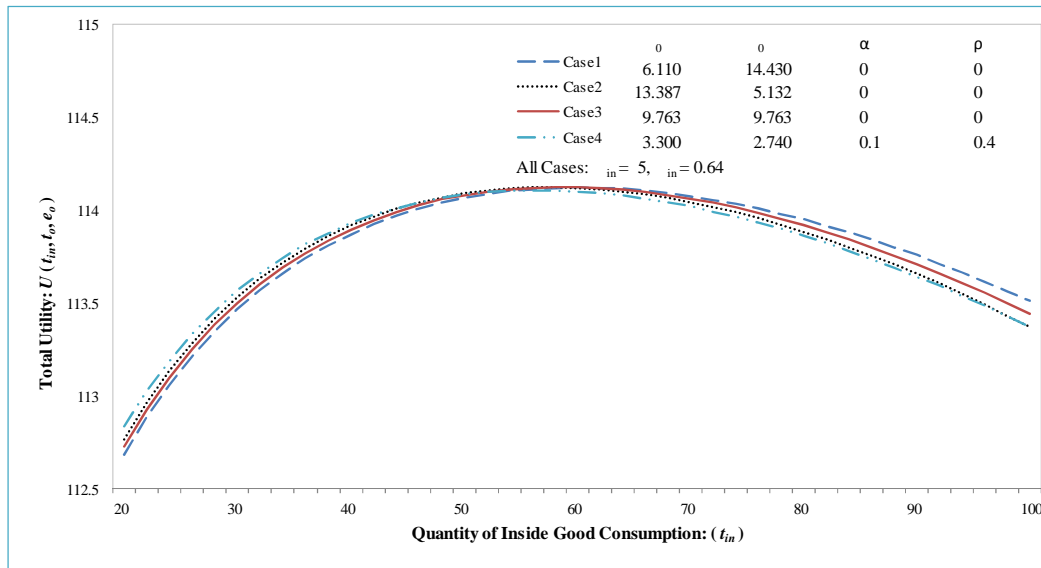


Figure 27 Total utility as a function of inside good consumption for different values of α_0, W_0, Γ , and ... (keeping all else same).¹⁰

¹⁰ The figure presents a “zoom-in” of the utility curves for inside good consumption values ranging from 20 to 100. Utility profiles over the entire consumption range (i.e., 0 to 365) make it very difficult to visually distinguish the four curves from one another.

Table 6 Descriptive statistics for estimation data

Household sociodemographic characteristics (in the estimation sample of 966 households)		
Household size	Average: 2.86	Std. Dev: 1.35
Age of householder (years)	Average: 45.6	Std. Dev.: 14.08
Household yearly income	Average: \$50,694	Std. Dev: \$30,544
Presence of children		38.4%
Householder is retired		12.6%
Household leisure travel characteristics (in the estimation sample of 966 households)		
Number of long distance leisure trips	Average: 2.2	Std. Dev: 2.07
1		51.6%
2 or more		48.4%
Number of destinations visited	Average: 1.58	Std. Dev: 0.93
1		62.7%
2		23.8%
3 or more		13.5%
Number of trips made to a destination	Average: 1.39	Std. Dev: 1.20
1		80.7%
2 or more		19.3%
Total annual vacation time (days)	Average: 6.40	Std. Dev: 5.40
Total annual expenditure on vacation	Average: \$693	Std. Dev: \$672
Trip-level characteristics (for 1530 leisure trips made by the 966 households)		
Primary mode of transportation	Auto: 84.7%	Air: 15.3%
Round trip ground distance (miles)	Average: 990	Std. Dev: 1,116
No. of nights away from home on trip	Average: 3.84	Std. Dev: 3.19
Monetary expenditure	Average: \$437	Std. Dev: \$400
Destination characteristics (for 210 destinations)		
Destination is an MSA	76.70%	
Ln (landarea in square miles)	Average: 5.92	Std. Dev: 2.88
Leisure employment per capita	Average: 0.12	Std. Dev: 0.14
Winter temperature (Fahrenheit)	Average: 42.3	Std. Dev: 16.53
Summer temperature (Fahrenheit)	Average: 82.03	Std. Dev: 8.79
Level of service characteristics (between 210 x 210 OD pairs)		
Highway distance (roundtrip)	Average: 2,622	Std. Dev: 1,749
Auto travel time (hours)	Average: 23	Std. Dev: 12.25
Air travel time (hours)	Average: 4.5	Std. Dev: 2.96
Auto travel cost (US dollars)	Average: \$154.24	Std. Dev: \$105.74
Air travel cost (US dollars)	Average: \$437.76	Std. Dev: \$301.69

Note. MSA = metropolitan statistical area

Table 7 Goodness of fit for models estimated in the study

	Log-likelihood (LL) at model convergence	No. of Parameters	Rho-square (\dots^2)	BIC	Predictive LL for 500 households
Models	(LL)	(K)	$1 - \frac{LL}{LL(C)}$	$-2LL + \ln(N)*K$	(PLL)
1. Time-constrained model for destination and mode choices	-10,620	21	0.14293	21,384	-6,484
2. Money-constrained model for destination and mode choices	-11,040	21	0.16368	22,224	-6,904
3. Time and Money constrained model for destination and mode choices	-10,048	21	0.17203	20,240	-6,114

Note. LL = Log-likelihood at model convergence; LL(C) = Log-likelihood with only constants in the model; K = No of parameters in the model; Rho-square (\dots^2) = $1 - \left\{ \frac{LL}{LL(C)} \right\}$; Bayesian Information Criterion (BIC) = $-2LL + \ln(N)*K$.

Table 8 Parameter estimates of the time- and money-constrained model

Baseline utility function (ψ_{jt}) specification	Parameter estimate	Std. err.
Constant for all destination-mode combinations	-8.614	0.110
Leisure Employment Per Capita at household location	-0.330	0.171
<i>Destination Specific Characteristics (Z_j)</i>		
Log of Land Area	0.170	0.006
Leisure Employment Per Capita at Destination	1.663	0.088
Dummy if destination in same state as HH residence	1.810	0.053
Dummy if destination in adjacent state to HH residence	1.156	0.044
Winter (January) temperature. 65°–75° Fahrenheit is base		
55°–65° Fahrenheit	-0.162	0.056
45°–55° Fahrenheit	-0.351	0.058
< 45° Fahrenheit	-0.378	0.066
Summer (June) temperature. 65°–75° Fahrenheit is base		
60°–65° Fahrenheit	-0.377	0.153
75°–80° Fahrenheit	-0.274	0.059
80°–85° Fahrenheit	-0.251	0.055
> 85° Fahrenheit	-0.310	0.059
<i>Mode Specific Characteristics (Z_{jt})</i>		
Alternative specific constant (Mode - auto as base)	-0.370	0.036
Origin is an MSA—on the air mode (auto is base)	0.085	0.016
Destination is an MSA—on the air mode (auto is base)	0.176	0.018
Round trip travel time (in days)	-0.332	0.023
<i>Scale parameter (σ) of the baseline utility function</i>	0.576	0.011
<i>Dissimilarity Parameter (θ)</i>	0.144	0.012
<i>Satiation Function (γ_j) Specification</i>		
Alternative specific constant	-3.242	0.010
Highway distance to destination (100's miles)	0.007	0.001

Note. MSA = metropolitan statistical area.

CHAPTER 4: A NATIONAL ANNUAL HOUSEHOLD LONG-DISTANCE PLEASURE-TRAVEL MODEL SYSTEM

4.1 Introduction

Pleasure is a major reason for long-distance travel in the U.S. Statistics from national travel surveys indicate that more than 55% of all long-distance travel is for pleasure purposes—sightseeing, recreation, relaxation, shopping, and visiting friends and relatives (BTS, 1997; BTS, 2003; McGuckin, 2009). In addition, long-distance pleasure travel (LDPT) garners particular attention due to its impact on the tourism and recreation industry. Therefore, LDPT has been studied extensively in the tourism literature and is slowly gaining explicit attention in transportation demand modeling.

Most long-distance travel demand analysis in the U.S. happens in the form of statewide travel demand models (Horowitz, 2008; Outwater et al., 2010) and inter-city travel analysis between specific city pairs (Bhat 1995; Koppelman and Sethi, 2005). A major gap in the US has been the unavailability of a nationwide long-distance travel demand model with the exception of a few recent attempts (Ashiabor et al., 2007; Baik et al., 2008; Epstein et al., 2008; Bradley et al., 2014; also see recent reviews by Zhang et al., 2012 and Van Nostrand et al., 2013). Several European countries have national travel models (see Lundquist and Mattson, 2002 for a review), including that for Sweden (Beser and Algers, 2001), Holland (HCG, 1990), Denmark (Foegerau, 2002), and the recently developed TRANS-TOOLS model system to analyze travel between European Union countries (Rich et al., 2009). Despite all these advances in modeling long

distance travel, most research in the transportation planning and travel demand modeling field has treated long-distance pleasure travel (LDPT) in limited ways, ignoring several behavioral aspects. A few important aspect addressed in this dissertation are discussed next.

4.1.1 Long-Distance Travel for VFR

As mentioned earlier, over half of all long-distance trips in the US are for pleasure. Among trips made for pleasure, most existing data suggests that more than half of these trips are for VFR (BTS, 1997; BTS, 2003) and the rest are for non-VFR leisure purposes such as sightseeing, outdoor recreation, rest/relaxation, and entertainment (from here on, for brevity, non-VFR leisure purposes are referred to as leisure). This is sufficient motivation to warrant studies focusing on VFR travel. However, most work in the transport modeling field does not explicitly recognize differences in travel behavior between VFR and leisure.

Contrary to the limited attention VFR travel has received in transportation modeling the tourism field has long recognized its importance. A seminal article by Jackson (1990) highlighted that the magnitude and influence of VFR travel has been underestimated and undervalued by tourism agencies. This led to the emergence of VFR travel as a new field of study in tourism. For example, as discussed in Becker (2011), a special edition of the *Journal of Tourism Studies* (1995) dedicated to this subject compiled studies on VFR travel in USA, Canada, Australia, and other countries. It highlighted the large share of its market share in their respective countries. Recognizing VFR travelers as an important source of tourism revenue (Braunlich and Nadkarni, 1995), led places and destination marketing organizations to have promotional campaigns separately for VFR visitors and leisure travelers. The differences in the expenditure patterns, travel and consumption behavior (Morrison et al., 1995; Backer 2009) of VFR and leisure travelers has been one of the reason for the above promotional campaigns. From

a transportation planning/modeling perspective, recognizing the behavioral differences between travel for different purposes helps in building a more robust capability to better plan the services and infrastructure to support pleasure travel.

4.1.2 Substitution Patterns

Apart from visiting friends and relatives, a majority of VFR travel involves some form of leisure (Quarmby, 2006), albeit some VFR travel is likely to be out of obligation. In addition, a considerable portion of VFR travel is likely to be motivated by leisure opportunities in those destinations (Backer 2009). Besides, it is common for obligatory VFR trips to be combined with leisure. Therefore, it is important to consider potential substitution between long-distance travel for VFR (which may or may not involve leisure as well) and travel exclusively for leisure. Doing so requires both VFR and leisure travel to be considered in a unified modeling framework. Indeed, whether VFR travel is of primary interest to the analyst or not, it is essential to consider VFR travel even if the expressed interest is only in analyzing long-distance leisure travel.

Another type of substitution effect is between long-distance pleasure travel (LDPT) and short-distance, day-to-day travel for leisure and socialization activities. Depending on accessibility to leisure activities and socialization within a close vicinity of a household's location, long-distance leisure travel of a household is likely to be traded off against local travel for leisure and social activities. Considering such substitution with local leisure travel helps in better estimating the travel demand for LDPT.

Further, leisure experiences at different destinations tend to be substitutable for each other, albeit there might be some destinations that provide unique experiences. At the same time, it is not necessary that households spend all their time and monetary budgets for pleasure at one single destination. Over a period of time, most people tend to visit multiple destinations due to

the variety of experiences offered by different destinations. This variety seeking nature combined with the substitutability of leisure experiences at different destinations leads to imperfect substitutability among destination choices for pleasure. For example, a household might visit multiple destinations for variety but not necessarily all destinations due to substitution effects and time/money constraints (Van Nostrand et al., 2013).

4.1.3 Timeframe of Analysis

Most literature on long-distance pleasure travel analysis is limited to shorter time frames such as a day or few weeks. However, long-distance pleasure trips are typically planned over longer time frames, as opposed to daily travel decisions for which shorter time frames may suffice. Analysis of the ATS data suggests that most households make less than 5 vacation trips per year. Typical one-day travel analysis provides a distorted picture of such infrequent activity. In this context, tourism and recreational demand literature has long suggested one year as appropriate for vacation travel analysis (Eugeno-martin 2003; van Middlekoop et al., 2004). Only recent studies in transportation modeling consider annual timeframe for long-distance pleasure travel analysis (LaMondia et al., 2008; Van Nostrand et al., 2013; Bradley et al., 2014; also see the holiday travel module of the TRANS-TOOLS model in Rich et al, 2009).

Consideration of an annual time frame comes with challenges in terms of modeling and analyzing long-distance pleasure travel. Specifically, most households tend to make multiple long-distance pleasure trips over a year. In the typical approach to modeling multiple trips over a given timeframe, the annual trip frequency is modeled first and then the travel choices (e.g., destination and mode choice) for individual trips are modeled separately for each trip. However, as discussed earlier, household decisions on long-distance pleasure trips over a year are likely to be interconnected due to time/money constraints, variety seeking, and substitution effects.

Furthermore, decisions of different travel dimensions—travel frequency, purpose, destination, mode, and length of stay—are likely to be made in conjunction with each other.

4.1.4 Current Research

The overarching goals of this research are two-fold. The first goal is to explore the merits of differentiating between long-distance pleasure travel for VFR and leisure purposes in the U.S. To this end, a descriptive analysis is undertaken to analyze household-level, annual long-distance travel characteristics collected from the ATS. While the data is about 20 years old now, it is the only widely available survey data with annual long-distance travel information. Furthermore, the authors' comparison of ATS data with more current data sources (NHTS) suggests that some aspects of travel behavior such as pleasure trip purpose distributions are relatively stable over time (also see BTS, 1997; BTS, 2003; NHTS 2001), while some other aspects such as duration of vacation trips have changed over time.

The second goal is to devise and test a household-level econometric model system that can be used to analyze households' annual long-distance pleasure travel patterns while considering the issues discussed earlier in this section. The model system includes the following household-level dependent variables of interest for analyzing LDPT demand – annual long-distance pleasure travel time budget, the destinations visited by the household (for pleasure) over an entire year, annual pleasure time allocated to each destination, travel purpose (VFR vs. leisure), travel mode, and trip frequency per year to each destination visited. Since the model system can be used to forecast all these variables at the household-level, the model system potentially can be used in a nationwide long-distance travel model system for forecasting pleasure travel OD flows in the U.S. for different demographic, land-use, and policy scenarios.

The proposed model system is a significant extension to Van Nostrand et al.(2013) model on simultaneously modeling household's annual leisure travel destination choices and the time allocation to each visited destination. First, the proposed model enhances the number of travel choices modeled, including the annual long-distance pleasure travel time budget (that was assumed to be exogenous in the previous studies) and the trip frequency to each visited destination. More importantly, it considers travel for VFR in addition to travel for leisure and sheds light on the differences in travel behavior between the two purposes. A household's VFR travel destination choices depend significantly on the spatial locations of the households' social contacts (friends and relatives), typically not gathered in travel surveys. To address this issue, aggregate migration data (from census) is explored as a surrogate measure for the spatial dispersion of households' social contacts for modeling VFR destination choices.

The empirical parameters of the proposed model system are estimated using data from the 1995 ATS. The model system is applied on a sample of the data to evaluate its predictions and to demonstrate its use for simulating the influence of a hypothetical policy scenario (of travel cost increase) on different choice dimensions associated with LDPT.

The remainder of this chapter proceeds as follows with Section 2 presenting a descriptive analysis, focusing on differences between VFR and leisure travel. Section 3 presents the modeling framework. Section 4 presents the model results along with policy simulations. Section 5 concludes this chapter.

4.2 Difference between VFR and Leisure Travel—A Descriptive Analysis

Out of over 80,000 survey households in the ATS sample, 48,527 reported at least one long-distance trip to a destination of 100 miles or greater distance from the household location; from 337,520 long-distance trips. Over half of these trips were made for pleasure purposes, most

of them to destinations within the U.S. and a small portion (3.5%) abroad; in this paper we focus on only domestic pleasure travel. Over half of all pleasure trips were made for VFR and the remaining for leisure. A majority of pleasure trips were undertaken by either car or air modes, with less than 5% by bus, rail, or water modes. Only households that traveled by car or air modes were considered in this analysis, because of the difficulty in gathering network data for the other modes.

Table 9 presents pleasure travel characteristics in the ATS data, separately for leisure and VFR to facilitate a comparison of travel behavior between the two purposes. These statistics are derived from a sample of 28,422 households who reported at least one pleasure trip; total 74,830 trips. 60% of these households reported at least one leisure trip over the year (total 34,088 leisure trips) and 65% reported at least one VFR trip (total 40,742 VFR trips). 25% of the households reported both leisure and VFR trips over the year, while the remaining 75% reported trips for only one purpose. This suggests the possibility of substitution effects between VFR and leisure travel.

In Table 9, descriptive statistics presented in the rows corresponding to “# Trips per annum per household” suggest that more than half of those who made pleasure trips made multiple pleasure trips in a year. However, the annual trip frequency for VFR (among households that made VFR trips) is higher than that for leisure. For both purposes, a significant proportion of households visited more than one destination per year (see the next set of rows), but a large percentage of them visited a destination, if they did so, only once (see the next set of rows corresponding to “# Trips made to a destination”). In other words, households are likely to visit multiple destinations per year and are less likely to re-visit a destination, suggesting variety seeking in household’s annual pleasure destination choices (Van Nostrand et al., 2013).

However, one can observe that the extent of variety seeking is higher for leisure than for VFR. Specifically, the number of destinations visited for VFR is smaller than that for leisure, presumably because the number of VFR destinations one could visit is limited by the presence (or absence) of friends/relatives at different destinations. On the other hand, the number of trips made to a destination (i.e., the extent of re-visiting a destination) is higher for VFR. Presence of friends/relatives at a place provides a greater motivation to revisit the place – for reconnecting with them, for potential cost savings on lodging expenses, and due to a greater familiarity.

In terms of annual time expenditure, households that made long-distance pleasure trips were found to spend an average of 10 days for pleasure travel (not reported in the table). Between the two pleasure purposes, households' time expenditure for VFR destinations (if they do so) is higher than that for leisure destinations.

The ATS did not collect information on the monetary expenditures for travel, lodging, dining, entertainment etc. at the destinations. For this study, these monetary expenditures were imputed using different sources, including: (1) the 1995 consumer expenditure survey data that includes such expenditures for out of town pleasure travel, (2) car travel costs estimated from reported mode choice, estimated travel distance, gasoline costs (from the US Energy Information Administration), (3) air travel costs estimated from Bureau of Transportation Statistics airline passenger ticket sample while considering travel party sizes reported in the ATS data, and carefully considered assumptions such as no lodging costs for VFR travel (see Pinjari and Sivaraman, 2013 for details on arriving at the monetary expenditures). It can be observed that the annual monetary expenditures for travel to VFR are considerably lower than that for travel to leisure, presumably because of the differences in the costs of lodging, dining and other expenses at the destinations (some of which are absorbed by the hosts for VFR travelers). In addition, it

was observed that the travel party size for VFR travel was in general smaller than the travel party size for leisure travelers, which would lead to differences in air travel costs.

The travel mode shares are also discernibly different between the two purposes as is the distribution of travel distances. The average travel distance for VFR destinations is higher than that for non-VFR (leisure) destinations. This could be due to households being captive to destinations where friends and relatives live (e.g., for the need to travel home for family occasions) that may be further away than leisure destinations. The differences in travel distances, in turn, along with lower party size for VFR travel, may be resulting in a greater share of VFR trips being undertaken by the air mode. All these differences have important implications to modeling long-distance leisure travel for these two purposes.

Long-distance trips for VFR and leisure purposes tend to exhibit different seasonal patterns. As expected, VFR trips peak during the last quarter with Thanksgiving and Christmas holiday seasons while leisure trips peak during the summer time.

When the households' VFR and leisure destinations were examined at the household level (not reported in the table), it was observed that a very small number of households visited the same destination separately for VFR and leisure purposes. Specifically, if a household visited a destination for VFR, the household would not visit the same destination for leisure in that year (and vice versa). This may very well be due to households combining leisure activities within their VFR trips thereby eliminating the need to visit the destination separately for leisure. Nevertheless, it is important to note that households are unlikely to visit a destination for one purpose (VFR or leisure) if they visited the destination for another purpose. This has implications to modeling pleasure destination and purpose choices in that the analyst must

consider strong substitution (or mutual exclusivity) between VFR and non-VFR leisure purposes across different trips to a destination.

Finally, as reported in Pinjari and Sivaraman (2013), households in the 1995 ATS were not changing their mode of travel across different trips to a same destination. The implication is that, a household's travel mode alternatives to a destination can be considered as strong substitutes for each other.

In summary, the descriptive analysis revealed interesting and importance differences in travel behavior between VFR and leisure travel purposes, as well as interesting substitution patterns in different travel choices for these two purposes. These findings inform the development of an appropriate modeling framework, as discussed in the next section, to analyze long-distance travel for VFR and leisure travel purposes.

4.3 Modeling Framework

The framework of the proposed household-level model system, as presented in Figure 28, comprises of three major steps. In the first step, household's annual time budget for long-distance pleasure travel is determined as a function of household socio-demographics and location characteristics. This model component takes the form of a stochastic frontier regression. In the second step, conditional on the annual long-distance pleasure travel (LDPT) time budget determined from the first step, annual long-distance pleasure travel destination choices, the purpose of travel to each visited destination, and the annual time allocated to each visited destination are modeled. All these choices are modeled simultaneously using a multiple-discrete continuous extreme value—multinomial logit (MDCEV-MNL) model. The use of MDCEV-MNL framework is guided by the findings from the descriptive analysis in the earlier section. Specifically, the MDCEV model component recognizes the variety seeking nature (or multiple

discreteness) in households' annual vacation destination choices, while the MNL component recognizes that the travel purpose alternatives for a chosen destination (VFR and leisure) are perfect substitutes in that households would not travel to a destination for one purpose if they visit that destination for another purpose. In the third step, conditional upon the destination, purpose, and annual time allocation choices, traditional discrete choice (MNL) models are used for determining the mode of travel and annual frequency of trips to each visited destination, as households were observed to not change their mode of travel across different trips to a destination. These models of mode and trip frequency choice are estimated separately for each purpose of travel (VFR and leisure) to recognize differences in mode choice and trip frequency choices by travel purpose.

As discussed earlier, all the above mentioned choices of a household are likely to be interconnected to each other. To recognize this, information predicted from the top-level model components is used to explain bottom level choices. In addition, log-sum variables are used to carry the information from bottom-level MNL models of mode and trip frequency choice to the MDCEV-MNL model of annual pleasure destination and purpose choice and time allocation model.

4.3.1 Stochastic Frontier Model of Household's Annual Long-Distance Pleasure Time

Budget

The household's LDPT time budget is estimated using a stochastic frontier regression approach. This approach has been widely used in firm-production economics (Aigner et al, 1977) for identifying the maximum possible production capacity (i.e., production frontier) as a function of various inputs. While the actual production levels and the inputs to the production can be

observed, a latent production frontier is assumed to exist. Such a production frontier is the maximum possible production that can be achieved given the inputs.

In travel behavior research, the stochastic frontier approach has been used to analyze: (1) the time-space prism constraints that people face (Kitamura et al., 2000), and (2) the maximum amount of time that people are willing to allocate to travel in a day (Banerjee et al., 2007). More recently, Augustin et al., (2013) used the approach to derive latent daily activity time budgets for analyzing individuals' out-of-home activity participation patterns.

Analogous to the above examples, we conceive a latent, household-level time frontier that governs households' annual long-distance pleasure travel choices. Such frontier, labeled the long-distance pleasure time frontier (LDPTF), is assumed as the maximum amount of time households are willing to allocate for long-distance pleasure travel in a year and, therefore, considered as the time budget governing households' long-distance pleasure travel. This frontier is, by design, greater than the observed total long-distance pleasure time expenditure (LDPTE), which is a sum of observed time expenditures for all long-distance pleasure destinations over a year for a household.

To explain the stochastic frontier model formulation, consider the following notation.

T_h = the observed total annual LDPTE for household h , assumed to be log-normally distributed;

\dagger_h = the unobserved LDPTF for household h , assumed to be log-normally distributed;

v_h = a normally distributed random component specific to household h , with variance σ_v^2 ;

u_h = a nonnegative random term assumed to follow half-normal distribution, with variance σ_u^2 ;

\mathbf{X}_h = a vector of observable individual characteristics; and β = a vector of coefficients of \mathbf{X}_h .

\dagger_h of a household is assumed to be a function of demographic and location characteristics, as:

$$\ln(\dagger_h) = \beta' \mathbf{X}_h + \epsilon_h \quad (42)$$

The unobserved LDPTF can be related to the observed LDPTE T_h as:

$$\ln(T_h) = \ln(\dagger_h) - u_h \quad (43)$$

Note that since u_h is nonnegative, the latent LDPTF is by design greater than observed LDPTE.

Combining Equations (42) and (43) results in the following stochastic frontier regression equation:

$$\ln(T_h) = \beta' \mathbf{X}_h + \epsilon_h - u_h \quad (44)$$

Once the model parameters are estimated (see Aigner et al., 1977 for details on estimating stochastic frontier regression parameters), using Equation 41, one can compute the expected value of LDPTF for household h as:

$$E[\dagger_h] = E\left[\exp(\beta' \mathbf{X}_h + \epsilon_h)\right] = \exp\left(\beta' \mathbf{X}_h + \frac{\sigma_v^2}{2}\right) \quad (45)$$

The expected LDPTF may be used as the time budget for the MDCEV-MNL model to analyze the household's annual, pleasure destination choice(s), the purpose of travel to each chosen destination, and annual time allocation for each chosen destination.

Before proceeding further, it is worth noting that one can model the observed LDPT expenditure (LDPTE), using the traditional regression approach, and use it as a budget in a subsequent MDCEV-MNL model. The problem with this approach, however, is that it is not easy (if not impossible) to incorporate the influence of destination-specific characteristics (such as leisure activity opportunities, costs of those activities, and costs of travel to destinations) in the LDPTE regression equation. As a result, one cannot accommodate the possibility that the LDPTE can change due to changes in destination characteristics. This is because the subsequent MDCEV-MNL model assumes the budget (which would be the estimated LDPTE in this case) as fixed and allocates it to different destinations. It would not allow the LDPTE to either increase or

decrease due to changes in destination-specific characteristics, but only leads to reallocations of the LDPTF among different destinations. Now, it is not easy to incorporate the effect of destination-specific characteristics in a stochastic frontier model of LDPTF either. However, as discussed in Augustin et al. (2013), an advantage of the stochastic frontier approach is that the LDPTF is by design greater than the observed LDPTF. Therefore, the budget estimated using the stochastic frontier approach provides a “buffer” for the LDPTF to increase or decrease. This can be easily accommodated in the second stage MDCEV-MNL model by designating an outside good that represents the difference between LDPTF and LDPTF.

4.4 MDCEV-MNL Model for Households’ Annual Choices of Long-Distance Pleasure-Travel Destinations (LDPT), Travel Purpose, and Time Allocation

Assume that each household makes its annual choices of LDPT destinations, travel purposes, and time allocations to maximize the following utility function:

$$U(\mathbf{t}) = \{\mathbb{E}_0 \ln t_o\} + \sum_{j=1}^J x_j \{\mathbb{E}_j \ln \{(t_j / x_j) + 1\}\}, \quad (46)$$

subject to time budget constraint:

$$t_o + \sum_{j=1}^J t_j = \dagger, \text{ and } t_o > 0 \text{ and } t_j \geq 0 \forall j = (1, 2, \dots, J). \quad (47)$$

In the above equations, $j (= 1, 2, \dots, J)$ is an index for long-distance pleasure destinations, whereas the index for households is suppressed for simplicity. \dagger is the annual LDPT budget (or frontier) for the household obtained from the stochastic frontier regression model and t_j is the annual household time allocated to the destination (includes travel to the destination and time spent at the destination). Note from the discussion earlier that the LDPT budget or frontier (\dagger) is

by definition greater than the LDPT expenditure ($\sum_{j=1}^J t_j$). Therefore, t_o is designated the outside good to accommodate any amount of annual LDPT budget that is not expended on long-distance pleasure travel.

In the utility expression in Equation 46, the term $x_j \mathbb{E}_j \ln \left\{ \left(t_j / x_j \right) + 1 \right\}$ captures the utility accrued by the household from allocating t_j amount of time for destination j . \mathbb{E}_j is the baseline marginal utility of time allocation to destination j (i.e., marginal utility at zero time allocation to destination j). Between two destination alternatives, the alternative with greater baseline marginal utility is more likely to be chosen. In addition, \mathbb{E}_j influences the time allocation to destination j , since a greater \mathbb{E}_j value implies a greater marginal utility of time allocation. The parameter x_j allows corner solutions (i.e., the possibility of not choosing a destination) and differential satiation effects (diminishing marginal utility with increasing consumption) for different destinations. Specifically, when all else is same, an alternative with a greater value of x_j will have a slower rate of satiation and therefore a greater amount of time allocation.

Recall from earlier discussion that the household visits a destination (if it does so) for only one travel purpose (VFR or leisure) regardless of the number of times the household visits the destination. To accommodate travel purpose within the modeling framework, let the index for VFR travel purpose be f and let the index for leisure travel purpose be l . Further, for each destination j , let \mathbb{E}_j be defined as (Bhat et al., 2009):

$$\mathbb{E}_j = \exp \left(u_{jf} W_{jf} + u_{jl} W_{jl} \right), \quad (48)$$

where, W_{jf} and W_{jl} are the utilities perceived by the household for visiting destination j for travel purposes VFR and non-VFR leisure, respectively, and u_{jf} and u_{jl} are dummy variables taking a value of 1 if the corresponding travel purpose is the reason why the household visits destination j . Note that $u_{jf} + u_{jl} = 1$ since only travel one purpose is chosen.

With the above definition of Ξ_j , as discussed in Bhat et al. (2009), since the household chooses only one travel purpose for each destination visited, the functional form of $U(\mathbf{t})$ implies that the household will choose the purpose that provides the maximum utility for each destination j in the process of maximizing $U(\mathbf{t})$. That is, $u_{jf}W_{jf} + u_{jl}W_{jl} = \max(W_{jf}, W_{jl})$, or $\Xi_j = \exp\left(\max(W_{jf}, W_{jl})\right)$. Therefore, the utility function in Equation 46 can be rewritten as:

$$U(\mathbf{t}) = \{\Xi_0 \ln t_0\} + \sum_{j=1}^J \left\{ x_j \exp\left(\max(W_{jf}, W_{jl})\right) \ln\left(\frac{t_j}{x_j} + 1\right) \right\} \quad (49)$$

To complete the model specification, let $\Xi_0 = \exp(v_0)$, $W_{jf} = (w'w_{jf} + v_{jf})$, and $W_{jl} = (w'w_{jl} + v_{jl})$, where $w'w_{jf}$ and $w'w_{jl}$ are the observed utility components corresponding to the choice of destination j for VFR and leisure purposes, respectively, and v_{jf} and v_{jl} are the corresponding unobserved components. In the observed utility components, w_{jf} and w_{jl} are the vectors of variables influencing the choice of destination j and the travel purpose. These vectors include (1) household sociodemographics explaining the choice of one purpose vs. the other, (2) destination-specific variables, such as leisure activity opportunities and weather, explaining the choice of destinations, and (3) destination and purpose-specific log-sum variables from a bottom-level discrete choice model of travel mode and trip frequency choice to carry the information on

travel times and travel costs by different modes of travel for different travel purposes to each destination.

In the above formulation, w is a vector of coefficients associated with aforementioned variables. Note that for variables that vary across destinations (i.e., destination-specific variables), some of the coefficients in w might be common (i.e., equal) for both purposes, while other coefficients might be different by travel purpose. If the coefficient of a destination-specific variable is different between the two travel purposes, it is implied that the variable has a different influence on the destination choice depending on the purpose of travel. For variables that do not vary across destinations, such as household sociodemographics, for identification, only one purpose-specific coefficient is estimated while considering the other purpose as the base category. Finally, the error terms v_{jf} and v_{jl} are assumed to be Gumbel and independent and identically distributed across different destinations. For each destination j , however, the error terms may share common unobserved destination-specific attributes generating correlations between v_{jf} and v_{jl} . Thus, for each destination j , the following distribution may be used:

$$F(v_{jf}, v_{jl}) = \exp \left\{ - \left[e^{-v_{jf}/\sigma} + e^{-v_{jl}/\sigma} \right]^\sigma \right\} \quad (50)$$

where σ is the dissimilarity parameter indicating the level of correlation between v_{jf} and v_{jl} .

Given the above utility specification and error term distributional assumptions, following Bhat et al. (2009), one can derive the MDCEV-MNL model probability expression for the households' choice of annual long-distance pleasure travel destinations, travel purposes, and time allocations. The MDCEV component of such probability expression is for the annual destination choice(s) and corresponding time allocations while the MNL component of this expression is for the choice of travel purpose for each destination. In addition to the above, a

third supplemental component to this model system includes multinomial logit models of travel mode choice and trip frequency, conditional on the annual vacation destination choices and the travel purposes. Two different logit models of travel mode and trip frequency choice are estimated separately for each purpose of travel (VFR and leisure) to recognize differences in mode choice and trip frequency choices by travel purpose.

4.5 Empirical Analysis

The empirical parameters of the proposed model system were estimated using a random sample of 2665 households from ATS sample. The sample characteristics are not discussed here as they are similar to that of the larger sample discussed before. To define the destination choice alternatives, the U.S. was divided into 210 destinations comprising 162 MSAs and 48 non-MSA destinations. Travel purpose was considered in the form of a binary variable with VFR and leisure categories. For each destination, auto and air were considered as the two primary modes of travel. The maximum possible trip frequency for each travel purpose was limited based on the mode of travel. Higher trip frequency was allowed for VFR trip purpose (than leisure purpose) and for travel by car mode (than by air mode). In addition to the ATS, several secondary data sources were used to derive level of service variables (travel times and costs) and lodging, dining, and entertainment costs at destinations, and a variety of destination characteristics (See Van Nostrand et al., (2013) and Pinjari and Sivaraman (2013) for details).

4.6 Model Estimation Results

4.6.1 Stochastic Frontier Model Estimates for Households' Annual Long-Distance

Pleasure Time Budgets

The household-level annual long-distance pleasure travel (LDPT) time budget was estimated using the stochastic frontier model (Table 10). Households with younger and older householders were found to have higher LDPTFs than those with middle age (25-45 years) householders. This is perhaps because middle age group householders tend to have greater time constraints due to familial and career responsibilities. This was also evident from Hu and Morrison (2002), who identified elder households to spend more time, traveling to multiple destination compare to other age groups, likely constrained due to familial or economic reasons. Households with better educated householders were associated with greater LDPTFs perhaps due to a more stable full-time employment (than the less educated group) and employer-perks such as annual paid vacation time to pursue long-distance pleasure travel. LDPTFs increased with household income levels due to increase in the amount of disposable income for long-distance pleasure travel. Households living in metropolitan areas showed larger LDPTFs than those living in rural areas perhaps due to a greater accessibility and less expensive travel options for long-distance travel. While across census regions, households in the west and mid-west are estimated to have larger LDPTF than northeast and south, likely due to the general need to travel further for pleasure opportunities in these regions.

4.6.2 Annual LDPT Destination Choice, Purpose Choice, and Time-Allocation Model

The MDCEV-MNL model of annual LDPT choices is presented in Table 11. The first two variables in the baseline utility function, with a common coefficient on all destinations and purposes (VFR and leisure) capture the extent to which the time budget (LDPTF) is expended for

long-distance pleasure travel. Households with greater leisure opportunities in the vicinity tend to spend less time for LDPT (as captured by the negative coefficient on leisure employment per capita at household location). This result highlights substitution between local leisure activities and the time spent on LDPT.

The next two variables vary across destinations (influencing destination choice), but their coefficients do not vary across travel purposes. Among them, the logarithm of land area is included as a control for the size of the destinations. Next, the coefficient on the costs of lodging/dining/other activities suggest, as expected, that pricy destinations are less preferred.

The next set of variables specific to destinations influences the choice of purpose. Among these, the log-sum variables from the bottom-level, travel mode and trip frequency model carry information of travel times and costs into the destination and purpose-specific utilities. Given these level-of-service (LOS) variables have negative coefficients in the bottom-level models, the positive coefficients on the log-sum variables suggest that increasing travel time (or cost) to a destination decreases the attractiveness of that destination. Between the two purposes, the corresponding coefficient is smaller in magnitude for VFR. This suggests that the destination choices are relatively less sensitive to travel impedance for VFR compared to that for leisure. This echoes the finding in the descriptive analysis that travel distances for VFR are greater than that for leisure and that air travel mode share is higher for VFR travel. Coefficients on the next three dummy variables – destination in the same/adjacent state as household residence and travel distance greater than 3000 miles – also suggest that travel for VFR is more likely to involve travel to farther destinations than for leisure.

Leisure employment per capita at the destination is used in the model as a surrogate for leisure opportunities at the destination, with coefficients suggesting that destination choice for

pleasure travel is positively influenced by leisure opportunities, albeit, with a greater influence on travel for leisure. As discussed earlier, VFR travel is likely to combine leisure and is therefore also influenced by the presence of such opportunities.

The next set of variables on temperatures reveal interesting differences in the choice of destinations for VFR and leisure travel. Particularly, households are observed to be less averse to travel to destinations with more extreme temperatures for VFR when compared to that for leisure, presumably because the primary motivation for former is visiting (sometimes due to obligation) with less regard to weather.

We explored the use of migration volumes between household location and destinations as a potential surrogate for spatial dispersion of social networks. As cited in previous studies (Jackson, 1990; Nadkarni and Braunlich 1995), households showed higher propensity to undertake pleasure travel to destination with stronger social ties (migration volume), be it for VFR or leisure, albeit the former being more influenced by it. Further, households are found to have higher propensity to travel to distant destinations (in non-contiguous states) contingent on stronger social ties. More research is warranted to understand these effects and better ways to incorporate social network information in such models.

Among demographics, low-income households are more likely to travel for VFR, presumably to save the costs of lodging etc. through the hospitality of hosts at the destination. Households with full-time employed householder are less likely to travel for VFR perhaps because they have a steady income allowing them to take more expensive vacations for leisure. Non-family households are more likely to travel for VFR purpose, presumably due to a greater need for socialization with friends and relatives. The satiation function includes one variable – distance to the destination from the household location on the highway network. The positive

coefficient on this variable suggests that annual time allocated to farther destinations tend to be higher (if such destinations are traveled to) than the time allocated to closer destinations (see Nadkarni and Braunlich, 1995 for a similar finding).

Overall, the model parameters have intuitive interpretations and shed light on the determinants of household choices of annual LDPT, including the potential usefulness of migration data in explaining LDPT choices.

4.6.3 MNL Model of Travel Mode and Trip Frequency

This section details a joint model of travel mode and trip frequency, conditional upon the earlier modeled choices, estimated separately for VFR and non-VFR trip purposes. Households, in potential could make recurrent trips for either purpose to a destination, as was evident from prior studies such as Meis, Joyal and Trites (1995) and Sidrelis and Moore (1998). The former exclusively assessing VFR travel to Canada, and latter focusing on recreational leisure travel to lakes with each of them being undertaken in different time frames. The aforementioned travel decisions are examined in a comprehensive manner in this paper, as similar travel patterns were observed to occur within the U.S. for each of the pleasure purpose in an annum, as detailed in the section on data descriptive. Table 12 presents the model estimates, which in addition to estimating alternative specific constants, also estimates parameters for the level-of-service variables such as travel times and travel costs (expressed in logarithmic form) likely to influence decision to make one or more trips via a given travel mode to a destination. As expected faster and cheaper travel modes were preferred over slower and expensive travel modes, with trip frequency found to decrease with increasing travel time and travel cost. Further, the sensitivity to travel cost decreased with increase in household income, captured via an interaction variable between log-travel cost and income. A dummy variable for origin or destination being a MSA

suggested that the air mode was more preferred for travel from or to metropolitan locations than non-MSAs.

4.7 Demonstration of Model Application

The above described model system was applied to a sample of 1000 households that were not used for model estimation to predict the annual LDPT choices modeled in this study. To predict the annual LDPT budget (i.e., LDPTF) for each household, the expression in Equation (46) was used. The predicted LDPTF is used to apply the MDCEV-MNL model for predicting households' annual destination choices, travel purpose and time allocation for each visited destination. The MDCEV-MNL predictions were performed by first completing the MDCEV - component predictions of annual destination choices and corresponding time allocation (using the log-sum variable from MNL component of travel purpose choice as a variable in the MDCEV model utility equations). Conditional on the MDCEV predictions, the MNL-component was applied to predict the choice of travel purpose for each predicted destination. Subsequently, the MNL model of travel mode and trip frequency was applied to predict the mode and trip frequency for each chosen destination given the purpose of travel.

Using the model predictions, one can build an OD table of trip flows and construct trip length distributions for different travel purposes as shown in the top part of Figure 29. It can be observed from comparing the observed and predicted (base case) trip length distributions that the predicted trip length distributions by travel purpose are reasonably close to the observed distributions; albeit the predictions show over (under)-prediction of long (short)-distance trips. When the travel costs were tripled from the base case, the trip length distribution shows a clear shift toward shorter length trips. This is a reflection of households increasing (decreasing) their trip frequency to closer (farther) destinations (because the travel cost variable appears in the

travel mode and trip frequency model) as well as a choice of destinations that are more closer (because the log-sum variable from the travel mode and trip frequency choice informs the travel costs to the MDCEV-MNL model of destination and travel purpose choice).

Figure 30 shows the distributions of household-level annual time expenditure (no. of days) for the destinations visited by households. Again, the predictions match reasonably well with observed distributions; albeit an over prediction of longer time investments. Increasing the travel costs shows a clear decrease in the time investment as well. This is due to a combination of decrease in the number of trips to a destination and decrease in the amount of annual time spent at each destination due to the same reasons as mentioned earlier for trip length predictions.

4.8 Summary

The proposed makes a two-fold contribution to the literature on long-distance pleasure travel modeling for transport planning. First, it highlights the importance of distinguishing travel for visiting friends and relatives (VFR) from leisure travel purposes and sheds light on the differences in long-distance travel between the two purposes. VFR travel is different from leisure travel in many ways, including the extent of travel (over half of all pleasure travel is for VFR), the extent of variety seeking in annual destination choices, distances traveled, time allocation, monetary expenditures, travel mode shares, trip frequency, travel party size, and seasonality of travel. In addition, VFR travel is found to have interesting substitution patterns with leisure travel. Specifically, the vast majority of households reporting pleasure travel in ATS sample did not visit a destination for one purpose (VFR or leisure) if they traveled to the same destination for another purpose.

The second contribution is the formulation of a household-level econometric model system that can be used to analyze households' annual long-distance pleasure travel choices

while considering an annual analysis time-frame, multiple discreteness in annual pleasure destination choices, and substitution between VFR and leisure travel. The model system includes the following household-level dependent variables of interest for analyzing long-distance pleasure travel demand – annual pleasure time budget, the destinations visited over an entire year, pleasure time allocated to each destination, the purpose for which destination is visited (VFR or leisure), the mode of travel (car or air), and frequency of trips per year to each destination visited.

The model estimated using ATS (1995) data sheds light on various determinants of long-distance pleasure travel choices, including the potential usefulness of migration data for modeling VFR and leisure destination choices. It also revealed differences in travel behavior for these two purposes and supported the need for distinguishing them. Prediction exercises and policy simulations with the model demonstrate its potential applicability in a national long-distance travel model system for forecasting nationwide pleasure travel demand.

This research can be strengthened in different ways. First, more extensive validation and testing of the proposed model system (perhaps with more recent data) is required before the model system can be used for practical forecasting or policy assessments. Second, current research assumes travel party size as exogenous information to estimate travel costs by the air mode. To address this, travel party size could be incorporated as an additional model component. Third, all leisure travel purposes—recreation, entertainment, sightseeing etc.—were lumped into one category. A finer categorization of travel purpose would be useful. Fourth, the proposed modeling framework was estimated sequentially, with log-sum variables providing feedback from lower-level choices to upper-level choices. Joint estimation of the entire model system would be an interesting undertaking.

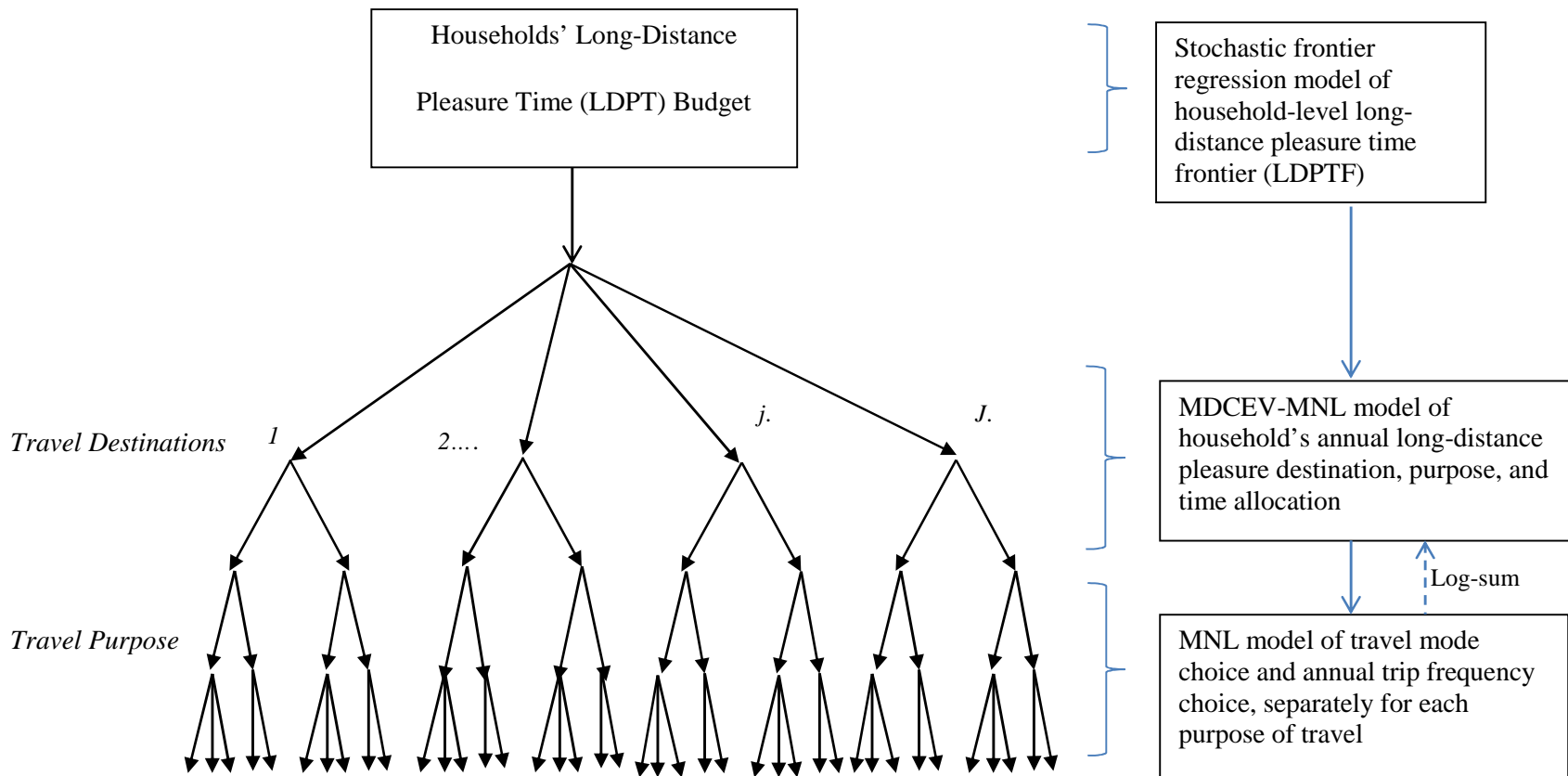


Figure 28 Structure of the household-level annual long-distance pleasure-travel-demand model system.

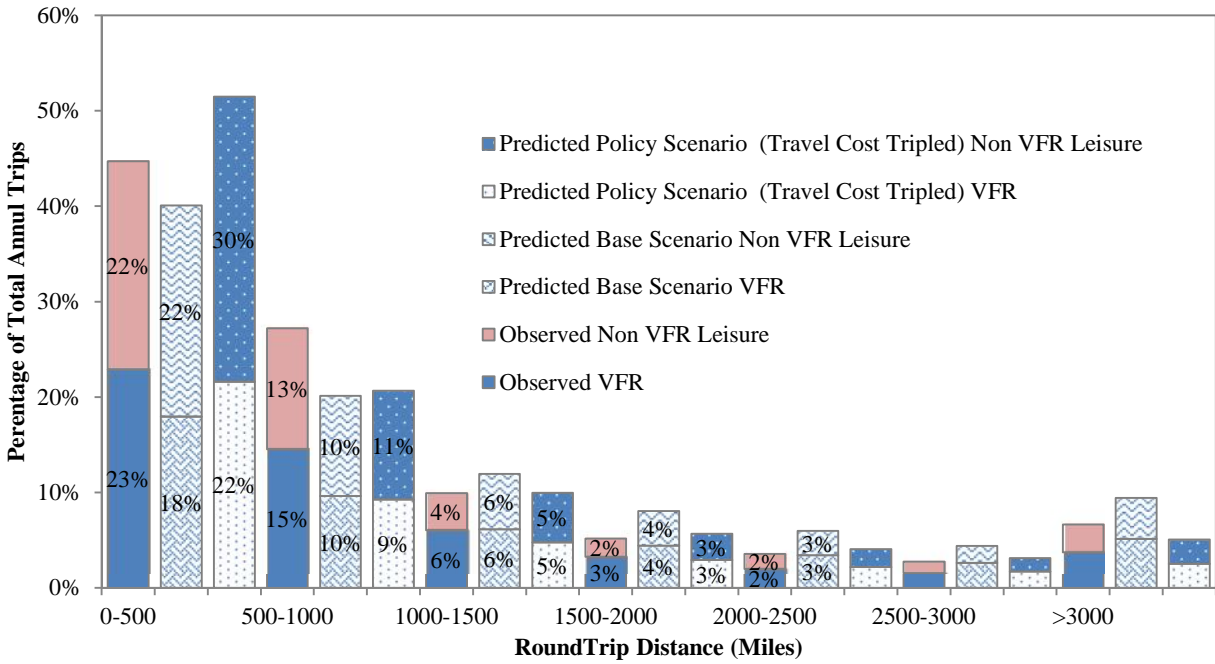


Figure 29 Prediction exercise results (annual trips) using the proposed model system.

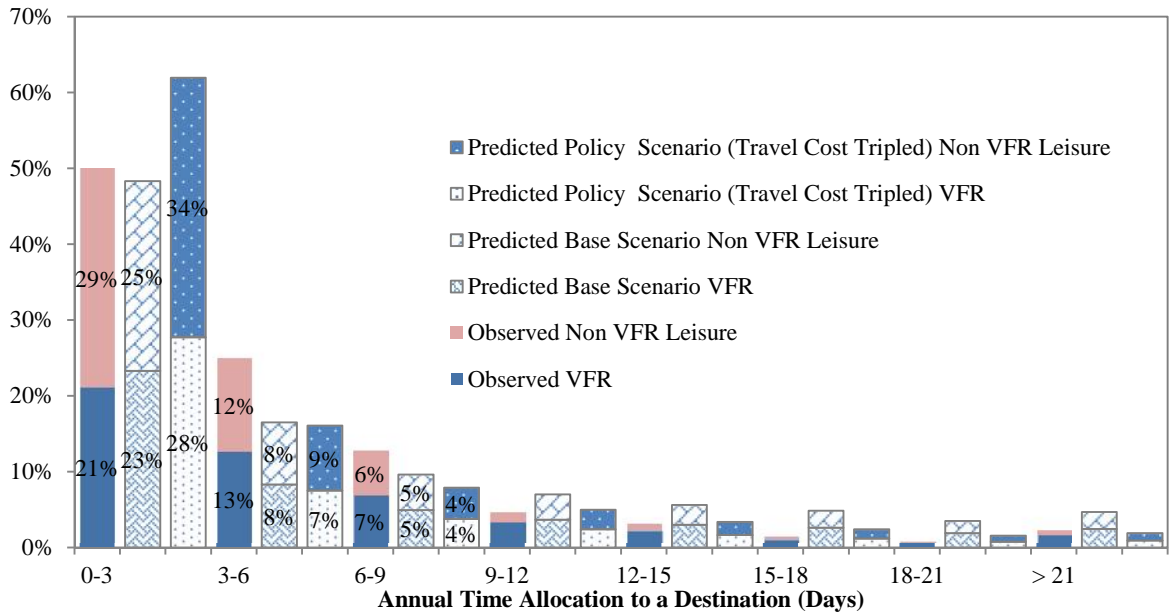


Figure 30 Prediction exercise results (annual time allocation) using the proposed model system.

Table 9 Descriptive statistics of long-distance non-visiting friends and relatives leisure and visiting friends and relatives travel characteristics in the 1995 American travel survey data

	For non-VFR leisure		For VFR	
Total # households in the sample making trips	17,821		18,580	
Total # trips in the sample	34,088		40,724	
# Trips per annum per household	Avg. = 1.97		Avg. = 2.19	
1	54.6%		50.2%	
2	22.0%		21.9%	
3	10.7%		11.0%	
4	5.4%		6.4%	
5 or more	7.3%		10.5%	
# Destinations visited per annum per household	Avg. = 1.50		Avg. = 1.39	
1	66.1%		70.8%	
2	22.9%		21.8%	
3	7.8%		5.7%	
4 or more	3.2%		1.7%	
# Trips made to a destination per annum	Avg. = 1.32		Avg. = 1.58	
1	82.7%		72.7%	
2	10.4%		14.1%	
3	3.5%		5.8%	
4 or more	3.4%		7.4%	
Annual household total time expenditure	Avg. = 6.69 days		Avg. = 8.62 days	
Annual household time expenditure for a destination	Avg. = 4.48 days		Avg. = 6.21 days	
Household time expenditure per 1 trip	Avg. = 3.40 days		Avg. = 3.93 days	
Annual household total monetary expenditure	Avg. = \$786		Avg. = \$397	
Annual monetary expenditure for a destination	Avg. = \$525		Avg. = \$286	
Household monetary expenditure per 1 trip	Avg. = \$399		Avg. = \$181	
Primary mode of transportation	Auto: 84%	Air: 16%	Auto: 77%	Air: 23%
Average travel party size per household trip	Auto: 2.04	Air: 1.74	Auto: 2.01	Air: 1.49

Table 9 (Continued)

# Trips per annum to a destination by mode of travel	For non-VFR leisure		For VFR	
	Auto	Air	Auto	Air
1	80%	94%	68%	88%
2	12%	5%	16%	9%
3	4%	1%	7%	2%
4 or more	4%	0%	9%	1%
Travel distance (miles)	Avg. = 507 miles		Avg. = 598 miles	
80–500	71%		62%	
500–1000	15%		19%	
1000–2000	11%		14%	
> 2000	2%		5%	
Seasonality of household trips				
% of trips ending in 1st quarter (Jan–March)	17%		19%	
% of trips ending in 2nd quarter (April–June)	26%		24%	
% of trips ending in 3rd quarter (July–September)	39%		27%	
% of trips ending in 4th quarter (October–December)	19%		30%	

Note. VFR = visiting friends and relatives.

Table 10 Parameter estimates for annual out-of-home time frontier for pleasure

Variables	Est.	T-Stat
Constant	2.5916	38.808
15–24 year old householders (25–45 age as base)	0.1182	1.187
46–64 year old householders (25–45 age as base)	0.1579	3.719
65 year and older householders (25–45 age as base)	0.2809	5.079
Householders with an associate’s degree or more	0.2428	6.069
African American (White and others as base)	-0.2293	-2.907
Hispanic (other ethnicities as base)	-0.3167	-3.35
Low-income households (< 25K per annum, medium income as base)	-0.0768	-1.727
High-income households (> 75 K per annum, medium income as base)	0.0910	1.465
Metropolitan household	0.0723	1.897
Midwestern household (Northeast census division as base)	0.1292	2.212
Southern household (Northeast census division as base)	0.0938	1.725
Western household (Northeast census division as base)	0.2671	4.604
Lambda (efficiency)	2.5213	14.013
Sigma-u		1.45
Sigma-v		0.57
Log-likelihood at constants	-3876.24	
Log-likelihood at convergence	-3816.57	

Table 11 Multiple discrete–continuous extreme-value multinomial logit model of long-distance pleasure-travel destination, purpose, and time allocation

	VFR		Non-VFR Leisure	
Baseline utility () specification	Est.	T-Stat	Est.	t-Stat
<i>Variables common to all destinations and purposes</i>				
Constant	-6.488	-40.16	-6.488	-40.16
Leisure employment per capita at the household location	-0.729	-4.01	-0.729	-4.01
<i>Destination characteristics with common effect across purposes</i>				
Logarithms of land area of the destination	0.269	44.07	0.269	44.07
Per-night cost of lodging/dining/other activities at destination (\$100's)	-0.119	-3.63	-0.119	-3.63
<i>Destination characteristics w/ differential effect across travel purposes</i>				
Log-sum variable from mode and trip frequency model	0.487	17.56	0.585	21.10
Dummy if destination in same state as household (HH) residence	1.450	17.75	1.705	24.16
Dummy if destination in adjacent state to HH residence	1.119	20.19	1.251	23.46
Round trip Distance to destination greater than 3000 miles (dummy)	-1.08	-13.51	-1.451	-16.69
Leisure employment per capita at destination	1.953	9.83	2.651	21.99
Winter (January) temperatures (monthly avg. of max daily values) with 67°–75° Fahrenheit as base				
55°–65° Fahrenheit	-0.528	-6.08	-0.512	-6.71
45°–55° Fahrenheit	-0.640	-7.48	-0.949	-10.57
35°–45° Fahrenheit	-0.749	-8.02	-1.163	-11.35
< 35° Fahrenheit	-0.764	-7.85	-1.115	-10.37
Summer (June) temperatures (monthly avg. of max daily values) with 65°–75° Fahrenheit as base				
60°–65° Fahrenheit	-1.492	-4.11	0.019	0.08
75°–80° Fahrenheit	-0.546	-7.33	-0.593	-7.81
80°–85° Fahrenheit	-0.448	-5.91	-0.395	-4.99
85°–90° Fahrenheit	-0.759	-9.42	-0.763	-9.07
> 90° Fahrenheit	-0.640	-7.38	-0.546	-5.78
Origin-Destination (OD) migration volume (1000's) * O&D in same or contiguous state	0.006	2.66	0.005	2.56
OD migration volume (1000's) * O&D in non-contiguous state	0.022	6.09	0.020	5.40

Table 11 (Continued)

	VFR	Non-VFR Leisure	
<i>Household demographic characteristics</i>			
Low income (VFR)	0.202	3.69	— —
Nonfamily household (VFR)	0.093	1.84	— —
Householder is Employed Full Time (VFR)	-0.284	-5.37	— —
Alternate specific constant for VFR	-0.176	-1.03	— —
<i>Dissimilarity parameter</i> μ (t-stat is against a value of 1)		0.871 (t-stat = 7.05)	
Satiation Function () Specification		0.123 (t-stat = 45.74)	
Highway distance to destination (100's miles)			
Model Fit Measures			
Log-likelihood at convergence : $L(\hat{S}, \hat{\mu})$		-38,945	
Constants only		-45,801	
Rho squared = $1 - \{L(\hat{S}, \hat{\mu}) / L(0)\}$		0.1498	

Note. VFR = visiting friends and relatives.

Table 12 Trip and mode choice model specification by purpose

Variable descriptions	VFR		Non-VFR Leisure	
	Est.	<i>T</i> -Stat	Est.	<i>T</i> -Stat
Alternate specific constant: 1 Ground trip	0	0	0	0
Alternate specific constant: 2 Ground trip	-0.487	-5.89	-0.977	-11.12
Alternate specific constant: 3 Ground trip	-0.512	-4.25	-1.34	-9.31
Alternate specific constant: 4 Ground trip	-0.49	-3.19	-1.45	-7.42
Alternate specific constant: 5 Ground trip	-0.636	-3.25	—	—
Alternate specific constant: 1 Air trip	-1.02	-4.83	-2.32	-11.37
Alternate specific constant: 2 Air trip	-2.72	-10.19	-4.86	-15.65
Alternate specific constant: 3 Air trip	-4.14	-10.16	—	—
Log travel cost in dollars	-0.677	-12.24	-0.437	-9.93
Travel time in days	-2.43	-14.4	-4	-16.17
High income * log travel cost in dollars	0.117	1.57	0.238	3.55
Origin or destination is a MSA (Ground as base)	0.623	3.64	0.548	3.21
Number of cases		2,226		2,248
Log-likelihood with only constants		-3013.4		-2195
Log-likelihood at convergence		-2660.7		-1895.6
Rho-square		0.425		0.529
Adjusted rho-square		0.423		0.527

Note. VFR = visiting friends and relatives; MSA = metropolitan statistical area.

CHAPTER 5: CONCLUSION AND FUTURE WORK

Long distance travel within the U.S., which has had modest recognition in transportation until the recent past, has gained increased recognition in the recent past with the need to understand the viability of implementing mass transit options between cities and regions in the U.S. However, there exist no current data to understand such travel within the U.S, with the last long distance travel survey undertaken in 2001 (NHTS, 2001). One of the main reasons for not undertaking such survey has been the significant amount of resource required to undertake it, resulting from the need to collect such data over longer period (year). Further, there exists no national model to understand such household travel behavior, and those that exist are trip level models, constructed using the traditional four-step urban model framework.

This dissertation as its main objective, sought to first develop a comprehensive behavioral framework to model annual household long distance pleasure travel, which makes up more than ½ of all long distance travel (ATS, 1995). To accomplish this objective, this dissertation first analyzed data from historical surveys with different collection methodologies, and supplemental datasets to inform the development of a behavioral framework. In the due course, this dissertation also proposed an alternate method to weight (expand) the sample long distance trip data collected in surveys with shorter recall period. It was proposed because the existing weighting methods for surveys with shorter recall period (e.g.: NPTS, 1995) do not directly yield annual travel estimates. Further, this approach could be help us arrive at aggregate annual travel estimates similar to ATS (1995), without investing as much resource as required for ATS (1995) like survey. The following sections summarize the above efforts in more detail.

5.1 Data Collection and Measurement

Overall, long distance travel tends to be diverse in terms of its characteristics: trip length, trip frequencies, purpose. This makes it difficult for surveys to collect all such trips within a single framework. An ideal collection method would involve collecting such data over a year to obtain comprehensive information on household travel. However, this also makes the collection effort a resource intensive task for surveyors and significant burden for its respondents.

The ATS (1995) was last such effort, and is still being used, in particular to understand, annual household travel in terms of their choice of destination, mode, trip frequencies etc. It is also used in this dissertation to construct a behavioral model of annual long distance pleasure travel, summarized in the subsequent sections in this chapter. There have been other efforts to collect long distance travel with shorter recall and collection period (single wave) (NPTS, 1995; NHTS, 2001). These datasets although not as comprehensive as ATS (1995), do gather sufficient data to arrive at aggregate estimates of annual demand for such travel, but has its own drawback, and requires significant post data processing to arrive at annual travel estimates.

5.1.1 Long and Short Period of Data Collection

In the due course of analyzing multiple travel surveys to construct a theoretical framework, the data collected using shorter recall period (e.g.: NPTS, 1995) were found to result in inappropriate annual estimates of long distance travel. More importantly, these estimates were getting affected due to single wave of data collection, and pursuing multiple waves could significantly increase the resource required to undertake such effort. Thus, this research proposed an alternate method to estimate annual travel from data collected using NPTS (1995): single wave of two-week data collected from each sample respondent. This dataset was considered as it could be validated against ATS (1995), and was collected around the same time-period compared

to the more recent NHTS (2001). However, before arriving at comparable estimates, numerous definitional issues had to be addressed to make consistent comparison with ATS (1995). In particular, the trip length definition had to be modified to consider only those trips that are 100 or more miles in straight-line distance across both datasets. Further only person trips over 5 years and older were considered for comparison from ATS (1995), as NPTS (1995) only reports the trips of individuals 5 years and older.

Further, more importantly, the weights had to be modified, as the current NPTS (1995) period weights is estimated by applying a straight-forward period factor to the daily person trip weights to arrive at annual long distance travel estimate. This weighting method does not yield annual estimates (NPTS handbook, 1998). Specifically, in this method, a multiplier (365/14 day collection period) is applied to the daily person trips weights before it is applied to the sample trips reported in the two-week period. In essence, this estimate suggests the population represented by sample in two week to exhibit the same travel behavior during the rest of the 25 two-week periods in a year. The above might not hold true for pleasure travel, which tends to be very seasonal and accounts for more than ½ of all annual long distance trips. Hence, an alternate weighting method is proposed in this dissertation to arrive at annual estimate.

This approach to expand the sample to arrive at annual estimate did lead the revised NPTS (1995) estimates to be within 6% of the annual estimate from ATS (1995). However, it still is significantly lower than ATS (1995), and more importantly NPTS (1995) trip length distribution revealed far less long distance trips to distant (long haul /over 200 miles) destinations than ATS. One of the reasons for under-representation of distant long distance trips could be attributed to the under-sampling of the population from the west. This was evident through the comparison of the sample distribution of ATS and NPTS against the 1995 population

across census region, which revealed significant under-representation of the population from the western region. Further, western households, more so than households from other regions were identified to travel the furthest, as was evident from the analysis of ATS (1995) and BLS consumer expenditure (CEX) data by region.

Besides the above, the other major reason for the lower share of long distance trips to distant (long haul) destination could be attributed to the recall period. The ATS (1995) had a longer three-month recall period compared to NPTS (1995) having a two-week recall period. As is intuitive, a shorter two week recall period, might in effect lead to under reporting of occasional (infrequent) trips to distant destination (e.g.: sightseeing, rest/relaxation), unless the recall period encompasses holidays. In contrast, the two - week recall period increases the likelihood of reporting more of the routine trips to short haul destinations (e.g.: commute to work, shopping etc.). This was evident, with NPTS (1995) reporting significantly more short haul trips (within 200 miles). While the ATS (1995) reports significantly more trips to long haul destinations (over 500 miles). To this effect, as Giesbrecht and Bose (2005) suggested, it would be appropriate to have multiple recall periods, with shorter (longer) recall for short (long) haul trips.

5.1.2 Travel Motivations and Patterns

The reporting of long distance trips besides being affected by period of collection and recall period are also bound to be affected by the definition and enumeration of travel purposes. For instance, pleasure trips such as those made for visiting could exhibit distinct trip rates and trip length. For example, a household could make frequent VFR trips to socialize with out-of-town friends over weekends, and make an occasional VFR trip to distant relative over traditional holidays (Thanksgiving and Christmas). Similarly, leisure trips could be distinct in terms of trip length based on specific motivations. For example, a household could make frequent short haul

leisure trips for outdoor recreation, entertainment or shopping, and make an occasional long haul leisure trips for sightseeing or rest/relaxation to a distant exotic destination in a year.

The ATS (1995), unlike other surveys, allows differentiating these segments through introduction of vacation as flag variable for same purposes depending on whether it is made on (or off) vacation. This segmentation is not feasible with other surveys because, vacation itself is introduced as a travel purpose, making it difficult to understand household preferences for travel within such segments for a given purpose. The above variable, also provides more clarification for respondents in addressing their questions, and should be taken into account in future collection efforts.

Further, the other aspect of data collection that is relevant from the perspective of assessing long distance travel, in particular for pleasure is the heterogeneity exhibited across socio-demographics and regions. An exploratory analysis of pleasure travel from ATS (1995) reveals individuals to exhibit distinct travel preferences across income groups, age and their region of residence. The older (55 and over) and younger working adults (35 and under) are found to exhibit higher preference to travel for visiting compared to leisure, with the middle age group exhibiting higher preference to undertake leisure travel. Interestingly, the above behavior was also evident from the examination of expenditure trend data analyzed using CEX data from BLS.

Further, CEX data from BLS, importantly revealed, transportation cost followed by cost at destination, such as that for lodging and food, to be the major factor influencing a households' decision to undertake long distance travel. These aspects are found to influence a households' travel purpose, their choice of destination and travel mode, with potential for them to minimize at least their destination expenditure on pursuing VFR travel. Further, CEX data also revealed

individuals to curb their travel during adverse economic conditions (e.g.: recession), with most of such reduction coming from reduced expenditures on transportation.

A significant share of this reduction came from households residing in the west, identified to make up a major share of such travel using airplanes (McGuckin, 2012). The reduction in such travel is bound to significantly reduce travel expenditures. The understanding gained from these theoretical assessments further informed the development of model to analyze annual household long distance pleasure travel behavior accounting for household socio-demographics, budget and relative transportation and destination expenditures.

5.2 Methodological Contribution

Utilizing the theoretical understanding gained from the assessment of long distance pleasure travel, a novel econometric model was developed to analyze a households' annual choice of destination(s) and mode of travel to such destination(s) subject to time and budget constraints for leisure travel. To this effect, a unified RUM framework was formulated that can be used as a joint MDC-SDC modeling framework to analyze discrete-continuous choices from a combination of perfect and imperfect substitutable choice alternatives. The MDC (Multiple Discrete Continuous) choices in the empirical specification of this research were one or more destination(s) to which a household would allocate their time for leisure. While the SDC (Single Discrete Choice) is a households' travel mode choice to a destination, i.e. air or ground as the travel mode choice; in that a household only chooses one of two available choices in this empirical context.

The key to this formulation is a utility form that is linear with respect to consumption across perfectly substitutable alternatives (travel mode choice in this dissertation) and nonlinear with respect to consumption across imperfectly substitutable alternatives (destination choices in

this dissertation). The linear utility form ensures single discretion among the perfect substitutes (i.e., only a single choice alternative is chosen). On the contrary, the nonlinear form accommodates multiple discretions among imperfect substitutes (i.e., multiple alternatives could be chosen). Furthermore, this formulation accommodates price variation across its choice of destination and mode. In this dissertation, this was introduced as per night cost to spend at destination across each destination-mode choice combination, with this cost calculated as sum of per night lodging and food cost, and travel cost amortized over the nights spent at destination.

Besides the formulation, a procedure to apply the proposed framework for forecasting purposes was also implemented. The overall formulation allows for single as well as multiple linear budget constraints, which in the empirical context of this dissertation are the pleasure time and monetary budgets. To our knowledge, this is the first formulation in the econometric literature to account for multiple linear budget constraints and price variation to model discrete-continuous choices from a combination of perfect and imperfectly substitutable choice alternatives. An empirical model was implemented using the above proposed formulation to estimate a households' annual leisure travel behavior for single as well as multiple budget constraints presented in the next section.

Following the above task, this dissertation further constructed a comprehensive model system to estimate a household's annual pleasure travel demand within the U.S. This was developed accounting for both leisure (non-VFR) as well as visiting (VFR) purpose; along with the likelihood of them making one or more trips via air or ground as principal travel mode for each purpose through additional nested MNL models. However, this model is only developed for a single budget constraint, and does not account for price variation across destinations at this time. This is because the additional nested structure makes it difficult to incorporate price

variations across the multitude of choice combinations (210 destinations * 2 mode * number of trips across modes). Alternatively, travel cost variables in this model were introduced as explanatory variables in the utility functions rather than being treated as part of price variation.

Overall, the national pleasure travel demand model incorporating both leisure (non-VFR) and VFR were estimated in two stages, with the first stage estimating a household's annual pleasure time budget, which served as input to second stage single constraint multiple-discrete continuous model. The model in the second stage with time budget constraints as input from first stage modeled a households' annual choice of destination, purpose of travel (VFR or leisure) to such destination, travel mode (air/ground) to each of these destinations and frequency of such trips. As is intuitive, it would be very difficult to implement a multiple time and budget constraint model with so many choices and the consequent price variations across them, unlike the leisure only model which just had destination and mode choices.

5.3 Empirical Application

Following from above, this dissertation implemented two empirical models to assess annual household pleasure travel demand. The first model (presented in Chapter 3) applied the RUM framework to analyze a households' annual discrete (destination and mode) and continuous (duration at destination) leisure choices accounting for expected costs across these choices subject to multiple (time and money) budget constraints. The expected cost here represented the price variation across destinations, in that it accounted for the cost to travel to each of the potential destination choices based on available mode as well as the cost to spend a night at such destination. The second model (presented in Chapter 4) was conceived with a vision to have a comprehensive national pleasure model system. It is constructed so as to estimate a households' annual pleasure choices of: purpose (motivation) to travel to a

destination, mode choice, and trip frequencies for particular purpose-destination via a given mode pursued within their annual budget for pleasure. This model as stated earlier was only implemented for single time budget constraints, with cost being introduced as variable in the utility specification.

5.3.1 Empirical Application of RUM Framework for Discrete Continuous Choices with Perfect and Imperfect Substitutes

In this first empirical application, the RUM framework is applied to develop a joint model of annual, long-distance vacation destination and mode choices. The model simultaneously analyzed the vacation destinations that a household visits over an entire year, along with the time allocation and the travel mode to each of the visited destinations based on household time and monetary budget constraint, and their expected expenditures at these destinations.

The formulation assumes that, over an annum, households allocate a part of the total time (365 days) and money (annual income) available with them to one or more vacation destinations and make the mode choices to travel to such destination in a manner that it maximizes the total utility derived from their choices. The framework recognizes that the vacation destinations are imperfect substitutes, in that a household can potentially choose to visit multiple destinations over a year. On the contrary, the travel mode alternatives to a destination are perfect substitutes in that only one primary mode is chosen to travel to a destination. The choices are made within a framework that recognizes that households operate under both time and money budget constraints.

The proposed modeling framework is applied to the 1995 ATS data to estimate the empirical model parameters, with the United States divided into 210 alternative long-distance

vacation destination choices for each household in the sample. The ATS data provides information on the different vacation destinations visited (and the time spent on each trip) by the surveyed households over the period of an entire year. The destination attributes, lodging costs and other travel costs on vacation for each household in the sample to each of 210 potential destinations were constructed by synthesizing information from the ATS and variety of other data sources, including the CEX (presented in Appendix A).

The resulting empirical model estimates were reasonable, and shed light on the various factors influencing households' annual vacation destination and mode choices along with related time and money allocations. The results demonstrate the applicability of the proposed framework for modeling discrete-continuous choices in the presence of a mix of imperfect and perfect substitutes in the choice set, while considering multiple budget constraints and price variation across such choices. In addition, the analysis demonstrates the benefit of considering both time and money budget constraints simultaneously in analyzing households' vacation travel choices. Importantly, considering the time and money constraints simultaneously lead to significant model improvements (over the single-constrained models), goodness of fit in the estimation sample as well as the predictive performance (as measured by predictive log-likelihood) on a validation sample. Besides, the time and money-constrained model demonstrated a greater capture of the variation in household preferences than the models that ignored one of the two constraints.

5.3.2 Annual Household Pleasure-Travel-Demand Model System

The second empirical model development provides a comprehensive assessment of household pleasure travel choices made over a year. To this effect, a two-stage model is estimated as presented in the last section, with the first stage estimating a households' annual

pleasure time budget using a stochastic frontier model. The estimated budget from this first stage is further allocated simultaneously to one or more destination. Specifically, the model allocates a households' pleasure time budget to one or more destination(s) to pursue visiting or leisure accessed via air or ground as their principal mode of travel on one or more occasion (trips) in a year. In essence, this research extends our prior work (Van Nostrand et al., 2013) along three fronts, (a) estimating an annual pleasure budget, (b) incorporation of multiple travel purpose and (c) accounting for the possibility of making one or more trips for each travel purpose across ground and air mode as travel choices.

In addition to the above extension, the model also examined the influence of social ties using census migration data. This forms a significant factor in influencing VFR travel, which has been gaining recognition in tourism section since Jackson (1990) and Buchanan (1996), but has received modest interest in transportation. Further, unlike prior studies, this dissertation examine the relative attractiveness of one destination to another based on extent of migration, using migration volumes, rather than just examining its influence of travel to a particular destination. This is achieved through introduction of migration volumes between destinations as utility variable in this model. In essence, this variable represents the relative attractiveness of a destination not just for visiting but for leisure purposes as well.

Besides estimation, the model was also validated using a forecasting procedure developed by Pinjari and Bhat (2011), after extending it to accommodate the joint MDCEV-MNL model structure proposed for this pleasure travel demand model. The forecasts revealed expected shifts across trip length distribution and time allocation across purposes under price rise scenarios. In general, households were found to reduce their travel to distant destinations with three times increase in travel cost, likely with them avoiding travel to distant destinations from this

prediction procedure. This form of behavior was also evident from CEX (BLS) expenditure trend analysis undertaken in Chapter 2. Overall, this behavioral model holds promise to be a useful tool for planning and scenario analysis across transportation and tourism with further improvements discussed in the following section.

5.4 Future Work

On the methodological front, the RUM econometric model could be extended in several important directions. First, development of rigorous constrained maximum likelihood estimation techniques for the proposed formulation would help alleviate some of the estimation difficulties encountered in this study. Second, development of efficient forecasting procedures for the proposed formulation with multiple budget constraints will enable the use of the estimated model for practical forecasting and policy analysis purposes. Third, the empirical application in the dissertation does not treat travel costs as fixed costs (in that travel costs vary with the no. of days spent at the destination). Instead, the travel costs are combined with variable costs such as lodging and dining costs by assuming travel costs could be amortized over the number of vacation days spent at the destinations. However, treating travel costs as fixed and different from variable costs makes the consumer's utility maximization problem nonlinear and unsmooth with respect to the money budget constraint and makes it difficult to use KKT conditions for solving the problem. Thus accommodating nonlinear and unsmooth budget constraints in random utility maximization-based discrete-continuous choice models is an important avenue for future research (see Parizat & Shachar, 2010 for a recent attempt at this).

On the empirical front, the above RUM framework implemented to analyze a joint destination and mode choice for leisure travel could be extended to analyze both VFR and Non-VFR (leisure) travel of households. This would involve extending the present RUM formulation

to incorporate additional nests to accommodate the choices of travel purpose and trip frequencies for each purpose across mode. It would make the model more robust and comprehensive in evaluating the household decisions for VFR and non-VFR travel, which at this time is implemented on a single pleasure budget constraint. However, this would also require arriving at appropriate price estimates for each of the potential combinations accounting for destination, mode, purpose and trip frequencies. Further, the estimation of the above proposed model extension could be enhanced through accommodation of heterogeneity across destination observed to affect the model fit in the single constraint pleasure model system proposed in this dissertation. Besides the above, the model could further be improved through additional policy specific variables such as information on the number of specific attractions that influence household travel decisions such as national monuments, parks, professional sports leagues etc. Overall, this model system could be further extended to incorporate other significant trip purposes such as business and commute to be able to comprehensively analyze and estimate annual household long distance travel demand.

REFERENCES

- Ahn, J., Jeong, G., and Kim, Y. (2008). A forecast of household ownership and use of alternative fuel vehicles: A multiple discrete-continuous choice approach. *Energy Economics*, 30, 2091–2104. doi:10.1016/j.eneco.2007.10.003
- Aigner, D., Knox Lovell, C. A., and Schmidt, P. (1977). Formulation and estimation of stochastic frontier production models. *Journal of Econometrics*, 6, 21–37. doi:10.1016/0304-4076(77)90052-5.
- Amemiya, T. (1974). Multivariate regression and simultaneous equation models when the dependent variables are truncated normal. *Econometrica*, 42, 999–1012. doi:10.2307/1914214
- Augustin, B., A.R. Pinjari, A. Faghih Imani, V. Sivaraman, N. Eluru, and R.M. Pendyala (2014). Stochastic Frontier Estimation of Budgets for Kuhn-Tucker Demand Systems: Application to Activity Time-Use Analysis. Technical Paper, University of South Florida.
- American Travel Survey, 1995.
- Armario, E. M.(2008). Tourist satisfaction: An analysis of its antecedents. Universidad, Sociedad y Mercados Globales. Asociacion Espanola de Direccion y Economia de la Empresa.
- Ashiabor, S., Baik H., and Trani, A. (2007). Logit models for forecasting nationwide intercity travel demand in the United States. *Transportation research record. Journal of the Transportation Research Board*, 1–12
- Augustin, B., Pinjari, A. R., Sivaraman, V., Imani, A. F., Eluru, N., and Pendyala, R. M. (2014). Stochastic Frontier Estimation of Budgets for Kuhn-Tucker Demand Systems: Application to Activity Time-Use Analysis. (*In Review*)
- Backer, E. R. (2008). VFR travellers—Visiting the destination or visiting the hosts? *Asian Journal of Tourism and Hospitality Research*, 2(1), 60–70. Retrieved from <http://www.cecc.com.au/clients/sob/research/docs/ebacker/AJTHR.pdf>
- Backer, E. R. (2009). VFR travel: An assessment of VFR versus non-VFR travellers (Doctoral thesis). Retrieved from <http://epubs.scu.edu.au/theses/122/>
- Backer, E. (2012). VFR travel: It is underestimated. *Tourism Management*, 33 (1), 74-79. Elsevier Ltd. doi:10.1016/j.tourman.2011.01.027

- Baik, H., Trani, A. A., Hinze, N., Swingle, H., Ashiabor, S., and Seshadri A. (2008). Forecasting model for air taxi, commercial airline, and automobile demand in the United States. *Transportation Research Record: Journal of the Transportation Research Board*, 2052, 9–20. doi:10.3141/2052-02
- Banerjee, A., X. Ye, and R.M. Pendyala (2007). Understanding Travel Time Expenditures Around the World: Exploring the Notion of a Travel Time Frontier. *Transportation*, 34(1), 51-65.
- 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. R. (2005). A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions. *Transportation Research Part B: Methodological*, 39, 679–707. doi:10.1016/j.trb.2004.08.003
- Bhat, C. R. (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, 274–303. doi:10.1016/j.trb.2007.06.002
- Bhat, C. R., Castro, M., and Khan, M. (2013). A new estimation approach for the multiple discrete-continuous probit (MDCP) choice model. *Transportation Research Part B: Methodological*, 55, 1–22. doi:10.1016/j.trb.2013.04.005
- Bhat, C. R., Castro, M., and Pinjari, A. R. (2013). Allowing for non-additively separable and flexible utility forms in multiple discrete-continuous models. Retrieved from <http://repositories.lib.utexas.edu/handle/2152/21072>
- Bhat, C. R., Sen, S., and Eluru, N. (2009). The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. *Transportation Research Part B: Methodological*, 43, 1–18. doi:10.1016/j.trb.2008.06.009
- Bhat, C. R., Srinivasan, S., and Sen, S. (2006). A joint model for the perfect and imperfect substitute goods case: Application to activity time use decisions. *Transportation Research Part B: Methodological*, 40, 827–850. doi:10.1016/j.trb.2005.08.004
- Bockstael, N. E., Strand, I. E., Jr., and Hanemann, M. W. (1987). Time and the recreational demand model. *Land Economics*, 69, 293–302.
- Bradley, M., Outwater, M. L., Ferdous, N., Bhat, C., Pendyala, R., Hess, S., Daly, A., and LaMondia, J. (2014). A tour-based national model system to forecast long distance passenger travel in the United States. (Working Paper).
- Braunlich, C. G., and Nadkarni, N. (1995). The importance of the VFR market to the hotel industry. *Journal of Tourism Studies*, 6(1), 38–47.
- Brons M., Pels, E., Nijkamp, P., and Rietveld, P. (2001). Price elasticities of demand for passenger air travel (Tinbergen Institute Discussion Paper, No 01-047/3).

- Buchanan, I. (1996, February). Examination of the relationship between VFR travel and population of Australian tourism regions. Paper presented at the Australian Tourism and Hospitality Research Conference, Southern Cross University, New South Wales, Australia.
- Bureau of Economic Analysis. (1995). Gross domestic product by state. Retrieved from <http://www.bea.gov/regional/gsp/>
- Bureau of Labor Statistics. (1995). Employment, hours, and earnings—State and metro area. Retrieved <http://www.bls.gov/data/#employment>
- Bureau of Labor Statistics, 2010.
- Bureau of Transportation Statistics. (1995a). 1995 American travel survey technical documentation. Washington, DC: Author.
- Bureau of Transportation Statistics. (1995b). Airline origin and destination survey (DB1B). Washington, DC: Author.
- Bureau of Transportation Statistics. (1997). 1995 American travel survey profile. Retrieved 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). NHTS – Long distance travel quick facts. Retrieved from http://www.bts.gov/programs/national_household_travel_survey/long_distance.html
- Bureau of Transportation Statistics. (2002).
- Bureau of Transportation Statistics (2003). Highlights of the 2001 National Household Travel Survey. Washington, DC.
- Bureau of Transportation Statistics. (2006).
- Castro, M., Bhat, C. R., Pendyala, R. M., and Jara-Diaz, S. (2012). Accommodating multiple constraints in the multiple discrete–continuous extreme value (MDCEV) choice model. *Transportation Research Part B: Methodological*, 46, 729–743. doi:10.1016/j.trb.2012.02.005
- Chiang, J. (1991). A simultaneous approach to the whether, what, and how much to buy questions. *Marketing Science*, 10, 297–314. doi:10.1287/mksc.10.4.297
- Chintagunta, P. K. (1993). Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Science*, 12, 184–208. doi:10.1287/mksc.12.2.184
- Chu, X., and Polzin, S. E. (2002). An application for measuring vehicle travel by visitors. *Journal of Transportation and Statistics*, 5(2/3), 83–90. Retrieved from <http://www.cutr.usf.edu/pdf/Chu.pdf>
- Contrino, H., McGuckin, N., and Banks, D. (2000). Exploring the full continuum of travel: Data Fusion by recursive partitioning regression. Presented at the 9th International Conference on Travel Behaviour Research, Gold Coast, Queensland, Australia.

- Daly, A. (1982). Estimating choice models containing attraction variables. *Transportation Research Part B: Methodological*, 16, 5–15. doi:10.1016/0191-2615(82)90037-6
- Dann, G. M. S (1977). Anomie, ego-enhancement and tourism. *Annals of Tourism Research*, 4(4),184–194.
- Deaton, A., and Muellbauer, J., (1980). *Economics and consumer behavior*. Cambridge, England: Cambridge University Press.
- Decrop, A., and Snelders D. (2004). Planning the summer vacation: An adaptable process. *Annals of Tourism Research*. 31(4), 1008–1030.
- Dubin, J. A., and McFadden, D. L. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, 52, 345–362. doi:10.2307/1911493
- Edwards, Y. D., and Allenby, G. M. (2003). Multivariate analysis of multiple response data. *Journal of Marketing Research*, 40, 321–334. doi:10.1509/jmkr.40.3.321.19233
- Eluru, N., Pinjari, A. R., Pendyala, R. M., and Bhat, C. R. (2010). An econometric multi-dimensional choice model of activity-travel behavior. *Transportation Letters*, 2, 217–230. doi:10.3328/TL.2010.02.04.217-230
- Energy Information Administration. (1995). Retail gasoline historical prices. Retrieved from http://www.eia.doe.gov/oil_gas/petroleum/data_publications/wrgp/mogas_history.html
- Epstein, J. M., Parker, J., Cummings, D., and 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, 341–354. doi:10.1177/146735840300400407
- Fang, H. A. (2008). A discrete-continuous model of households' vehicle choice and usage, with an application to the effects of residential density. *Transportation Research Part B: Methodological*, 42, 736–758. doi:10.1016/j.trb.2008.01.004
- Fosgerau, M. (2001). PETRA—an activity based approach to travel demand analysis. In L. Lundqvist and L. G. Mattsson (Eds.), *National transport models: Recent developments and prospects* (pp. 134–145). Stockholm, Sweden: The Swedish Transport and Communications Research Board.
- Griffin, K. (2001). Recent trends in global and Irish tourism. *Geographic Viewpoints—Journal of Association of Geography Teachers of Ireland*, 29, 22–30.
- Giesbrecht, L., and Bose, J. (2005). Developing a sustainable and flexible long-distance travel survey. *Proceedings from the Survey Research Methods Section, American Statistical Association*.
- Habib, K. M. N., and Miller, E. J. (2008). Modelling daily activity program generation considering within-day and day-to-day dynamics in activity-travel behaviour. *Transportation*, 35, 467–484. doi:10.1007/s11116-008-9166-8

- 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.
- Hanemann, W. M. (1978). A methodological and empirical study of the recreation benefits from water quality improvement (Doctoral dissertation). Department of Economics, Harvard University.
- Hanemann, W. M. (1984). Discrete/continuous models of consumer demand. *Econometrica*, 52, 541–562. doi:10.2307/1913464
- 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.
- Heckman, J. J. (1974). Shadow prices, market wages, and labor supply. *Econometrica*, 42, 679–694. doi:10.2307/1913937
- Heckman, J. J. (1979) Sample selection bias as a specification error. *Econometrica*, 47, 153–161. doi:10.2307/1912352
- Herriges, J. A., Kling, C. L., and Phaneuf, D. J. (2004). What’s the use? Welfare estimates from revealed preference models when weak complementarity does not hold. *Journal of Environmental Economics and Management*, 47, 55–70. doi:10.1016/S0095-0696(03)00058-5
- Heung, C. S. V., Hailin, Q., and Chu, R. (2001). The relationship between vacation factors and socio-demographic and travelling characteristics: the case of Japanese leisure travellers. *Tourism Management*, 22, 259–269.
- Horowitz, A. J. (2008). Statewide travel demand forecasting. Paper presented at the Meeting federal surface transportation requirements in statewide and metropolitan transportation planning conference. Retrieved from [http://onlinepubs.trb.org/onlinepubs/archive/NotesDocs/NCHRP08-36\(76C\)_ModelingFR.pdf](http://onlinepubs.trb.org/onlinepubs/archive/NotesDocs/NCHRP08-36(76C)_ModelingFR.pdf)
- Hu, B. A., and Morrison, A. M. (2002). Tripography: Can destination use patterns enhance understanding of the VFR market? *Journal of Vacation Marketing*, 8, 201–220. doi:10.1177/135676670200800301
- Hu, B. A., Morrison, A. M., and O’Leary, J. T. (2002). An analysis of the variety seeking behaviour of the U.S. visiting friends and relatives market. *Journal of Tourism Studies*, 13(2), 28–40.
- Hwang, Y. H., Gretzel, U., and Fesenmaier, D. R. (2006). Multicity trip patterns: Tourists to the United States. *Annals of Tourism Research*. 33(4), 1057–1078.
- Inter University Consortium for Political and Social Research. (2011). Bureau of Labor Statistics Consumer Expenditure Survey, 1995. Retrieved from <http://www.icpsr.umich.edu/icpsrweb/ICPSR/index.jsp>
- Jackson, R. T. (1990). VFR tourism: Is it underestimated? *Journal of Tourism Studies*, 1(2), 10–17.

- Jaggi, B., Weis, C., and Axhausen, K. W. (2013). Stated response and multiple discrete-continuous choice models: Analysis of residuals. *Journal of Choice Modelling*, 6, 44–59. doi:10.1016/j.jocm.2013.04.005
- Kim, J., Allenby, G. M., and Rossi, P. E. (2002). Modeling consumer demand for variety. *Marketing Science*, 21, 229–250. doi:10.1287/mksc.21.3.229.143
- King, B. (1994). What is ethnic tourism? An Australian perspective. *Tourism Management*, 15(3), 173–176. doi:10.1016/0261-5177(94)90101-5
- Kitamura, R., T. Yamamoto, K. Kishizawa, and R.M. Pendyala (2000). Stochastic Frontier Models of Prism Vertices. *Transportation Research Record*, No. 1718, 18-26.
- Koppelman, F., and 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, 825–853. doi:10.1016/j.trb.2004.10.003
- Kozak, M. (2002). Comparative analysis of tourist motivations by nationality and destinations. *Tourism Management*, 23(3), 221–232.
- Kuriyama, K., Shoji, Y., and Tsuge, T. (2011, November). Recreation demand models with spatial heterogeneity: A spatial Kuhn-Tucker model. Paper presented at European Association of Environmental and Resource Economists, 18th Annual Conference, Rome, Italy.
- LaMondia, J., Bhat, C., and Hensher, D. (2008). An annual time use model for domestic vacation travel. *Journal of Choice Modelling*, 1(1), 70–97. doi:10.1016/S1755-5345(13)70023-X
- LaMondia, J., Snell, T., and Bhat, C.R. (2009). Travel behaviour and values analysis in the context of vacation destination and travel mode choices: European Union case study. *Transportation Research Record: Journal of the Transportation Research Board.*, 2156, 140–149. doi:10.3141/2156-16.
- Lee, L.-F. (1983). Generalized econometric models with selectivity. *Econometrica*, 51, 507–512. doi:10.2307/1912003
- Lee, L.-F., and Pitt, M. M. (1986). Microeconomic demand system with binding nonnegativity constraints: The dual approach. *Econometrica*, 54, 1237–1242. doi:10.2307/1912330
- Lee, S., Kim, J., and Allenby, G. M. (2010). A category-level model of asymmetric complements. *Social Science Research Network*, working paper. doi:10.2139/ssrn.1596218
- Leiper, N. (2004). *Tourism management* (3rd ed.). Santa Cruz, CA: Arnold.
- Lundquist, L., and Matteson, L.G. (2002). *National Transport Models: Recent Developments and Prospects*: Springer.
- Maddala, G. S. (1983). *Limited-dependent and qualitative variables in economics*. Cambridge, England: Cambridge University Press.
- Mattsson, L. G. (2002). *National transport models: recent developments and prospects*: Springer.

- Mäler, K-G. (1974). *Environmental economics: A theoretical inquiry*. Baltimore, MD: Johns Hopkins University Press.
- Mallet, W.J. (2001). Long distance travel by low-income households. *Transportation Research Circular E-C026- Personal Travel: The long and short of it*, 169-177.
- Manchanda, P., Ansari, A., and Gupta, S. (1999). The “shopping basket”: A model for multicategory purchase incidence decisions. *Marketing Science*, 18, 95–114. doi:10.1287/mksc.18.2.95
- McGuckin, N. (2009). Long-Distance travel in the United States. Presented at the TRB Annual Conference, Washington, D.C.
- McGuckin, N. (2013). Long distance travel data: Challenges and opportunities. Presented at the 92nd Annual Meeting of the Transportation Research Board, Washington, D.C.
- Morrison, A.M., Hsieh, S. and O’Leary, J.T. (1995). Segmenting the visiting friends and relatives market by holiday activity participation. *Journal of Tourism Studies*, 6(1), 44-63.
- Meis, S., Joyal, S., and Trites, A. (1995). The U.S. repeat and VFR visitor to Canada: Come again Eh!. *Journal of Tourism Studies*. 6(1), 27–37.
- Microsoft: MapPoint North America. (2009).
- Miller, E. J. (2003). The trouble with intercity travel demand models. Presented at the 83rd Annual Meeting of the Transportation Research Board, Washington, D.C.
- Moeckel, R., and Donnelly, R. (2010). Nationwide estimate for long distance travel. Presented at the Third International Conference on Innovations in Travel Modeling (ITM) of the Transportation Research Board, Tempe, AZ.
- Morley, C. L. (1992). A microeconomic theory of international tourism demand. *Annals of Tourism Research*, 19, 250–267. doi:10.1016/0160-7383(92)90080-9
- Morrison, A. M., Hsieh, S., and O’Leary, J. T. (1995). Segmenting the visiting friends and relatives market by holiday activity participation. *Journal of Tourism Studies*, 6(1), 44–63. Retrieved from <http://www.jcu.edu.au/>
- Moscardo, G., Pearce, P., Morrison, A., Green, D., and O’Leary J. T. (2000). Developing a typology for understanding visiting friends and relatives markets. *Journal of Travel Research*, 38, 251–259. doi:10.1177/004728750003800307
- National Cooperative Highway Research Program, 2004.
- National Cooperative Highway Research Program, 2012.
- National Household Travel Survey, 2001. Retrieved from <http://nhts.ornl.gov/publications.shtml>
- National Household Travel Survey, 2009. Retrieved from <http://nhts.ornl.gov/publications.shtml>
- National Oceanic and Atmospheric Administration. (2011). Ocean and coastal resource management. Retrieved from <http://coastalmanagement.noaa.gov/>
- Nationwide Personal Transportation Survey, 1977. Retrieved from <http://nhts.ornl.gov/download.shtml>

- Nationwide Personal Transportation Survey, 1983. Retrieved from <http://nhts.ornl.gov/download.shtml>
- Nationwide Personal Transportation Survey, 1987. Retrieved from <http://nhts.ornl.gov/download.shtml>
- Nationwide Personal Transportation Survey, 1990. Retrieved from <http://nhts.ornl.gov/download.shtml>
- Nationwide Personal Transportation Survey, 1995. Retrieved from <http://nhts.ornl.gov/download.shtml>
- Nationwide Personal Transportation Survey, 1998. Retrieved from <http://nhts.ornl.gov/download.shtml>
- National Personal Transportation Survey Data book, 1998. Retrieved from <http://nhts.ornl.gov/publications.shtml>
- Nicolau, J., and Mas, F. (2005) Stochastic modelling: 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., and Modulgula, V. (2010). California Statewide Model for High-Speed Rail. *Journal of Choice Modelling*, 3(1), 58.
- Parizat, S., and Shachar, R. (2010). When Pavarotti meets Harry Potter at the Super Bowl? (Working paper). Tel Aviv, Israel: Tel Aviv University.
- Pennington-Gray, L. (2003). Understanding the domestic VFR drive market in Florida. *Journal of Vacation Marketing*, 9, 354–367. doi:10.1177/135676670300900405
- Phaneuf, D. J., Kling, C. L., and Herriges, J. A. (2000). Estimation and welfare calculations in a generalized corner solution model with an application to recreation demand. *The Review of Economics and Statistics*, 82, 83–92. doi:10.1162/003465300558650
- Phaneuf, D. J., and Smith, V. K. (2005). Recreation demand models. In K.-G. Mäler and J. R. Vincent (Eds.), *Handbook of environmental economics* (Vol. 2, pp. 671–751). Amsterdam, Holland: Elsevier.
- Pinjari, A. R. (2011). Generalized extreme value (GEV)-based error structures for multiple discrete-continuous choice models. *Transportation Research Part B Methodological*, 45, 474–489. doi:10.1016/j.trb.2010.09.004
- Pinjari, A. R., and Bhat, C. R. (2011). Computationally efficient forecasting procedures for Kuhn-Tucker consumer demand model systems: Application to residential energy consumption analysis. Retrieved from <http://www.cae.utexas.edu/>
- Pinjari, A. R., and Sivaraman, V. (2013). A discrete-continuous choice model with perfect and imperfect substitutes. Retrieved from http://abdulpinjari.weebly.com/uploads/9/6/7/8/9678119/pinjari_sivaraman_new_perfect-imperfect_formulation_july-15-2013.pdf
- Poel, P., Masurel E., and Nijkamp, P. (2004). The importance of friends and relations in tourist behaviour: A case study on heterogeneity in Surinam. Amsterdam, The Netherlands: VU University.
- Quarmby, D. (2006). Tourism for transport professionals or ‘what you hadn’t realised you didn’t know about tourism.’ Presented at the European Transport Conference.

- Rich, J., Brocker, J., Hanson, C., Korchenewych, A., Nielsen, O., and Vuk, G. (2009). Traffic flow: Scenario, traffic forecast and analysis of traffic on the TEN-T, taking into consideration the external dimension of the Union. Retrieved from http://energy.jrc.ec.europa.eu/transtools/docs/TenC408_001_ver20090721.pdf
- Satomura, T., Kim J., and Allenby, G. M. (2011). Multiple-constraint choice models with corner and interior Solutions. *Marketing Science*, 30, 481–490. doi:10.1287/mksc.1110.0636
- Savageau, D., and Loftus, G. (1997). *Places rated almanac: Your guide to finding the best places to live in North America* (5th ed.). New York, NY: Macmillan Travel.
- Sharp, J. (2011). Long distance travel surveys: historical overview and challenges for future surveys. NHTS Workshop. Retrieved from <http://onlinepubs.trb.org/onlinepubs/conferences/2011/NHTS1/Sharp.pdf>
- Srinivasan, S., and Bhat, C. R. (2006). A multiple discrete-continuous model for independent and joint-discretionary-activity participation decision. *Transportation*, 33, 497–515. doi:10.1007/s11116-006-8078-8
- U.S. Census Bureau, 1995.
- U.S. Census Bureau. (2000). Demographic profiles: 100-percent and sample data. Retrieved from <http://www.census.gov/census2000/demoprofiles.html>
- U.S. Travel Association, (2014): U.S. Travel Answer Sheet, June 2014. Accessed August 1, 2014 from <http://www.ustravel.org/news/press-kit/travel-facts-and-statistics>
- Uysal, M., and Hagan, L. A. R. (1993). Motivation of pleasure travel and tourism. *Encyclopedia of Hospitality And Tourism*. 798,810.
- Van Middelkoop, M., Borgers, A., and Timmermans, H. J. P. (2004). Merlin: Microsimulation system for predicting leisure activity travel patterns. *Transportation Research Record*, 1894, 20–27. doi:10.3141/1894-03
- Van Nostrand, C. (2011). A discrete-continuous modeling framework for long-distance, leisure travel demand analysis (Master's thesis). University of South Florida, Tampa. Retrieved from <http://scholarcommons.usf.edu/etd/3389/>
- Van Nostrand, C., Sivaraman, V., and Pinjari, A. R. (2013). Analysis of long-distance vacation travel demand in the United States: A multiple discrete-continuous choice framework. *Transportation*, 40, 151–171. doi:10.1007/s11116-012-9397-6
- Vasquez-Lavin, F., and Hanemann, M. (2009, March). Functional forms in discrete/continuous choice models with general corner solutions. *Proceedings of the 1st International Choice Modeling Conference*, Harrogate, England.
- von Haefen, 2007.
- von Haefen, R. H., and Phaneuf, D. J. (2003). Estimating preferences for outdoor recreation: A comparison of continuous and count data demand system frameworks. *Journal of Environmental Economics and Management*, 45, 612–630. doi:10.1016/S0095-0696(02)00024-4

- von Haefen, R. H., Phaneuf, D. J., and Parsons, G. R. (2004). Estimation and welfare analysis with large demand systems. *Journal of Business and Economic Statistics*, 22, 194–205. doi:10.1198/073500104000000082
- Wales, T. J., and Woodland, A. D. (1983). Estimation of consumer demand systems with binding non-negativity constraints. *Journal of Econometrics*, 21, 263–285. doi:10.1016/0304-4076(83)90046-5
- Yoon, Y. and Uysal, M. (2005). An examination of the effects of motivation and satisfaction on destination loyalty : a structural model. *Tourism Management*, 26(1):45-56.
- Zhang, L., F. Southworth, C. Xiong and A. Sonnenberg (2012) Methodological Options and Data Sources for the Development of Long Distance Passenger Travel Demand Models: A Comprehensive Review. *Transport Reviews*. 32(4):399-433.

APPENDICES

Appendix A: Estimation of Per Night Cost at Destination

There are two types of unit prices for each vacation destination and travel mode alternative – time-prices and money-prices (i.e., the q_{jl} and p_{jl} variables in the time and money constraints of Equation 33). For the current analysis, the time-prices (q_{jl}) are considered to be unity in that the amount of time needed to spend 1 day of vacation time is equal to 1 day. This makes an implicit assumption that the time spent traveling to a destination j is part of the vacation time t_j . That is, households derive utility not only from the time spent at a vacation destination, but also from the time spent traveling to the destination. This is reasonable because traveling for vacation might not be as onerous (it might in fact be fun) as compared to commuting.

However, doing so does not account for the possibility that households tend to prefer to visit closer destinations as opposed to farther destinations. To account for such preferences, the baseline utility functions incorporate the travel time to the destination (by the corresponding mode) as an explanatory variable. One would expect a negative coefficient on this variable. The synthesis of money-prices (p_{jl}), on the other hand, required several assumptions and significant data gathering and processing, as described below. The money-price p_{jl} is the monetary expenditure a household needs to incur to spend unit time (i.e., a day) at a vacation destination j traveled by mode l (note that the subscript for the household is suppressed for simplicity in notation). These prices comprise two components – (a) destination prices that do not depend on the mode of travel and (b) travel prices that depend on the mode of travel. The destination prices, in turn, have two components: (a1) lodging prices (i.e., lodging costs per day) and (a2) nonlodging prices (i.e., costs per day for dining, recreation, entertainment, etc.).

The process used to synthesize the information on money-prices for each household to travel to each available destination by each available travel mode is described below. First, the lodging costs and nonlodging costs per day at each destination were synthesized from the 1995 CEX data using a two stage process. In the first stage, the per-day lodging costs for each household was derived using a regression equation relating the per-day costs to the household's sociodemographic characteristics (income, household size, and residential Census region). This regression equation was estimated using household-level microdata on annual vacation expenditures from the 1995 CEX data (see Table a1). Similarly, the per-day nonlodging costs were derived using another regression equation estimated with the CEX data on nonlodging vacation expenditures (see Table a2).

Both the above mentioned regression equations recognize the variation in per-day costs by household characteristics. Thus, this approach recognizes that not every household incurs the same costs at a destination. Rather, households make the lodging choices and other expenditure choices according to their income and other characteristics. However, the regression equations do not recognize the variation in the lodging and nonlodging prices across the different destinations (because the CEX data does not provide information on which destinations were visited by the households). To accommodate the price-variation across destinations, in the second stage, the regressed per-day costs for each household were scaled by a factor capturing how pricy (or less expensive) each destination is compared to an average destination (as measured by the median per-day costs at different destinations).

To implement this second state strategy, the median values of lodging costs of vacationing at each of the 210 destinations were obtained from a hotel guide database made available by VisitUSA.com (<http://www.visitusa.com/state-hotels/index.htm>). The lodging prices

and nonlodging prices obtained in the above manner were added up to obtain the destination prices. Call such destination price as p_j , where j is the index for destination.

Second, using the 1995 ATS data, the number of days spent at a destination were regressed, using an ordered logit model, as a function of the household characteristics (age of householder, household size, income, presence of children), distance between origin and destination, and an indicator if the destination is an MSA (see Table A3). The resulting ordered response model estimates were used, for each household in the estimation sample, to estimate the expected number of days (n_j) that the household would spend at each destination ($j = 1, 2, \dots, J$) if the household visited that destination. Third, the money-price p_{jl} of spending a day visiting a destination j by mode l was computed as: $p_{jl} = \frac{p_j n_j}{(n_j + tt_{jl})} + \frac{tc_{jl}}{(n_j + tt_{jl})}$, where tc_{jl} and tt_{jl} are the round trip travel cost and travel time, respectively, to travel to a destination j (from the household's origin) by travel mode l . The first component of this money-price formula can be viewed as the destination price, while the second component can be viewed as the travel price. Note from the formula for p_{jl} that the money-prices are computed assuming that the travel costs (tc_{jl}) can be amortized over the duration spent visiting a destination ($n_j + tt_{jl}$).

In reality, however, travel costs are fixed costs and do not depend on the number of days spent at the destination. The current formulation cannot consider such fixed costs, a reason why we assumed that travel costs could be amortized over the no. of days spent at the destination. Enhancing the model formulation to relax this assumption and consider travel costs as *fixed* is an important avenue for further research.

Table A1 Regression estimates for lodging costs per day on vacation (CEX data)

Explanatory variables	Coefficient	t-stat
Constant	58.922	2.85
Region of Residence : Midwest (Northeast as base)	-7.473	2.77
Region of Residence : West (Northeast as base)	-5.787	2.78
HH Income < \$30K (high income is base category)	-15.743	2.65
HH Income between \$30K and < 75K (high income is base category)	-13.134	2.53
Household size < 3 (3 or more member household as base)	-8.473	1.91

Note. The dependent variable, lodging costs per day was derived as the total annual lodging expenditure on vacation divided by the total number of days per annum on vacation.

Table A2 Regression estimates for nonlodging costs per day on vacation (CEX data)

Explanatory variables	Coefficient	t-stat
Constant	74.464	4.61
Region of Residence: Midwest (Northeast is base category)	-10.417	4.12
Region of Residence: West (Northeast is base category)	-9.432	4.14
HH Income < \$30K (high income is base category)	-19.984	3.93
HH Income between \$30K and < 75K (high income is base category)	-17.121	3.76
Household size between 3 and 4 (1–2 member household is base)	11.813	3.07
Household size greater than 5 (1–2 member household is base)	20.603	4.49

Note. The dependent variable, non-lodging costs per day was derived as the total annual expenditure on food, drinks, recreation, entertainment and other local expenditures (other than lodging) during vacation divided by the total number of days per annum on vacation.

Table A3 Ordered logit model for no. of days spent at destination (ATS data)

Explanatory variables	Coefficient	t-stat
Age of householder	0.009	3.66
Household size*Low income household	-0.117	-4.22
Presence of Kids	0.187	2.74
Distance to the destination	0.044	16.02
Destination is an MSA	-0.494	-7.72

Note: Estimates of the thresholds in the ordered logit model are not reported as they do not carry significant interpretation.

Appendix B: National Pleasure Travel Demand Model Estimates

Table B1 Visiting model estimates with and without migration volumes

MDCEV (Destination and Time Allocation) - MNL (Purpose) Choice Model				
Baseline Utility Function (\mathbb{E}_j) Specification	With migration		Without migration	
	Est.	T-Stat	Est.	T-Stat
<i>Variables common to all destination-purpose combinations</i>				
Alternative specific constant	-6.488	-40.16	-6.383	-39.92
Origin Leisure Employment Per Capita	-0.729	-4.04	-0.651	-3.57
Destination Specific Characteristics				
Log of Land Area	0.269	44.07	0.270	44.22
Per Night Cost at Destination (\$100's)	-0.119	-3.63	-0.127	-3.89
Purpose Specific Characteristics				
Alternate specific constant for visiting (Visiting)	-0.176	-1.03	-0.137	-0.81
Log sum from mode and trip frequency model	0.487	17.56	0.499	18.01
Dummy if destination in same state as HH residence	1.450	17.75	1.402	17.63
Dummy if destination in adjacent state to HH residence	1.119	20.19	1.076	19.53
Destination Leisure Employment Per Capita	1.953	9.83	1.987	9.94
Winter (January) temperature. 65°–75° Fahrenheit				
55°–65° Fahrenheit	-0.528	-6.08	-0.554	-7.29
45°–55° Fahrenheit	-0.640	-7.48	-1.007	-11.31
35°–45° Fahrenheit	-0.749	-8.02	-1.225	-12.16
< 35° Fahrenheit	-0.764	-7.85	-1.187	-11.25

Table B1 (Continued)

Summer (June) temperature. 65°-75° Fahrenheit				
60°–65° Fahrenheit	-1.492	-4.11	0.011	0.045
75°–80° Fahrenheit	-0.546	-7.33	-0.586	-7.67
80°–85° Fahrenheit	-0.448	-5.91	-0.393	-4.98
85°–90° Fahrenheit	-0.759	-9.42	-0.752	-9.10
> 90° Fahrenheit	-0.640	-7.38	-0.547	-5.84
<hr/>				
Origin to Destination Migration Volume (1000's)	0.006	2.66		
O–D Migration Volume (1000's)*Non Contiguous State	0.022	6.09		
Low Income (Visiting)	0.202	3.69		
Nonfamily Household (Visiting)	0.093	1.84		
Householder is Employed Full Time (Visiting)	-0.284	-5.37		
Round Trip Distance to Destination Greater than 3,000 miles	-1.08	-13.51	-1.11	-13.77
<i>Scale parameter (t-stats are against a value of 1)</i>		1		1
<i>Dissimilarity Parameter (t-stats are against a value of 1)</i>	0.8714	7.05	0.8893	7.14
Satiation Function (\mathcal{Y}_j) Specification				
Highway distance to destination (100's miles)	0.123	45.74	0.123	46.00
Log-Likelihood at Convergence		-38,945		-39,023
Rho Squared		0.1498		0.1481

Table B2 Leisure model estimates with and without migration volumes

MDCEV (Destination and Time Allocation) - MNL (Purpose) Choice Model				
Baseline Utility Function (\mathbb{E}_j) Specification	With migration		Without migration	
	Est.	T-Stat	Est.	T-Stat
<i>Variables common to all destination-purpose combinations</i>				
Alternative specific constant	-6.48	-40.16		
Origin Leisure Employment Per Capita	-0.729	-4.04		
Destination Specific Characteristics				
Log of Land Area	0.269	44.07		
Per Night Cost at Destination (\$100's)	-0.119	-3.64		
Purpose Specific Characteristics				
Alternate specific constant for visiting (Visiting)				
Log sum from mode and trip frequency model	0.585	21.10	0.592	21.54
Dummy if destination in same state as HH residence	1.705	24.16	1.677	24.84
Dummy if destination in adjacent state to HH residence	1.251	23.46	1.217	23.33
Destination Leisure Employment Per Capita	2.651	21.99	2.666	22.43
Winter (January) temperature. 65°–75° Fahrenheit				
55°–65° Fahrenheit	-0.512	-6.71	-0.594	-6.84
45°–55° Fahrenheit	-0.948	-10.57	-0.713	-8.37
35°–45° Fahrenheit	-1.163	-11.35	-0.829	-8.95
< 35° Fahrenheit	-1.115	-10.37	-0.839	-8.77
Summer (June) temperature. 65°–75° Fahrenheit				
60°–65° Fahrenheit	0.020	0.08	-1.569	-4.27
75°–80° Fahrenheit	-0.593	-7.81	-0.553	-7.40
80°–85° Fahrenheit	-0.395	-5.00	-0.436	-5.80
85°–90° Fahrenheit	-0.763	-9.01	-0.745	-9.34
> 90o Fahrenheit	-0.546	-5.78	-0.657	-7.68
Origin to Destination Migration Volume (1000's)	0.005	2.56		

Table B2 (Continued)

Origin to Destination Migration Volume (1000's)	0.005	2.56		
O-D Migration Volume (1000's)*Non Contiguous State	0.020	5.40		
Low Income (Visiting)				
Nonfamily Household (Visiting)				
<hr/>				
Householder is Employed Full Time (Visiting)				
Round Trip Distance to Destination Greater than 3,000 miles	-1.451	-16.69	-1.486	-17.04
<i>Scale parameter</i> (t-stats are against a value of 1)		1		1
<i>Dissimilarity Parameter</i> (t-stats are against a value of 1)	0.871	7.05	0.889	7.14
Satiation Function (γ_j) Specification				
Highway distance to destination (100's miles)	0.123	45.74	0.123	46.00
Log-Likelihood at Convergence		-38,945.43		-39,022.72
Rho Squared		0.1498		0.1481
<hr/>				