

# **Incorporating the Influence of Latent Modal Preferences in Travel Demand Models**

By  
Akshay Vij

A dissertation submitted in partial satisfaction of the  
requirements for the degree of  
Doctor of Philosophy  
in  
Engineering - Civil and Environmental Engineering  
in the  
Graduate Division  
of the  
University of California, Berkeley

**Committee in charge:**  
Professor Joan Walker, Chair  
Professor Paul Waddell  
Professor John Canny  
Professor Adib Kanafani

Fall, 2013



## **Abstract**

### **Incorporating the Influence of Latent Model Preferences in Travel Demand Models**

By

Akshay Vij

Doctor of Philosophy in Engineering – Civil and Environmental Engineering  
University of California, Berkeley

Professor Joan Walker, Chair

Latent modal preferences, or modality styles, are defined as behavioral predispositions towards a certain travel mode or set of travel modes that an individual habitually uses. They are reflective of higher-level orientations, or lifestyles, that are hypothesized to influence all dimensions of an individual's travel and activity behavior. For example, in the context of travel mode choice different modality styles may be characterized by the set of travel modes that an individual might consider when deciding how to travel, her sensitivity, or lack thereof, to different level-of-service attributes of the transportation (and land use) system when making that decision, and the socioeconomic characteristics that predispose her one way or another. Travel demand models currently in practice assume that individuals are aware of the full range of alternatives at their disposal, and that a conscious choice is made based on a tradeoff between perceived costs and benefits associated with alternative attributes. Heterogeneity in the choice process is typically represented as systematic taste variation or random taste variation to incorporate both observable and unobservable differences in sensitivity to alternative attributes. Though such a representation is convenient from the standpoint of model estimation, it overlooks the effects of inertia, incomplete information and indifference that are reflective of more profound individual variations in lifestyles built around the use of different travel modes and their concurrent influence on all dimensions of individual and household travel and activity behavior.

The objectives of this dissertation are three-fold: (1) to develop a travel demand model framework that captures the influence of modality styles on multiple dimensions of individual and household travel and activity behavior; (2) to test that the framework is both methodologically flexible and empirically robust; and (3) to demonstrate the value of the framework to transportation policy and practice.

In developing an appropriate framework, the dissertation builds on Latent Class Choice Models (LCCMs) used previously in the literature, synthesizing recent advances in the sub-domains of taste heterogeneity and choice set generation, and contributing methodologically to the sub-domains of preference endogeneity and simultaneous choice models. With regards to preference endogeneity, discrete choice models in the past have usually subscribed to the neoclassical

assumption that preferences, as denoted by taste parameters and choice sets, are characteristics of the decision-maker that are exogenous to the choice situation and stable over time. Such a representation would be adequate if an individual's modality style were expected to be invariant across time. However, modality styles are subject to external influences, such as changes to the transportation system. For example, one would expect the introduction of transportation policies such as the London Congestion Charge or the implementation of infrastructural initiatives such as Bogota's Transmilenio bus rapid transit system to lead to changes in modality styles, and an apposite framework should be able to model and predict these changes. With regards to simultaneous choice models, discrete choice models in the past have introduced correlation across multiple dimensions through the covariance structure of the utility specification corresponding to each of the dimensions. Though such an approach has been shown to result in a significant improvement in fit, the covariance structure is a black box that it does not offer any insight to the underlying source of correlation. We introduce correlation through the modality styles construct, conditioning multiple dimensions of individual and household travel and activity behavior on a single overarching modality style, and thereby offering a behavioral rationale to the underlying source of correlation.

The proposed framework has the following structure: modality styles are specified as latent classes. Heterogeneity across modality styles is captured by allowing taste parameters and choice sets corresponding to the class-specific choice models to vary across classes. Preferences are endogenized by defining class membership as a function not only of the characteristics of the decision-maker, as is standard practice, but also of the consumer surplus offered by each class, which in turn is a function of alternative attributes, taste parameters and choice sets. Choices across multiple dimensions are correlated by conditioning the class-specific choice models for all dimensions of interest on the class membership model.

We apply the framework to study the relationship between individual modality styles and travel mode choice behavior using two very distinct travel diary datasets from two very culturally and geographically distinct regions. The first dataset was collected in Karlsruhe, Germany and comprises a relatively small sample of 119 individuals surveyed over a fairly long observation period of six weeks. Estimation results indicate the presence of habitual drivers who display a strong bias for using the automobile and multimodal individuals who exhibit variation in their modal preferences. Multimodal behavior is further distinguished by those who appear to be sensitive to travel times and those who appear to be insensitive. The second dataset was collected in the San Francisco Bay Area in the United States and consists of a relatively large sample of 26,350 individuals surveyed over a fairly short observation period of two days. Estimation results uncover six modality styles that are distinguishable from one another by the kinds of individuals that belong to each of the six modality styles, their latent preferences for different travel modes and the relative importance that they attach to different level-of-service attributes of each of the travel modes. For example, two of the six modality styles comprising 30% of the sample population only consider the car when deciding how to travel. These two modality styles, labeled inveterate drivers and car commuters, can further be distinguished from one another by their value of travel time. Inveterate drivers have a very low value of in-vehicle travel time of 0.55 \$/hr for mandatory tours and are insensitive to in-vehicle travel times for non-mandatory tours. Car commuters have a value of in-vehicle travel time of 6.95 \$/hr for mandatory tours and are insensitive to travel costs for non-mandatory tours, indicating a very high value of in-vehicle

travel time for the same. Consistent with findings in the social sciences and multiple streams within economics that have shown preferences to be endogenous, the case study shows that a decision-maker's value of time is sensitive to the level-of-service of the transportation system, and an increase in overall travel times can induce decision-makers to lower their value of time.

The framework is subsequently adapted to study the evolution and persistence of modality styles and travel mode choice behavior in a dynamic context. Individual modality styles are still represented as latent classes, but an individual is allowed to have different modality styles at different time periods. The evolutionary path is hypothesized to be a Markov process such that an individual's modality style in the current time period is dependent only on her modality style in the previous time period. As before, travel mode choices for a particular time period are conditioned on the individual's modality style for that time period. The framework is empirically tested using travel diary data collected in Santiago, Chile. The dataset comprises a sample of 220 individuals surveyed over four one-week periods that span a time period of twenty-two months that includes the introduction of Transantiago, a complete redesign of the city's public transit system. Estimation results identify three modality styles: unimodal auto users who only consider the automobile, unimodal transit users who only consider the public transit system and have a low value of time, and multimodal users who consider all travel modes and have a high value of time. The case study further finds that the distribution of individuals across modality styles is highly sensitive to a shock to the transportation system such as that represented by the introduction of Transantiago. Results from a sample enumeration show that nearly a quarter of the sample population changed its modality style post-Transantiago.

For all three datasets, estimation results find that modality styles are strongly correlated with more long-term travel and activity decisions, such as level of vehicle ownership and residential location. In examining the influence of individual modality styles on travel mode choice behavior, the model framework for both the static and the dynamic context took one or more of these decisions as exogenous inputs. However, such a causal representation risks endogeneity, leading us to reverse the representation and include these dimensions explicitly as dependent variables. In doing so, we recognize that dimensions such as level of vehicle ownership represent decisions that are not made by individuals in isolation from other members of the household. An individual's preferences and choices are strongly shaped by the opinions and behaviors of the people around her, particularly when a choice is made collectively by a group of individuals, as in the case of a household. Therefore, interaction between household members must be understood to influence, among other things, individual modality styles. To reflect this influence, we introduce the household modality styles construct, characterized by the modality styles of the respective individuals that make up the household. We build upon the LCCM framework described previously, replacing the individual modality styles construct with the household modality styles construct and conditioning both individual and household level dimensions on the household's modality style, therefore introducing correlation between preferences of the individuals that constitute the household.

The framework is used to examine the relationship between household modality styles, level of vehicle ownership, transit season pass possession and travel mode choice behavior using travel diary data from Karlsruhe, Germany. The dataset comprises a sample of 96 male and female household heads belonging to 48 households surveyed over a six-week observation period. Estimation results identify four distinct household modality styles. The model uncovers both

significant correlation between modal preferences of heads of the same household and notable differences as well. In general, female household heads are found to be less reliant on the automobile for their mobility requirements than their male counterparts. Short-term individual decisions, such as mode choice, are found to be inextricably linked with more long-term individual and household decisions, namely level of vehicle ownership and transit season pass possession, both of which vary considerably across different modality styles.

Modality styles have important implications for transportation policies and infrastructural initiatives seeking to change existing patterns of travel mode choice behavior. Travel demand models constitute an important component of the planning and policy-making process, being widely used to make forecasts, which in turn are driven by the assumptions that these models make about how individuals arrive at decisions. We use estimation results from the BATS 2000 dataset to compare forecasts from different model specifications for scenarios evaluating the impact of increased auto congestion and improvements to the public transit system on travel mode choice behavior. Findings reveal that models of travel mode choice behavior that ignore the influence of modality styles can overestimate expected gains from transport policies and infrastructural initiatives seeking to reduce automobile use by factors of between one-and-a-half and three. The dissertation further demonstrates how incremental improvements in the transportation system, unless accompanied by corresponding shifts in the distribution of individuals across different modality styles, will result in far smaller changes in travel behavior than would be predicted by a traditional model of travel mode choice. This dissertation makes the case that what is needed is a dramatic change to the transportation system that forces individuals to reconsider how they travel.

Though the applications presented in the dissertation restrict their attention to the influence of modality styles on four specific dimensions of individual and household travel and activity behavior - travel mode choice for work/mandatory tours, travel mode choice non-work/non-mandatory tours, transit season pass possession and level of vehicle ownership, our results serve as a good starting point for a more comprehensive framework that recognizes the influence of modality styles on all dimensions of individual and household travel and activity behavior. The model framework developed by this dissertation is shown to be both methodologically flexible and empirically robust. Using a model of individual modality styles and travel mode choices in a static context as the foundation, we were able to expand the framework in multiple directions, extending it to a dynamic context, including additional dimensions of decision-making such as transit season pass possession and level of vehicle ownership, and incorporating the influence of intra-household interactions on individual preferences. Despite differences in observation period, sample size, local topography and cultural context across the three datasets, the framework was consistently found to outperform traditional models of travel behavior in terms of both statistical measures of fit and behavioral interpretation. The dissertation concludes with a discussion on how the framework might be extended further to include dimensions such as destination choice, vehicle miles traveled and residential location. We identify some of the major hurdles to their inclusion and suggest possible solutions, laying out an extensive road map for future research in the area. When complete, the line of work initiated by this dissertation is expected to result in a comprehensive model of individual and household travel and activity behavior that integrates travel demand and land use analysis through the modality styles construct with the objective of offering a deeper understanding of decision-making and greater predictive power than current models in practice.

*To Nanu,  
and the memory of breaking oranges.*

# Contents

<b>List of Figures</b>	<b>iv</b>
<b>List of Tables</b>	<b>vi</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.2 Objective	4
1.3 Travel Demand Models: State-of-the-Practice	4
1.3.1 Taste Heterogeneity	5
1.3.2 Choice Set Generation	7
1.3.3 Preference Endogeneity	8
1.3.4 Simultaneous Choice Models	10
1.3.5 Models of Group Decision-Making	12
1.4 Contributions	13
1.5 Dissertation Outline	14
<b>2 Methodological Framework</b>	<b>15</b>
2.1 Methodological Shortcomings of Existing Travel Demand Models	15
2.2 Proposed Methodological Framework	16
2.3 Contributions to the Literature	22
<b>3 Individual Modality Styles</b>	<b>24</b>
3.1 Case Study I: Karlsruhe, Germany	24
3.1.1 The Dataset	25
3.1.2 Descriptive Analysis	26
3.1.3 Model Specification	28
3.1.4 Estimation Results	29
3.1.5 Conclusions	34
3.2 Case Study II: San Francisco Bay Area, United States	34
3.2.1 The Dataset	35
3.2.2 Model Specification	36
3.2.3 Estimation Results	36
3.2.4 Conclusions	41
<b>4 Policy Implications</b>	<b>43</b>
4.1 Scenario I: Increased Auto Congestion	43
4.2 Scenario II: Improvements to the Public Transit System	47
4.3 Conclusions	53
<b>5 Dynamics of Modality Styles</b>	<b>54</b>
5.1 Literature Review	54
5.2 Proof of Concept	55

5.2.1	The Dataset	55
5.2.2	Model Framework	58
5.2.3	Estimation Results	61
<b>6</b>	<b>Household Modality Styles</b>	<b>65</b>
6.1	Literature Review	66
6.2	Proof of Concept	67
6.2.1	The Dataset	68
6.2.2	Model Framework	72
6.2.3	Estimation Results	75
<b>7</b>	<b>Conclusions</b>	<b>82</b>
7.1	Findings and Contributions	82
7.2	Directions for Future Research	85
7.3	Concluding Remarks	90
	<b>Bibliography</b>	<b>91</b>

## List of Figures

1.1	Car advertisements taken from American publications over the second half of the last century	2
1.2	Schematic diagram of the San Francisco Chained Activity Modeling Process (SF-CHAMP)	11
3.1	Scatter plot of individual mode shares for work and non-work tours for the MOBIDRIVE dataset	27
3.2	Model of travel mode choice for the MOBIDRIVE dataset	28
3.3	Model of travel mode choice for the BATS 2000 dataset	35
3.4	Illustrations showing the kinds of individuals that belong to each of the six modality styles and their latent preferences for different travel modes	38
4.1	Change in the percentage of tours, mandatory and non-mandatory, made by car (drive alone and shared ride) across different modality styles as a function of increasing car travel times, as predicted by the six-class LCCM with feedback	45
4.2	Change in the distribution of modality styles across the sample population as a function of increasing car travel times, as predicted by the six-class LCCM with feedback	46
4.3	The mean value of in-vehicle travel time across the sample population for mandatory tours and non-mandatory tours as a function of increasing car travel times, as predicted by the six-class LCCM with feedback through consumer surplus	47
4.4	Change in the percentage of mandatory tours (top) and non-mandatory tours (bottom) made by car (drive alone and shared ride) as a function of increasing travel times, as predicted by the multinomial logit model with sociodemographic variables, the six-class latent class choice model without feedback, and the six-class latent class choice model with feedback through consumer surplus	48
4.5	Change in the percentage of mandatory tours (left) and non-mandatory tours (right) made by car (drive alone and shared ride) as a function of proposed improvement to the public transit system, as predicted by the multinomial logit model with sociodemographic variables, the six-class latent class choice model without feedback, and the six-class latent class choice model with feedback through consumer surplus	49
4.6	Change in the distribution of individuals across modality styles as a function of proposed improvement to the public transit system, as predicted by the six-class latent class choice model with feedback through consumer surplus	50
4.7	The required distribution of individuals across modality styles that would produce an equivalent shift in travel mode shares away from the automobile as each of the three proposed public transit improvements	51
5.1	Shifts across travel modes between subsequent waves of the Panel	57
5.2	An HHMM showing the influence of individual modality styles on travel mode choice behavior over time	58

5.3	Results from a sample enumeration illustrating the distribution of survey respondents across the three modality styles at each of the four waves comprising the Santiago Panel	64
6.1	Scatter plot of mode shares for work tours for the male and female household head; each house figure on the plot represents one household in our sample	70
6.2	Scatter plot of mode shares for non-work tours for the male and female household head; each house figure on the plot represents one household in our sample, except for the transit and bike plots where the number of individuals in the lowest quintile is written into the plot inside the white circle. The grey diagonal band is to help identify households where the two household heads have relatively similar mode shares.	71
6.3	Model of travel mode choice, transit season pass possession and level of vehicle ownership for the MOBIDRIVE dataset	72
6.4	Sample enumeration results for travel mode choice for work and non-work tours as predicted by the preferred four-class model specification	79
7.1	The influence of modality styles in the context of a comprehensive forecasting model system (adapted from Bowman, 1995)	83
7.2	A comprehensive travel demand model depicting the influence of household and individual modality styles on residential location, car ownership and travel mode choice behavior; other dimensions of individual and household travel and activity behavior, such as destination choice or vehicle miles travelled, can be plugged into the model framework analogously	87

## List of Tables

3.1	Summary statistics for different model specification	29
3.2	Class-specific travel mode choice mode for work tours	31
3.3	Class-specific travel mode choice mode for non-work tours	31
3.4	Class membership model	32
3.5	Aggregate elasticity of demand with respect to travel times	33
3.6	Summary statistics for different model specification	37
3.7	Class-specific travel mode choice model for mandatory tours	39
3.8	Class-specific travel mode choice model for non-mandatory tours	39
3.9	Class membership model	40
5.1	Summary statistics for different model specifications	62
5.2	Class-specific travel mode choice model	63
5.3	Class membership model	63
6.1	Summary statistics for different model specifications	76
6.2	Class-specific travel mode choice model for work tours	77
6.3	Class-specific travel mode choice model for non-work tours	77
6.4	Class-specific choice model for transit season pass possession	78
6.5	Class-specific choice model for level of vehicle ownership	78
6.6	Class membership model	78

## Acknowledgements

There are a number of people that I'd wish to express my thanks to. First and foremost, to Joan Walker, for being so spectacular. To the other members of my committee: Paul Waddell, for both his enthusiasm and his insights regarding the work; Adib Kanfani, for his sage advice on research and academic life; and John Canny, for his ideas and his patience with my limited programming ability. To my fellow collaborators: André Carrel, Fletcher Foti and Sebastián Raveau, for their help with the datasets. To my funding sources: the University of California Transportation Center, the National Science Foundation, the California Center for Innovative Transportation and the Institute of Transportation Studies, for paying me to do what I love most. To the friends that I've made during my stay here at Berkeley: Julia, Haotian, Sebas, Darren, DJ, GKB, Jeff, Nakul, Will, Aditya, Ipsita, Stefano, Musti, Hao Lu, Xiaofei, Yi, Tierra and Rui, these last few years wouldn't have been the same without the company of any of you, for better or for worse. To friends back home, for providing me with an annual reprieve from work. To Daniel Rodriguez, for his companionship. And finally, to my parents, for their mostly unconditional love and support.

# **Chapter 1**

## **Introduction**

Individuals are fundamentally different. Empirical evidence increasingly indicates the existence of higher-level orientations, or lifestyles, that concurrently influence all dimensions of an individual's travel and activity behavior (Kitamura et al., 1997; Krizek and Waddell, 2002; Lanzendorf, 2002; Choo and Mokhtarian, 2004; and Johansson et al., 2006). Where a man lives is not a very different question to ask from how many cars does a man need. An individual's proclivity to recycle, her desire to reside in high-density mixed-use neighborhoods, and her inscrutable ability to endure a slow and uncomfortable bus ride as part of her morning commute to work every day, are different manifestations of the same system of beliefs. How that system of beliefs, or the lifestyle that the individual subscribes to, influences her travel behavior is the question motivating this research.

One of the first formal definitions of individual lifestyles was proffered by Sobel (1981), who characterized lifestyles as sets of expressive, observable behaviors, and regarded consumption as the activity that best captured different lifestyles. It was proposed by Giddens (1991) that individuals embrace differing patterns of consumption behavior not only because they fulfill varying utilitarian needs, but because they give material form to a particular narrative of self-identity. Therefore, the myriad choices that an individual is daily confronted with result in decisions not only about how to act but who to be. Within this framework, latent modal preferences, or modality styles, are introduced as that component of an individual's lifestyle that relates to travel behavior and, more specifically, travel mode choice. They may be thought of as lifestyles built around the use of a particular travel mode or set of travel modes.

Nothing is perhaps more emblematic of the American lifestyle than the automobile. The United States is home to a fifth of the world's passenger vehicles (Davis et al., 2011) and has an average ownership rate of 1.86 passenger vehicles per household (FHWA, 2009). The National Household Travel Survey (NHTS) for 2009 finds that 88% of commute trips and 83% of other trips are made by the car, with average vehicle occupancies of 1.13 and 1.74, respectively. A Federal Transit Administration report that examines local familiarity with public transportation systems across the United States (Wirthlin Worldwide and FJCandN, 2000) finds that one-in-four Americans know nothing about public transportation in their neighborhood, one-in-three Americans has never used public transportation in their lives, and only one-in-two Americans can claim complete familiarity with the local public transportation network.

The automobile's profound impact on the physical space we live in and the cultural landscape that we inhabit is inescapable. Architecture, advertising, art, cities, design, sexuality, literature, music, cinema – very little is exempt from the automobile's influence (Wollen and Kerr, 2002). Figure 1.1 shows a series of car advertisements taken from American publications over the

All over America  
the word for style is  
Studebaker

## ANNOUNCING New 1948 Studebakers

**T**HREE flies faster than most of us realize. It's just a little more than eighteen months since you first read the thrilling Studebaker announcement of the new four-door dream car in production. Now the new Studebaker has arrived.

New 1948 Studebakers, including glorious new Champion and Commander convertibles, are sweeping upon the scene with a new sense of style and beauty.

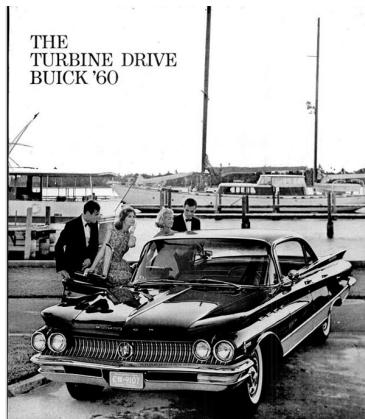
They're more than fresh 1948 interpretations of the "new look." They're the result of a year's intensive work on the domestic scene to over a year and a half of the most extensive new-car success in motoring history.

Riding in a Studebaker is like a vacation into the heart of discriminating America. Studebaker styling has established the new standard of taste and beauty.

See these latest Studebaker achievements at your first opportunity. They're superb 1948 examples of the new kind of styling in which Studebaker is impressively and so inspiringly bold.



### THE TURBINE DRIVE BUICK '60



OUTSIDE: the unmistakable look of elegance at Buick's all-time comfort design as fresh and clean as sea air. INSIDE: extravagant comfort and Buick's all-time best interior decor... richer, roomier, with more attention to thoughtful details! THE TURBINE DRIVE BUICK '60

### The new Camaro Type LT. You deserve a car this good.



For 1973, there's a new kind of Camaro... called Type LT. The "LT" is for luxury, the "T" is for touring. Inside, the Type LT is a true standard. Full foam, cloth-trimmed bucket seats. Power steering, wood-grain vinyl covering. Even a power side panel and doors and Hide-A-Way door panels. For an even quieter, more luxurious ride, insulation has been added from the front to the rear. And for the touring side of LT, there's Camaro's famous independent front suspension. Fully adjustable front and rear coil springs. Left-side remote control. And a choice of three special instrumentation packages. So now you can have your luxury and drive it, too.

1973 Chevrolet. Building a better way to see the U.S.A. Chevrolet

## Ask the person | who drives one.



1990 Grand Caravan SE: Or ask J.D. Power & Associates, they gave Caravan the highest customer satisfaction rating of any American minivan. Or ask us. We'll give you a 100% money-back guarantee if you don't like it. Only \$1,000 extra for options on a Caravan LE at no extra cost! Pretty impressive, huh?

Lynsey over there why she's in the wayback!  
"On top of this, I'm an attorney and I've got a big caseload. And I need our Caravan for that, too."  
"You mean you're driving a Caravan? Neither can we, Leslie. Neither can we." **The new spirit of Dodge.**

\*Just look around you, our Caravan is the best selling minivan of all time. Or ask J.D. Power about their 1989 Light Duty Truck Customer Satisfaction With Perfect Quality & Driver Service Survey (1988 models). Or talk to your State, who told us that you save \$1,000 on a Caravan LE with the 100% package, based on the list price of items of equal value. Didn't we? And the insurance rates are \$100 per month higher for safety.

**Figure 1.1:** Car advertisements taken from American publications over the second half of the last century: (1) A double page spread for a Studebaker from a 1948 issue of Life; (2) a full-page spread for a Buick from a 1960 issue of Life; (3) a full-page spread for a Camaro from a 1972 issue of Life; and (4) a double-page spread for a Caravan from a 1990 issue of National Geographic

second half of the last century. In chronological order, the first, a double page spread from a 1948 issue of Life, has four young and beautiful people roasting marshmallows over a beach bonfire while their car, a Studebaker, stands gleaming on the sand under the moonlit night. The second, a full page spread from a 1960 issue of Life, shows two elegantly dressed couples returning perhaps from a pierside wedding to their equally elegant Buick. The third, a full page spread from a 1972 issue of Life, depicts a family of adventurers posing in their scuba diving gear beside their Camaro. The fourth and final advertisement, a double page spread from a 1992 issue of National Geographic, shows a suburban working mother standing proudly beside her Caravan, explaining in the text below the photograph how she needs the car to "Get the kids to school, our two plus three more from down the street. With volleyball afterwards. Not to mention groceries... On top of this, I'm an attorney and I've got a big caseload. And I need our Caravan for that too."

What used to be a quintessentially American obsession has, through radio channels, television sets and the Internet, leaked into the popular imagination of cultures all around the world. The automobile's continued preeminence in much of the developed world, and its more recent proliferation in many developing countries, is a source of grave concern to the health of our cities and the global environment at large. It is widely agreed that the current pattern of human growth is unsustainable, and that a determined effort needs to be made to encourage individuals to forego driving in favor of greener modes of transportation. The growing social costs imposed by the automobile through its impacts on congestion and safety, and the increased relevance of issues of equity and livability, have together contributed to a renewed interest in alternative modes of travel, such as public transit and bicycling, and their potential to offer a more sustainable solution to our mobility requirements.

However, incremental changes in the level-of-service of alternative modes aimed at inducing a shift in travel mode choice behavior often come unstuck in the face of such firmly rooted daily patterns that revolve around the use of the automobile. Human beings are creatures of habit. When an action has been repeated frequently in stable contexts in the past, minimal thought is required to initiate, implement, and terminate it (Wood et al., 2002). Any attempt to influence choices will fail if the choices are non-deliberate (Gärling and Axhausen, 2003). For example, an increase in bus frequencies or the introduction of bike lanes is of little to no consequence to individuals who drive because they have always driven; such individuals will continue to drive even when new information has changed the contextual environment in which the original decision to drive might have been made (Aarts et al., 1997; Axhausen et al., 2001; Simma and Axhausen, 2003; Thøgersen, 2005; Kuhnimhof et al., 2006; Kuhnimhof, 2009). It is ironic then that what first attracts many individuals to the automobile are the ideas of free will and self-determination, but the behavior itself is sustained over time by automatic, unconscious mental processes (Bargh and Chatrand, 1999).

Travel demand models constitute an important component of the planning and policy-making process, being widely used to make forecasts, which in turn are driven by the assumptions that these models make about how individuals arrive at decisions. Existing travel demand models place an overriding emphasis on travel times and travel costs as determinants of travel mode choice, but individuals don't just drive because the car is the fastest or the cheapest mode. Some may drive because the car facilitates a lifestyle that involves a suburban single-family home with separate bedrooms for each of the kids and a safe and quiet environment in which to raise them, as amply demonstrated by the woman pictured in the 1992 Caravan advertisement. Others might drive because of the sense of freedom and independence afforded by the car, the ability to take off whenever and wherever you please, the open road as symbolic of the American dream. The automobile is not just a mode of transportation. To the people who drive one, and to those who dream of the day they can, the automobile epitomizes a way of life. Carmakers have always known this. And if transportation practitioners and policy-makers are to succeed in persuading individuals to drive less then it's imperative that travel demand models too recognize that the decision to use a particular travel mode involves a more fundamental choice between very different and divergent lifestyles.

With that as motivation, the remainder of the chapter is organized as follows: Section 1.2 formally introduces the modality styles construct; Section 1.3 states the objective of this research; Section 1.4 reviews current practice within the travel demand modelling profession;

Section 1.5 summarizes the major contributions of this research to the body of literature on travel behavior and discrete choice analysis; and Section 1.6 concludes the chapter with an outline to the dissertation.

## **1.2 Objectives**

Latent modal preferences, or modality styles, are defined as lifestyles built around the use of a particular travel mode or set of travel modes. They are hypothesized to influence all dimensions of individual and household travel and activity behavior. For example, in the context of travel mode choice, different modality styles may be characterized by the set of travel modes that an individual might consider when deciding how to travel, her sensitivity, or lack thereof, to different level-of-service attributes of the transportation (and land use) system when making that decision, and the socioeconomic characteristics that predispose her one way or another.

Individuals with different modality styles likely respond differently to transportation policies and infrastructural initiatives aimed at changing their travel and activity behavior. When considering different policy options, it is therefore important to have an understanding of the distribution of modality styles in the population and of the possible responses. As mentioned previously, travel demand models constitute an important component of the planning process. Forecasts from travel demand models are regularly employed by Metropolitan Planning Organizations to determine the required capacity that new infrastructure must satisfy, and to facilitate the economic, environmental and social impact assessments that usually accompany the debate on how to allocate funds between competing initiatives. A greater comprehension of the many factors that shape behavior and a more precise understanding of expected responses are essential to the successful design of systems that serve the immediate needs of the population while satisfying long-term objectives.

The objectives of this research are three fold: (1) to develop a travel demand model framework that captures the influence of modality styles on multiple dimensions of individual and household travel and activity behavior; (2) to test that the framework is both methodologically flexible and empirically robust; and (3) to demonstrate the value of the framework to transportation policy and practice.

## **1.3 Travel Demand Models: State-of-the-Practice**

The field of travel demand analysis has progressed far beyond the four-step models that defined the profession for much of the last century. Following McFadden's seminal work in the field of discrete choice theory during the sixties and seventies (McFadden, 2001), focus shifted from market-level data and aggregate behavior to individual-level data and more disaggregate behavior. The family of discrete choice models includes the many different model forms used to study problems that examine behavior of decision-makers faced with a choice between a finite and countable set of alternatives. The Random Utility Maximization (RUM) model has been the model of choice for studies on individual and household travel and activity behavior. The RUM model formulates the utility of each alternative as a function of the attributes of the alternative, the characteristics of the decision-maker and some stochastic component that is unobserved by

the analyst and/or purely random. The model assumes decision-makers are utility maximizing in that they choose that alternative which offers the greatest utility.

Early applications of the RUM model, both within the field of travel behavior and outside it, almost exclusively used some model form belonging to the Generalized Extreme Value (GEV) branch of models, owing largely to the computational tractability offered by these models. The multinomial logit and nested logit models proved by far to be the most popular (Carrasco and Ortúzar, 2002), earning their colloquial appellation of the workhorses of discrete choice analysis. Numerous studies have since devoted attention towards improving the specification of the logit model. Extensions include the incorporation of flexible error structures and random taste heterogeneity through the use of either the mixed logit or the multinomial probit model; the inclusion of latent variables representing latent biological, psychological and sociological constructs, such as attitudes, values, norms and affects; the introduction of latent classes to capture latent segments that differ from each other with regards to, for example, the taste parameters; the combination of stated and revealed preference data to capitalize on the benefits offered by either type of data; and the representation of individual decision-making behavior in a dynamic context to capture interdependencies between decisions made at different stages in time.

The shift towards disaggregate models of decision-making has been seen as a significant step forward. However, travel demand models currently in practice continue to be deficient in several critical ways. Over the course of the next few paragraphs, we review five major shortcomings to existing travel demand models that this dissertation seeks to address. We identify relevant recent advances within the broader body of literature on discrete choice analysis that this dissertation leverages in an attempt to address some of these shortcomings, and point out gaps that this dissertation helps partially to fill through methodological contributions of our own. Sections 1.3.1 and 1.3.2 review literature on the related ideas of taste heterogeneity and choice set generation; Section 1.3.3 discusses the often neglected issue of preference endogeneity; Section 1.3.4 describes recent studies on simultaneous choice models; and Section 1.3.5 summarizes studies that have examined group decision-making.

### **1.3.1 Taste Heterogeneity**

Differences in modality styles are expected to manifest themselves most prominently through differences in tastes. For example, in the context of travel mode choice behavior, auto-oriented individuals might be more sensitive to access, egress and waiting times than transit-oriented individuals. Similarly, individuals predisposed towards non-motorized travel modes might have a lower than average value of time. Travel demand models often represent heterogeneity in the choice process through the use of observable socioeconomic variables, such as gender and income, either as alternative-specific variables or by interacting them with level-of-service attributes (see, for example, Bowman, 1998). Since modality styles can evolve with changing socioeconomic conditions, these variables serve as useful proxies for attitudes and motivations underlying observed behavior. More recent examples that have relied on the idea of systematic taste heterogeneity include Vovsha and Petersen (2009), who proposes a framework that captures the effect of socioeconomic variables on two long-term travel and activity decisions: household car ownership and individual transit pass holding.

However, capturing heterogeneity systematically may be insufficient when tastes vary with unobservable variables or purely randomly, and can result in inconsistent parameter estimates (Chamberlain, 1980). This inadequacy has resulted in the growth in popularity of mixed logit models. Mixed logit models are continuous mixture multinomial logit models that can approximate any random utility model (McFadden and Train, 2000). They allow for random taste heterogeneity, in addition to unrestricted substitution patterns and a rich error correlation structure. Early applications of the mixed logit model to incorporate unobservable taste heterogeneity in discrete choice models involved only one or two dimensions of integration (see, for example, Ben-Akiva et al., 1993). Advances in computational power, and corresponding leaps in simulation methods, have since helped set off a veritable explosion in the development and application of these models (see Train, 2009). Numerous distributions have been employed, the most popular being the normal and the lognormal, and attempts have also been made to describe these distributions as functions of covariates to improve fit and ease interpretation. For example, Bhat (2000) specifies the means of the random parameters as functions of observed individual characteristics. Greene et al. (2006) extend this framework to include heterogeneity in the variance, or heteroskedasticity, of the random parameter distribution.

While the use of continuous mixture distributions often provides an excellent fit to the data, it has been argued that the correlation structure is a black box in that the cause of the distribution is not readily apparent (Walker and Ben-Akiva, 2011). Other criticism of the random parameters approach has drawn attention to its requirement of the analyst to make an a priori assumption about the mixture distribution for each randomly distributed coefficient (Hess and Rose, 2006). Fosgerau (2005) and Hess et al. (2005) discuss some of the deleterious effects of a wrongly specified distribution on parameter estimates and the attendant model interpretation. Since distributional assumptions exert influences of their own on the results (Hess, 2005), it has also been argued that simply knowing that a parameter is distributed randomly across respondents might be of limited utility to policy makers (Hess et al., 2009).

Efforts to overcome some of the limitations of continuous mixture models described above, and to provide insights into individual preferences, have led to interest in Latent Class Choice Models (LCCMs). LCCMs are nonparametric (or semiparametric) finite mixture discrete choice models. They were first developed in the field of marketing sciences as tools to identify relatively homogenous consumer segments that differ substantially from each other in terms of their behavior in the marketplace (Kamakura and Russell, 1989). LCCMs consist of two components: a class membership model and a class-specific choice model. The class membership model formulates the probability that a decision-maker belongs to a particular segment, or class, as some function of the characteristics of the decision-maker. Conditioned on the class that the decision-maker belongs to, the class-specific choice model formulates the probability that the decision-maker chooses a particular alternative as some function of the attributes of all of the alternatives in the choice set. Heterogeneity in the choice process is captured by allowing taste parameters (and choice sets and/or decision rules) to vary across the class-specific choice models for different classes. The popularity of LCCMs as a way of incorporating heterogeneity in disaggregate models of decision-making may be ascribed to two factors: (1) unlike parametric discrete choice model forms such as mixed logit and multinomial probit that also allow for random taste heterogeneity, LCCMs do not require the analyst to make prior assumptions about the distribution of parameters across decision-makers; and (2) the characterization of latent classes by the type of decision-makers that belong to a particular class

and the relative importance that they attach to different alternative attributes can lend greater behavioral insight to the underlying sources of heterogeneity. For these same reasons, we will be building on existing LCCMs in our attempt to incorporate greater heterogeneity within existing representations of travel and activity behavior.

### 1.3.2 Choice Set Generation

Latent modal preferences, or modality styles, are characterized as distinct segments in the population that differ from one another in terms of their awareness of and proclivity towards different modes. These differences likely manifest themselves through the alternatives that enter the individual's decision protocol. Therefore, central to the idea of modality styles is the notion of heterogeneous choice sets. For example, in the context of level of auto ownership, somebody who's strongly predisposed towards the automobile will probably not entertain the idea of not owning a car. Somebody who is more multimodal, on the other hand, might *a priori* exclude higher levels of auto ownership from the choice set. Travel demand models typically assume that individual-specific choice sets can be deterministically estimated by the analyst, thereby ignoring unobservable heterogeneity that might be ascribed to lack of information or habitual behavior. Swait and Ben-Akiva (1986) demonstrate how an incorrectly specified choice set can lead to biased estimates of individuals' sensitivity to level-of-service attributes; Cantillo and Ortúzar (2005) find that the use of standard mixed logit models with pre-specified choice sets can lead to potentially severe estimation and forecasting errors.

Choice set generation has received considerable attention in the realm of route choice behavior, where the number of possible alternatives can virtually number to infinity. Notable among these studies is the implicit availability and perception (IAP) model developed by Cascetta and Papola (2000), which penalizes the utility of an alternative based on its perceived availability. Swait (2009) proposed an ideologically similar model form where the utility of an alternative is specified as a continuous probability density function with one or two mass points, the mass points allowing for an alternative to be either extremely unattractive or entirely dominant.

An alternative formulation to the continuous probability density function is Manski's two-stage theoretical framework (1977), analogous to an LCCM where the first stage consists of estimating the probabilities of all possible subsets of the universal choice set. Since even a small number of alternatives can generate an intractable number of choice sets, early applications of Manski's formulation employed a latent captivity representation, where the simplifying assumption was made that an individual is either captive to an alternative or is free to choose from the full choice set (see, for example, Gaudry and Dagenais, 1979 and Gopinath, 1995). Swait and Ben-Akiva (1987) proposed the use of random constraints conditioned on both individual characteristics, such as car ownership, and level-of-service attributes, such as distance to the bus stop, to generate the probability that an alternative is part of an individual's choice set or not. Using combinatorics, they then estimated choice set generation probabilities for every possible subset of the universal choice set. Ben-Akiva and Boccara (1995) expanded this framework to include the influence of attitudes and perceptions regarding the availability of different travel modes.

Given the discrete nature of heterogeneity hypothesized here, namely modality styles, LCCMs are particularly appropriate (Gopinath, 1995). While previous studies have used LCCM to

capture choice set heterogeneity, here we use LCCM to capture the higher orientation of modality style by characterizing preference heterogeneity as differences across decision-makers in terms of both taste parameters and choice sets. Therefore, unlike previous studies in choice set generation that have focused on the latter at the expense of the former, our approach allows for the existence of multiple classes that might consider the same choice set and still differ from each other in terms of their behavior. Second, we do not estimate choice set generation probabilities for every possible subset of the universal choice set. Computationally it is difficult to support such a specification, particularly as the number of alternatives increase. Moreover, not every subset is behaviorally meaningful. For example, a choice model with five alternatives can potentially give rise to thirty-one distinct subsets. From the perspective of a marketer or policy-maker looking to influence behavior, knowing what kinds of individuals belong to each of these subsets might not be as helpful as knowing what the most relevant subsets are and identifying the kinds of individuals that belong to these. With this objective in mind, we adopt a more exploratory approach. We estimate multiple model specifications where we vary both the number of classes and the constraints on the choice set for each class (and, as mentioned in Section 2.2.1, the taste parameters specific to each class). We continue to increase the number of classes until we arrive at a preferred model specification, based on a comparison across both statistical measures of fit and behavioral interpretation. Usually, the final number of classes tends to be much smaller than the total number of possible choice sets. Therefore, such an approach allows us to circumnavigate the problem of dealing with an intractable number of possible choice sets and arrive at the discrete choice sets that are most prevalent in the population. Further, as emphasized in Section 1.3.4, the latent modality style is a modeled as a personal characteristic that influences multiple travel-related decisions.

### 1.3.3 Preference Endogeneity

Discrete choice models have long relied upon the neoclassical assumption that preferences, as denoted by taste parameters and choice sets, are characteristics of the decision-maker that are exogenous to the choice situation and stable over time. Though the idea that preferences might vary across decision-makers has garnered much attention over the last two decades, led largely by a surge in the popularity of model forms such as the mixed logit and the LCCM, the related notion that a decision-maker's preferences might change in response to changes in the decision-making environment has languished in relative obscurity. The assumption has never been accepted within the social sciences (Hirschman, 1982; Hollis, 1987) and has additionally been criticized by studies in both public and welfare economics (Sen, 1973; Pollak, 1978) and behavioral economics (Tversky and Thaler, 1990; Bowles, 1998). Consider, for the sake of illustration, the case of travel behavior. National average commute times in the United States have increased from 21.7 minutes in 1980 to 23.4 minutes in 1990 to 25.5 minutes in 2000 (Pisarski, 2006). In the face of worsening freeway congestion, an individual making a commute trip by car might find herself thinking, "I wish my commute took as much time as it did before, but I also wish I could continue doing what I did before." There are two ways in which such an individual could choose to respond. The individual could switch to a different travel mode, take another route, make the trip during the off-peak or go to a different work location, without changing the time that it took her before to get to work. Or, if the individual wishes to continue doing what she did before, she could lower her value of time, thereby changing her preferences in response to a change in the decision-making environment but staying consistent in terms of

her behavior. In this case, the change in preferences is driven by cognitive dissonance. In other cases, it could be motivated by such disparate mechanisms as rationalization (or the tendency to make excuses to justify otherwise unacceptable behavior), habituation (or the decrease in response to a stimulus after repeated exposure), sensitization (the opposite of habituation in that repeated exposure to a stimulus may lead to a progressive increase in response), taste acquisition (or an appreciation for things that are unlikely to be enjoyed upon initial exposure), etc.

That preferences change over time in response to changes in the decision-making environment as a result of one or more of these behavioral mechanisms has never been a point of contention; debate has largely centered on whether economists should concern themselves with such changes (Weizsäcker, 1971). On the one hand, the assumption has allowed econometricians to forecast changes in observable behavior in response to changes in one or more variables that define the decision-making environment. On the other, as Samuel Bowles (1998) writes, “If preferences are affected by the policies or institutional arrangements we study, we can neither accurately predict nor coherently evaluate the likely consequences of new policies or institutions without taking account of preference endogeneity.” Therefore, it has been argued that the use of extant models for forecasting must necessarily be limited to short term horizons over which preferences can reasonably be assumed to be insensitive to changes in the decision-making environment. Depending upon the nature of the problem, forecasting horizons may vary anywhere between a week and several years. As an extreme example, Metropolitan Planning Organizations employ travel demand models estimated with cross-sectional travel diary data collected over one or two days to predict changes in travel demand and land use patterns over planning horizons of 20-30 years. Most would agree that the assumption that individual and household modality styles indicative of latent modal preferences are immune to changes in the transportation and land use system over a period of two or three decades is one that is not reasonable.

Very little work has been done towards incorporating preference endogeneity in discrete choice models. The only study that we are aware of is by Zhao (2009), who uses Integrated Choice and Latent Variable (ICLV) models to test the reciprocal influence between preferences and behavior in the context of car ownership. As Zhao posits, depending upon the nature of the latent factors underlying individual preferences and the observable behavior of interest, causal relationships may lead from preferences to behavior, from behavior to preferences, or be significant in both directions concurrently. In Zhao’s case, estimation results find that the relationship is indeed bidirectional. The pride that an individual derives from car ownership and the individual’s perception of the relative convenience of using a car against using public transit is found to contribute positively to car ownership, but the individual’s willingness to pay taxes to help protect the environment is found to contribute negatively to the same. In the reverse direction, high car ownership is found to decrease the individual’s willingness to pay taxes to help protect the environment and increase the individual’s perception of the relative convenience of car against public transit.

Our hypothesis differs from Zhao’s in that we argue that preferences are not a function of observable behavior per se, but rather of the decision-making environment in which the behavior is observed. Towards this end, we will be building on LCCMs to allow for preference endogeneity. Conventional LCCMs formulate class membership as some function of the decision-maker’s characteristics, but they ignore the impact of alternative attributes, which usually enter the class-specific choice models, on class membership. We introduce LCCMs with

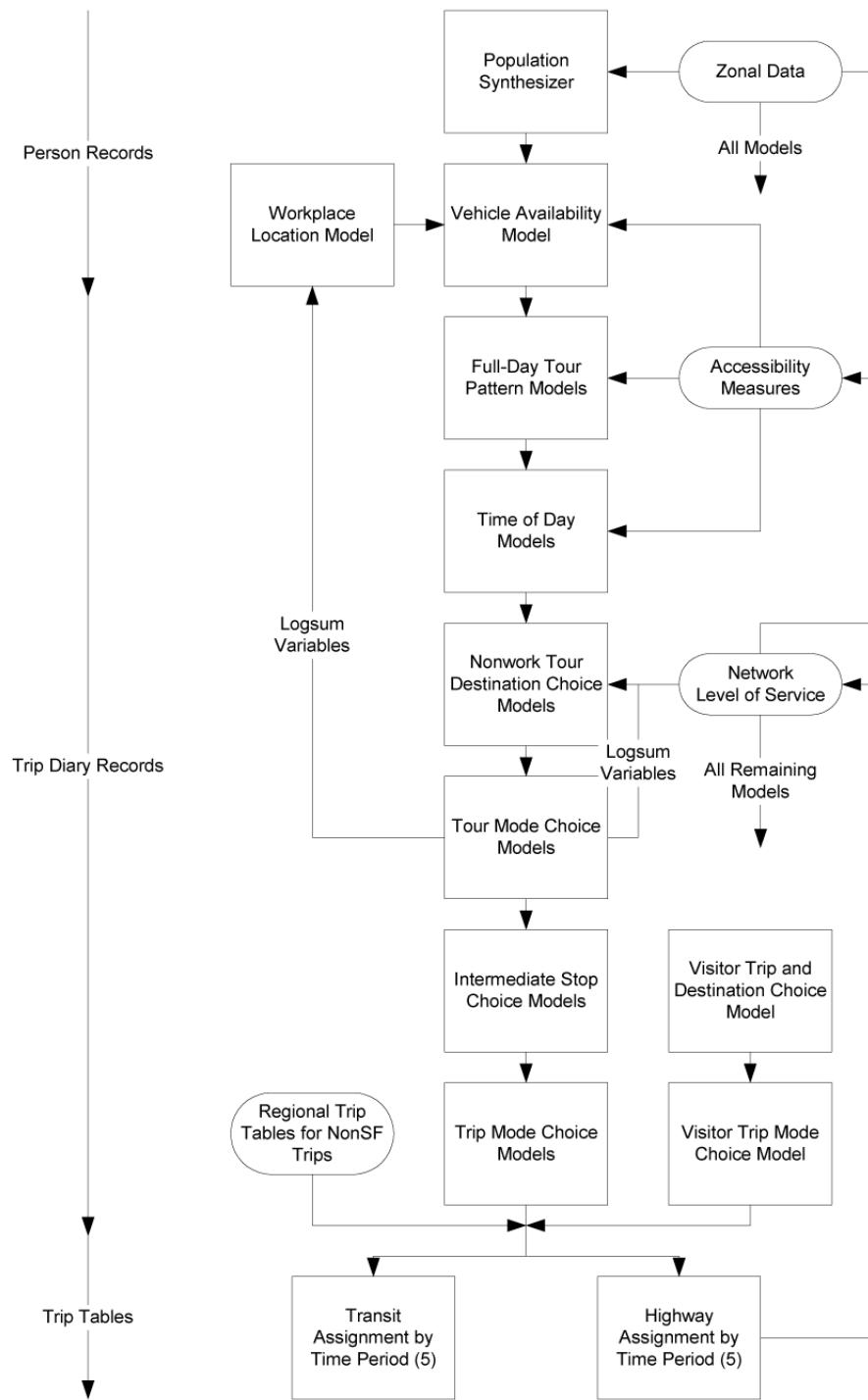
feedback from the class-specific choice models to the class membership model through the construct of consumer surplus. Class membership is hypothesized to be a function not only of the characteristics of the decision-maker but also of the consumer surplus offered by each class, which in turn is a function of alternative attributes, taste parameters and choice sets. Since classes differ from each other with regards to taste parameters and choice sets, different classes will offer differing levels of consumer surplus for the same objective choice situation. Furthermore, changes in alternative attributes will result in unequal changes in the consumer surplus offered by each class, which in turn will change the probability that a decision-maker belongs to a particular class, and consequently the distribution of decision-makers across classes over the forecasting horizon. Consequently, LCCMs with feedback allow preferences to be endogenous to the decision-making environment.

### 1.3.4 Simultaneous Choice Models

Travel demand models currently in practice represent different dimensions of individual and household travel and activity behavior as a series of sequentially nested logit models. Lower dimensions, such as travel mode choice, are conditioned on purportedly higher dimensions, such as destination choice and vehicle availability, creating a vertical chain of inter-connected nests that in their totality represent an individual's travel and activity behavior. Figure 1.2 shows a schematic diagram of the San Francisco Chained Activity Modeling Process (SF-CHAMP), the state-of-the-art travel demand model developed for the San Francisco County Transportation Authority (SFCTA) to forecast changes in travel demand for different planning applications (Cambridge Systematics, 2002).

Though nested logit model systems such as SF-CHAMP are convenient from the standpoint of estimation, they ignore the concurrent influence of modality styles on all dimensions of individual and household travel and activity behavior. Different modality styles are ultimately expected to manifest themselves through their effect on more long-term decisions, such as where to live and whether to buy a car. For example, individuals predisposed towards the automobile are likely to own a greater number of cars, and their auto-oriented lifestyles are perhaps best served by moving to auto-oriented suburban environments. Similarly, transit-oriented high-density urban developments with mixed land use probably hold a greater draw for individuals with modality styles that lean towards alternative modes of travel, such as transit, bicycling or walking. A sequential representation of decision-making would be impervious to the influence exerted by modality styles through confounding factors such as residential self-selection.

Recent interest in simultaneous choice models has prompted the use of mixture distributions to correlate choices across multiple dimensions. Bhat and Guo (2007) analyze the effects of the built environment on car ownership through the means of a mixed multinomial logit-ordered response structure. In a similar study by Pinjari et al. (2007), a simultaneous mixed logit model of residential location and mode choice for work tours is used to examine the effects of residential self-selection on the latter. Eluru et al. (2010a) develop a joint multiple discrete continuous extreme value (MDCEV) framework that models an individual's choices across the following five choice dimensions: activity type, time of day, mode, destination, and time use



**Figure 1.2:** Schematic diagram of the San Francisco Chained Activity Modeling Process (SF-CHAMP), taken from Cambridge Systematics (2002)

allocation. Eluru et al. (2010b) employ a copula-based framework to capture unobservable correlation between residential location and car ownership on the one hand, and vehicle miles traveled on the other.

However, criticism of the random parameters approach mentioned in Section 1.3.1 applies to each of these studies as well. Though the use of mixture distributions to correlate choices across multiple dimensions may result in a significant improvement in fit, the covariance structure is a black box in that it does not offer any insight to the underlying cause of correlation. Efforts to overcome some of the limitations of mixture models described above have led to interest in the behavioral mixtures approach (Walker and Ben-Akiva, 2011). Behavioral mixtures models employ latent constructs to represent the influence of higher-level attitudes and orientations on the choice process, such as modality styles in our case, and provide a behavioral rationale to the mixture distribution. The mixture distributions can be discrete or continuous. This study argues that any sample population may be decomposed into discrete segments that differ in their awareness of and proclivity towards different travel modes, and that these differences are indicative of an overarching modality style that influences all dimensions of an individual's travel and activity behavior. Though the focus of this dissertation will be on capturing correlation across three specific dimensions: travel mode choice for work tours, travel mode choice for non-work tours and level of vehicle ownership, the framework will be developed such that it can readily be extended to include other dimensions of individual and household travel and activity behavior, such as residential location and extent of daily travel.

### **1.3.5 Models of Group Decision-Making**

Individuals do not act in isolation from other members of their household. Rather, household members interact continuously with each other through the allocation of shared responsibilities (Townsend, 1987; van Wissen, 1989; Golob and McNally, 1997; Gliebe and Koppelman, 2002; Srinivasan and Bhat, 2005; and Wang and Li 2009); the division of common resources (Golob et al., 1996; Petersen and Vovsha, 2006; and Roorda et al., 2009); and joint participation in activities (Chandrasekharan and Goulias, 1999; Vovsha et al., 2003; and Kato and Matsumoto, 2009). However, interaction between household members is not limited to or defined by these three facets alone. Long-term decisions such as vehicle ownership and residential location are often made at the household-level. An individual's preferences and choices are strongly shaped by the opinions and behaviors of the people around her (Thorndike, 1938; Davis, 1976; Rose and Hensher, 2004; Zhang et al., 2009), particularly when the choice is made collectively by a group of individuals, as in the case of a household. Therefore, interaction between household members must also be understood to influence attitudes and beliefs towards, among other things, individual travel and activity behavior.

The body of literature that has focused on intra-household interactions in the context of the three dimensions mentioned previously is extensive. For recent reviews of the literature, the reader is referred to special issues of *Transportation* (Bhat and Pendyala, 2005) and *Transportation Research Part B: Methodological* (Timmermans and Zhang, 2009). However, there remains little work that has examined the effect of intra-household interactions on the formation and persistence of latent individual modal preferences. The subject of group decision-making has attracted some attention in the marketing sciences (Corfman and Lehman, 1987; Arora and

Allenby, 1999; Aribarg et al., 2010). These studies have argued that different individuals in the group exercise differing influence with regards to each of the attributes of the decision-making environment, and as a consequence the preferences of the group are likely to differ from the preferences of the individuals that constitute the group. Our work builds on studies in both of these domains through the use of hierarchical LCCMs that capture the interplay between the modality style of the household as a single unit of decision-making, the modality styles of each of its constituent individual members, and both long-term and short-term travel and activity decisions made variously at either the level of the household or the individual.

## 1.4 Contributions

This dissertation makes contributions along three key directions. First, the dissertation contributes to the body of literature on travel and activity behavior through the development of a travel demand model framework that explicitly recognizes the influence of latent modal preferences, or modality styles, that are indicative of more overarching differences in travel-related lifestyles. As Ryuichi Kitamura (1988) wrote in his seminal paper on lifestyles and travel behavior, “To deal with the challenge to urban transportation, one must first recognize that congestion is not the problem, but merely a symptom. The true problem is the life-style to which Americans aspire; the American dream is to live in a suburban single-family house on a half-acre lot with a three-car garage. If this is the root of the urban transportation problem, then obviously a fundamental solution to the issue of congestion cannot be reached without addressing the question of life-style.” This dissertation lays down the framework for a travel demand model that can facilitate the kind of analysis that Kitamura thought necessary two-and-a-half decades ago. The failure of transportation policies and infrastructural investments to overturn existing lifestyles built around the use of the automobile since then serves to demonstrate the importance of this work. The framework is applied to two very distinct travel diary datasets from two very culturally and geographically distinct regions. The first dataset was collected in Karlsruhe, Germany and comprises a relatively small sample of 119 individuals surveyed over a fairly long observation period of six weeks. The second dataset was collected in the San Francisco Bay Area in the United States and consists of a relatively large sample of 26,350 individuals surveyed over a fairly short observation period of two days. Estimation results indicate that the framework works equally well for both kinds of datasets, attesting to its robustness. Though the applications presented in the dissertation restrict their attention to the influence of modality styles on four specific dimensions of individual and household travel and activity behavior - travel mode choice for work/mandatory tours, travel mode choice non-work/non-mandatory tours, transits season pass possession and level of vehicle ownership, the framework can readily be extended to include additional dimensions, such as vehicle miles traveled and residential location.

Second, the dissertation contributes to the broader realm of discrete choice analysis. In developing a travel demand model framework that is both methodologically flexible and empirically robust, the dissertation synthesizes recent advances in the sub-domains of taste heterogeneity and choice set generation and contributes methodologically to the sub-domains of preference endogeneity (models in the past have assumed that preferences are characteristics of the decision-maker that are exogenous to the choice situation and stable over time; our framework characterizes preferences as endogenous to the choice situation and susceptible to change in response to one or more changes in the decision-making environment), simultaneous

choice models (choices across multiple dimensions have usually been correlated using black box representations that have overlooked the concurrent influence of behavioral constructs, such as modality styles in our case, on multiple dimensions of individual and/or household behavior), dynamic choice models (past studies have modeled the evolution of latent variables and observed choices over time but, unlike the framework proposed by this study, have not accounted for the influence of external conditions on the same) and models of group decision-making (our framework incorporates the influence of intra-household interactions on individual preferences).

And third, the dissertation contributes to transportation policy and practice. Travel demand models constitute an important component of the planning and policy-making process, being widely used to make forecasts, which in turn are driven by the assumptions that these models make about how individuals arrive at decisions. Findings from this dissertation reveal that models of travel mode choice behavior that ignore the influence of modality styles can overestimate expected gains from transport policies and infrastructural initiatives seeking to reduce automobile use by factors of between one-and-a-half and three. The dissertation further demonstrates how incremental improvements in the transportation system, unless accompanied by corresponding shifts in the distribution of individuals across different modality styles, will result in far smaller changes in travel behavior than would be predicted by a traditional model of travel mode choice. This dissertation makes the case that what is needed is a dramatic change to the transportation system that forces individuals to reconsider how they travel.

## **1.5 Dissertation Outline**

The dissertation is structured as follows: Chapter 2 presents the methodological framework that forms the basis of the different model forms developed through the remainder of this thesis – Latent Class Choice Models (LCCMs) with feedback through consumer surplus. LCCMs with feedback allow for preference heterogeneity (in terms of both taste parameters and choice sets) and preference endogeneity within existing representations of disaggregate behavior across multiple dimensions of interest that may or may not correlate with each other. Chapter 3 applies the framework to the study of individual modality styles and its influence on two specific dimensions of behavior: travel mode choice for work tours and travel mode choice for non-work tours, using travel diary datasets from Karlsruhe, Germany and the San Francisco Bay Area, United States. Chapter 4 analyzes two different policy scenarios as a way of illustrating some of the more important implications of our findings for transportation policies seeking to force a change in existing patterns of travel mode choice behavior. Chapter 5 adapts the framework developed in Chapter 2 to model the relationship between individual modality styles and travel mode choice behavior in a dynamic context. The framework is tested using travel diary data collected in Santiago, Chile over four one-week waves spanning a period of twenty-two months. Chapter 6 extends the framework developed in Chapter 2 to the level of the household, incorporating advances in models of group decision-making to explore the relationship between household modality styles, level of vehicle ownership, transit season pass possession and travel mode choice behavior. The framework is applied to the travel diary dataset from Karlsruhe, Germany. Finally, Chapter 7 concludes the dissertation with a summary of findings, contributions and directions for future research.

# **Chapter 2**

## **Methodological Framework**

The objective of this chapter is to develop a methodological framework that explicitly allows for both preference heterogeneity and preference endogeneity within existing representations of disaggregate behavior across multiple dimensions of interest that may be correlated with one another. We do so by building on conventional LCCMs used previously in the literature. The Chapter is structured as follows: Section 2.1 reviews major shortcomings to models of travel and activity behavior currently in practice that the framework presented in this chapter seeks to address; Section 2.2 presents the framework in full detail; and Section 2.3 concludes the chapter with a discussion of the contributions of the framework to extant literature.

### **2.1 Methodological Shortcomings of Existing Travel Demand Models**

Traditional models of travel mode choice assume that individuals are aware of the full range of alternatives at their disposal, and that a conscious choice is made based on a tradeoff between perceived costs and benefits associated with level-of-service attributes, and individual and household characteristics. Heterogeneity in the choice process is typically represented as systematic taste variation or random taste variation to incorporate both observable and unobservable differences in sensitivity to attributes. Often though, these models overlook the effects of inertia, incomplete information and indifference that are reflective of more profound individual variations in lifestyles built around the use of different travel modes. Given that lifestyle is a longer-term and partially subconscious choice, we argue that the assumption that individuals choose their mode of travel independently for every trip or tour likely does not hold true. Instead we introduce the construct of latent individual modal preferences, or individual modality styles, characterized by a certain travel mode or set of travel modes that an individual habitually uses.

For example, consider a unimodal auto user who views the world from behind the steering wheel, imagining distances in terms of driving times and locations in terms of parking availability. A unimodal auto user might not be aware of the alternatives at his disposal, or chooses not to consider them, irrespective of the nature of the trip. He knows merely to drive. At the other end of the spectrum, we have a multimodal user who thinks of the available destinations in conjunction with their accessibility by different modes, and optimizes her choice of mode prior to every trip. Even within multimodal users, there might be some who are more sensitive to out-of-vehicle time (i.e. access, egress and waiting times) and some who are more sensitive to in-vehicle travel time. Irrespective of what modality style an individual subscribes to, it is hypothesized that the individual's modality style is inextricably linked with other short and long-term travel decisions, and individuals with different modality styles likely respond differently to policies aimed at changing their travel and activity behavior. When considering

different policy options, it is therefore important to have an understanding of the distribution of modality styles in the population and of the possible responses.

A second shortcoming to travel demand models currently in practice has been their reliance on the neoclassical assumption that preferences, as denoted by taste parameters and choice sets, are characteristics of the decision-maker that are exogenous to the choice situation and stable over time. The assumption is perhaps as old as the first economic model of decision-making. Though the idea that preferences might vary across decision-makers has garnered much attention among econometricians over the last two decades, led largely by a surge in the popularity of model forms such as the Mixed Logit Model and the Latent Class Choice Model (LCCM) that have sought to incorporate greater heterogeneity within existing representations of the decision-making process, the related notion that a decision-maker's preferences might vary across decision-making environments has languished in relative obscurity. Consider, for the sake of illustration, that the population of interest initially comprises three modality styles: auto-oriented, transit-oriented and multimodal individuals. Changes in the level-of-service of different travel modes will affect each of the three modality styles differently. For example, increased freeway congestion will affect auto-oriented individuals the most and a reduction in transit services will affect transit-oriented individuals the most. The former may push auto-oriented individuals towards one of the other two classes, and the latter might trigger a similar effect on transit-oriented individuals. However, a failure to account for preference endogeneity implies that a traditional travel demand model would be oblivious to the potential impact of changes in the level-of-service of different travel modes on the redistribution of individuals across the three modality styles over the forecasting horizon.

Finally, as mentioned before, existing travel demand models often represent different dimensions of individual and household travel and activity behavior as a sequence of decisions made one after the other. While such a representation is convenient from the standpoint of model estimation, it overlooks the concurrent influence of modality styles on all dimensions of individual and household travel and activity behavior. These may include short-term decisions such as travel mode, destination, activity-chaining and travel time choice, medium-term decisions such as level of vehicle ownership or transit season possession, and more long-term decisions, such as residential location. For example, a habitual auto user may have a different perception of space, travel times and the activity chaining options than a habitual transit user. Similarly, a habitual auto user might be more likely to own a greater number of cars than a habitual transit user, just as the latter would be more likely to possess a transit season pass than the former. And habitual auto users would perhaps be best served by moving to suburban environments that cater to their auto-oriented lifestyles; whereas high-density urban developments with mixed land use would probably hold a greater draw for individuals with modality styles that lean towards alternative modes of travel, such as the habitual transit user. In building a model of individual and household travel and activity behavior, it is important to recognize that each of these many dimensions might be correlated.

## 2.2 Proposed Methodological Framework

LCCMs are nonparametric (or semiparametric) finite mixture discrete choice models. They were first developed in the field of marketing sciences as tools to identify relatively homogenous

consumer segments that differ substantially from each other in terms of their behavior in the marketplace (Kamakura and Russell, 1989). Advances in optimization routines and computational power and the ready availability of estimation software such as Latent GOLD Choice (Vermunt and Magdison, 2005) and Python Biogeme (Bierlaire, 2003) have together contributed to the growth and spread in the use of these models to other domains in the behavioral and social sciences, being employed by studies on subjects as varied as substance abuse (Chung et al., 2006), religiosity (Siegers, 2010), philanthropy (Brown et al., 2010) and theatre patronage (Grisolía and Willis, 2012). Within the field of travel demand analysis itself, LCCMs have been applied by studies on travel mode choice (Atasoy et al., 2011; Vij et al., 2011), vehicle ownership (Train, 2008; Hidrue et al., 2011), residential location (Walker and Li, 2007; Olaru et al., 2011), air travel (Teichert et al., 2008; Wen and Lai, 2010), freight (Puckett and Raschiute, 2010; Greene and Hensher, 2013), etc.

LCCMs consist of two components: a class membership model and a class-specific choice model. The class membership model formulates the probability that a decision-maker belongs to a particular segment, or class, as some function of the characteristics of the decision-maker. Conditioned on the class that the decision-maker belongs to, the class-specific choice model formulates the probability that the decision-maker chooses a particular alternative as some function of the attributes of all of the alternatives in the choice set. Heterogeneity in the choice process is captured by allowing taste parameters, choice sets and/or decision rules to vary across the class-specific choice models for different classes.

LCCMs currently in practice do not allow for preference endogeneity because there is no feedback from the class-specific choice model to the class membership model. Different classes value each of the alternative attributes differently. Therefore, changes in one or more of the alternative attributes will affect some classes disproportionately more than others, which might induce decision-makers to redistribute themselves across the classes. Classes that are better off in the wake of the changes will attract more decision-makers, and classes that are worse off will lose some decision-makers to these other classes. However, the absence of feedback from the class-specific choice model to the class membership model implies that a conventional LCCM would be oblivious to the potential impact of changes in the attributes of the alternatives on the distribution of decision-makers across classes over the forecasting horizon.

We propose that class membership is a function not only of the characteristics of the decision-maker but also of the consumer surplus that each decision-maker would derive from subscribing to different classes. Consumer surplus is a measure of the welfare that decision-makers gain from a choice situation. If decision-makers are utility maximizers, then consumer surplus is the expected maximum utility that a decision-maker derives from the choice situation, defined mathematically as some function of taste parameters and the choice set (in the case of multinomial logit models, consumer surplus is the familiar logsum term used in nested logit models and for welfare analysis). In an LCCM, classes differ from each other with regards to taste parameters and choice sets<sup>1</sup>. Therefore, different classes will offer differing levels of consumer surplus for the same objective choice situation. Furthermore, changes in alternative

---

<sup>1</sup> We overlook differences in decision rules, assuming for the remainder of the paper that all classes are utility maximizing. LCCMs that allow for other decision rules, such as elimination by aspects or satisficing, fall outside the scope of this study.

attributes will result in unequal changes in the consumer surplus offered by each class, which in turn will change the probability that a decision-maker belongs to a particular class, and consequently the distribution of decision-makers across classes over the forecasting horizon.

Why not use Hidden Markov Models (HMMs), which also allow class membership to change over time in response to changes in alternative attributes? LCCMs and HMMs are two very different model forms that are appropriate under very different circumstances: LCCMs are static models usually estimated using cross-sectional datasets and HMMs are dynamic models that require longitudinal datasets collected over extended periods of time for estimation. In many ways, HMMs currently in practice suffer from similar drawbacks as LCCMs currently in practice. Even though an HMM could potentially allow preferences to evolve over time, forecasting with HMMs would require that the transition matrix be stable over time. As a consequence, one cannot test the impact of specific scenarios on observable behavior unless the potential effect of these scenarios is captured implicitly in the data and, through it, the transition matrix. For the model to be truly general, the analyst would have to parameterize the transition matrix as some function of the decision-making environment, but then that leads us to the same problem of how to parameterize the class-membership model in the case of an LCCM as some function of the decision-making environment? We are not proposing LCCMs with feedback as an alternative to HMMs currently in practice. Rather, as part of future research, we intend to develop a dynamic version of LCCMs with feedback. This would essentially be an HMM where the transition matrix is parameterized as some function of the consumer surplus offered by each class over successive time periods. We return to this discussion in greater detail in Chapter 6, where we develop an HMM of travel mode choice using travel diary data collected over four one-week periods over a time frame of nearly two years.

Why do we elect to endogenize preferences through the inclusion of consumer surplus in the class membership model? For the same reasons that consumer surplus is the preferred way of linking sequential models in a nested logit framework, such as that used by most activity-based travel demand models currently in practice. In a sequentially nested logit model, each subsequent dimension is conditioned on the dimension immediately preceding it and receives feedback from the dimension immediately following it in the form of the logsum measure. In the case of LCCMs with feedback, there are only two dimensions, namely the class membership model and the class-specific choice model. Analogous to the sequentially nested logit model, the class-specific choice model is conditioned on the class membership model and the class membership model receives feedback from the class-specific choice model in the form of the logsum measure. The logsum is one among many ways to construct a composite variable that may be used to link choices across different dimensions. Ben-Akiva (1973) discusses the need to create composite variables and the comparative merits of the different ways in which an analyst may construct composite variables in the context of sequentially nested logit models. However, the arguments that Ben-Akiva uses extend straightforwardly to the case of LCCMs with feedback.

Why create composite variables? Why not incorporate alternative attributes directly as explanatory variables in the class membership model, in addition to including them as explanatory variables in the class-specific choice model? While such a representation would result in an explosion in the number of model parameters, more importantly it implies that a decision-maker has different marginal rates of substitution between the same attributes at the class-specific and class membership levels. Say, for example, that the value of time for a

particular class is 10\$/hour. If we were to specify the utility of that same class to be some direct function of travel times and costs, an increase in travel costs by \$10 could potentially have a very different impact on the utility of that class than an increase in travel times by one hour, without any behavioral justification as to why, given that an individual belonging to that class appears to value the two things equally.

Therefore, in creating composite variables the analyst must weigh each of the constituent variables by the relevant parameters to ensure that the marginal rate of substitution remains unchanged across levels. For that to be the case, the composite variable must be some function of the utilities of the alternatives belonging to the class-specific choice model. Two candidate measures appear most natural. The analyst could use the sum of the systematic utility of each alternative belonging to the class-specific choice model weighted by the class-specific probability of choosing that alternative. Alternately, the analyst could use the expected maximum utility as predicted by the class-specific choice model, which is mathematically equivalent to the logsum measure that we use to denote consumer surplus by. A third way would be for the analyst to use total utility from the class-specific choice model, but this is equivalent to calculating the weighted sum under the assumption that at the class membership level, the class-specific probability of choosing each alternative is the same, which is a harder assumption to justify. Therefore, we ignore it from our discussion. As Ben-Akiva (1973) shows, the weighted sum and the logsum give similar results. However, the former requires the analyst to compute both the class-specific utilities and choice probabilities when the latter only requires the analyst to compute the class-specific utilities. Therefore, the logsum is preferred.

LCCMs with feedback, like LCCMs without feedback, comprise two components as well: a class membership model and a class-specific choice model. We begin with the class-specific choice model, which predicts the probability that decision-maker n over choice dimension d and choice situation t chooses alternative j, conditional on the decision-maker belonging to latent class s, and is written as:

$$P(y_{ndtj} = 1 | q_{ns} = 1) \quad (1)$$

, where  $y_{ndtj}$  equals one if decision-maker n over choice dimension d and choice situation t chose alternative j, and zero otherwise, and  $q_{ns}$  equals one if decision-maker n belongs to latent class s, and zero otherwise. Let  $u_{ndtj|s}$  be the utility of alternative j over choice situation t for choice dimension d and decision-maker n given that the decision-maker belong to latent class s, which may be expressed as follows:

$$u_{ndtj|s} = \mathbf{x}'_{ndtj} \boldsymbol{\beta}_{ds} + \varepsilon_{ndtj|s} \quad (2)$$

, where  $\mathbf{x}_{ndtj}$  is a vector of attributes of alternative j over choice situation t for choice dimension d and decision-maker n;  $\boldsymbol{\beta}_{ds}$  is a vector of parameters for choice dimension d specific to the class s; and  $\varepsilon_{ndtj|s}$  is the stochastic component of the utility specification. Assuming that all decision-makers are utility maximizers, the class-specific choice model may be formulated as:

$$P(y_{ndtj} = 1 | q_{ns} = 1) = P(u_{ndtj|s} \geq u_{ndtj'|s} \forall j' \in C_{ndt|s}) \quad (3)$$

, where  $C_{ndt|s}$  is the choice set for choice situation t for choice dimension d and decision-maker n given that the decision-maker belongs to latent class s. LCCMs with feedback incorporate preference heterogeneity by allowing both the taste parameters  $\beta_{ds}$  and the choice set  $C_{ndt|s}$  to vary across modality styles. Depending upon the distributional assumptions regarding  $\varepsilon_{ndtj|s}$ , equation (3) can be reduced to different functional forms. For example, if we assume  $\varepsilon_{ndtj|s}$  to be i.i.d. Extreme Value across decision-makers, choice dimensions, choice situations, alternatives and latent classes with mean zero and variance  $\pi^2/6$ , then equation (3) yields the familiar probability expression for a multinomial logit model:

$$P(y_{ndtj} = 1 | q_{ns} = 1) = \frac{\exp(x'_{ndtj} \beta_{ds})}{\sum_{j' \in C_{ndt|s}} \exp(x'_{ndtj'} \beta_{ds})} \quad (4)$$

Equation (3) may be combined iteratively over alternatives and choice situations to yield the following conditional probability of observing the vector of choices  $y_n$  for decision-maker n and choice dimension d:

$$P(y_{nd} | q_{ns} = 1) = \prod_{t=1}^{T_{nd}} \prod_{j \in C_{ndt|s}} [P(y_{ndtj} = 1 | q_{ns} = 1)]^{y_{ndtj}} \quad (5)$$

, where  $T_n$  denotes the number of distinct choice situations observed for decision-maker n and choice dimension d. Moving on to the construct of consumer surplus, let  $CS_{nds}$  be the consumer surplus offered by latent class s to individual n from choice dimension d. As mentioned earlier, if decision-makers are utility maximizers then the consumer surplus is the expected maximum utility derived by the decision-maker, expressed as follows:

$$CS_{nds} = \frac{1}{T_{nd}} \sum_{t=1}^{T_{nd}} E \left[ \max_{j \in C_{ndt|s}} u_{ndtj|s} \right] \quad (6)$$

, where we've normalized consumer surplus for the number of observations  $T_{nd}$ . Since the location of the utilities corresponding to the class-specific choice models are set arbitrarily, and it is only the difference in utilities that is identifiable, the expected maximum utility is ill defined and is only valid up to a constant. In particular, different ways of setting the location of the utilities corresponding to the class-specific choice model can result in different values for the expected maximum utility. Normalizing consumer surplus for the number of observations  $T_{nd}$  ensures that the class membership model remains the same regardless of how the analyst chooses to set the location of the utilities corresponding to the class-specific choice model, as long as the analyst makes sure to include class-specific constants in the utility for each class in the class membership model. If  $\varepsilon_{ndtj|s}$  is assumed to be i.i.d. Extreme Value as before, then equation (6) results in the following logsum measure of consumer surplus:

$$CS_{nds} = \frac{1}{T_{nd}} \sum_{t=1}^{T_{nd}} \log \left( \sum_{j \in C_{ndt|s}} \exp(\mathbf{x}'_{ndtj} \boldsymbol{\beta}_{ds}) \right) \quad (7)$$

Additionally, the analyst may convert consumer surplus into units that are more meaningful than utils, such as dollars, by dividing the expression in equation (6) by the class-specific parameter corresponding to the appropriate monetary variable.

The second piece to the LCCM with feedback through consumer surplus is the class membership model, which predicts the probability that decision-maker  $n$  belongs to latent class  $s$ , and is written as:

$$P(q_{ns} = 1) \quad (8)$$

Let  $u_{ns}$  be the utility of latent class  $s$  for decision-maker  $n$ , which may be expressed as follows:

$$u_{ns} = \mathbf{z}'_n \boldsymbol{\gamma}_s + \sum_{d=1}^D CS_{nds} \alpha_{ds} + \varepsilon_{ns} \quad (9)$$

, where  $\mathbf{z}_n$  is a vector of characteristics of decision-maker  $n$ ,  $\boldsymbol{\gamma}_s$  is a vector of parameters associated with the decision-maker's characteristics,  $\alpha_{ds}$  is a parameter associated with the consumer surplus offered by the class for choice dimension  $d$ ,  $D$  denotes the number of choice dimensions and  $\varepsilon_{ns}$  is the stochastic component of the utility specification. The scale of the consumer surplus for any class is the same as the scale of the utilities for the class-specific choice model corresponding to that class. Since the scale of the utilities for class-specific choice models will differ across classes and choice dimensions,  $\alpha$  needs to be specified as a class-specific parameter that differs across different choice dimensions. Assuming as before that decision-makers are utility maximizing, the class membership model may be stated as:

$$P(q_{ns} = 1) = P(u_{ns} \geq u_{ns'} \forall s' = 1, \dots, S) \quad (10)$$

, where  $S$  is the number of latent classes in the sample population. If  $\varepsilon_{ns}$  is assumed to be i.i.d. Extreme Value across decision-makers and latent classes with mean zero and variance  $\pi^2/6$ , then equation (10) may be reduced to the following multinomial logit model:

$$P(q_{ns} = 1) = \frac{\exp(\mathbf{z}'_n \boldsymbol{\gamma}_s + \sum_{d=1}^D CS_{nds} \alpha_{ds})}{\sum_{s'=1}^S \exp(\mathbf{z}'_n \boldsymbol{\gamma}_{s'} + \sum_{d=1}^D CS_{nd{s'}} \alpha_{ds'})} \quad (11)$$

The inequality  $\alpha_{ds} \geq 0$  needs to be satisfied for all latent classes  $s = 1, \dots, S$  and all choice dimensions  $d = 1, \dots, D$  for the model to be consistent with utility-maximizing behavior. The alternative attributes  $\mathbf{x}_{ndtj}$  and the corresponding vector of parameters  $\boldsymbol{\beta}$  enter equation (10) indirectly through the consumer surplus construct. Since taste parameters and choice sets vary across classes, changes in alternative attributes will result in unequal changes in the consumer surplus offered by each class, which in turn will change the probability that a decision-maker

belongs to a particular class. Therefore, LCCMs with feedback allow preferences, as represented by taste parameters and choice sets, to be endogenous to the decision-making environment.

As with LCCMs without feedback, the number of classes  $S$  is determined by estimating models with different number of classes and using a combination of goodness-of-fit measures, such as the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC), and behavioral interpretation to select the most appropriate model. The modeling approach is exploratory in that both the number of classes and the behavior of each class emerge naturally from the process of testing different model specifications. Equations (5) and (8) may now be combined to yield the marginal probability  $P(\mathbf{y}_n)$  of observing the vector of choices  $\mathbf{y}_n$  for decision-maker  $n$ :

$$P(\mathbf{y}_n) = \sum_{s=1}^S P(q_{ns} = 1) \prod_{d=1}^D \prod_{t=1}^{T_{nd}} \prod_{j \in C_{ndt|s}} [P(y_{ndtj} = 1 | q_{ns} = 1)]^{y_{ndtj}} \quad (12)$$

Equation (12) shows how both choice situations across the same choice dimension and multiple choice dimensions for the same decision-maker are correlated through the class-membership model. Equation (12) may be combined iteratively over all decision-makers to give the likelihood function for the sample population as follows:

$$L(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}; \mathbf{y}, \mathbf{X}, \mathbf{Z}) = \prod_{n=1}^N \sum_{s=1}^S P(q_{ns} = 1) \prod_{d=1}^D \prod_{t=1}^{T_{nd}} \prod_{j \in C_{ndt|s}} [P(y_{ndtj} = 1 | q_{ns} = 1)]^{y_{ndtj}} \quad (13)$$

, where  $N$  denotes the number of individuals in the sample population. The unknown parameters  $\{\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\gamma}\}$  may be estimated by maximizing the likelihood function given by equation (13). For conventional LCCMs without feedback, the Expectation- Maximization (EM) algorithm provides a computationally robust method of optimization that takes advantage of the conditional independence properties of the model framework. The EM algorithm is particularly useful for LCCMs without feedback because in the M-step, each of the class-specific choice models and the class-membership model can be maximized independent of the other sub-models. For LCCMs with feedback through consumer surplus however, the class-specific choice models and the class-membership model can no longer be separated and maximized independently because the sub-models are joined through the consumer surplus construct. Consequentially, the EM algorithm is no more useful than traditional gradient-based optimization routines. In our case, all models were estimated in MATLAB using traditional gradient-based optimization routines.

### 2.3 Contributions to the Literature

Travel demand models constitute an important component of the planning and policy-making process, being widely used to make forecasts, which in turn are driven by the assumptions that these models make about how individuals arrive at decisions. In this chapter, we developed LCCMs with feedback from the class-specific choice models to the class membership model through the construct of consumer surplus, thereby allowing preferences to be both heterogeneous across decision-makers and sensitive to changes in the decision-making

environment as represented by changes in alternative attributes. LCCMs with feedback, like LCCMs without feedback, further allow for correlation across multiple choice dimensions through the class membership model. In developing the framework, we made use of recent advances in the fields of taste heterogeneity, choice set generation and simultaneous choice models and we contributed methodologically to the domain of preference endogeneity.

Discrete choice models have long assumed that individual preferences are exempt from the influence of the decision-making environment and therefore stable over time. That the assumption has survived as long as it has, despite notable advances in behavioral theory and computational methods, ought to be a matter of surprise. Public institutions and business corporations spend millions of dollars annually on marketing campaigns that seek actively to change individual preferences. The very existence of the arts and culture industry and, to a lesser degree, the technology industry, rests on the notion of ever evolving consumer preferences. Within the context of travel behavior itself, individual preferences could potentially be subject to the influence of changes in the level-of-service of the transportation system. And yet, extant travel demand models have continued to be deficient in their failure to account for preference endogeneity. LCCMs with feedback represent an operationalization of the notion of preferences as a “constructive, context-dependent process” (Tversky and Thaler, 1990) and allow for the use of these models for forecasting over longer horizons that are more consistent with the time scale of studies both within the field of transportation and land use behavior, and without. This point shall become clearer in Chapter 4, where we compare travel demand forecasts from LCCMs with feedback with more conventional model frameworks that do not allow for preference heterogeneity and/or preference endogeneity.

# **Chapter 3**

## **Individual Modality Styles**

To illustrate the benefits of the methodological framework developed in Chapter 2 vis-à-vis more traditional discrete choice model frameworks, we employ the framework to examine the relationship between individual modality styles and travel mode choice behavior using two very distinct datasets that differ from each other in terms of sample size, observation period, and geographical and cultural context. Section 3.1 applies the framework to travel diary data collected in Karlsruhe, Germany and comprises a relatively small sample of 119 individuals surveyed over a fairly long observation period of six weeks. Section 3.2 applies the framework to travel diary data collected in the San Francisco Bay Area in the United States and consists of a relatively large sample of 26,350 individuals surveyed over a fairly short observation period of two days. The datasets offer a unique opportunity both to compare travel mode choice behavior across different societies and to validate the robustness of the methodological framework.

### **3.1 Case Study I: Karlsruhe, Germany**

Travel demand models traditionally employ cross-sectional travel diary data recorded over one or two days, observation periods that might be too short to discern the effects of individual habits, routines and predispositions that are reflective of modality styles. Given our research objectives, a longer observation period could prove to be useful, and so the dataset that we first apply the model framework to consists of six-week travel diary surveys administered as part of the MOBIDRIVE research project in the German city of Karlsruhe (Axhausen et al., 2002). The dataset offers a unique opportunity to observe modality styles in a longer-term setting that is perhaps more consistent with the time scale of the modality styles construct.

The focus of this case study will be on capturing the influence of individual modality styles on two specific dimensions of travel behavior: travel mode choice for work tours and travel mode choice for non-work tours. Our analysis consists of two stages: First, we identify different modality styles within the sample population through a descriptive analysis, investigating possible correlations between travel mode choices for work and non-work tours. This is followed by the econometric analysis, where we apply the framework from Chapter 2 to deduce unobserved modality styles and their effects on travel behavior. The stages are linked as findings from the first stage inform the process of model development.

The Section is structured as follows: Section 3.1.1 describes the dataset in greater detail; Section 3.1.2 undertakes a descriptive analysis of the dataset that is used to inform the subsequent subsection on model specification; Section 3.1.3 elaborates upon the model specification; Section 3.1.4 presents estimation results for the preferred model specification; and Section 3.1.5 summarizes findings from the case study.

### 3.1.1 The Dataset

The dataset consists of six-week travel diary surveys administered as part of the MOBIDRIVE research project (Axhausen, 2002). The survey was conducted in the two German cities of Karlsruhe and Halle in the fall of 1999. A total of 317 persons over 6 years of age in 139 households participated in the study. The survey consisted of a face-to-face interview in which socio-demographic characteristics and household information were collected. This was followed by a self-administered travel diary survey in which participants recorded for each trip during the six-week study period the day the trip was made, trip purpose, modes used, departure and arrival times, accompanying individuals, etc. During post-processing, the level-of-service for all modes (walk, bike, auto, and transit) was generated from transportation network data for the city of Karlsruhe. More details on the survey and the resulting dataset can be found in Axhausen (2002).

Since level-of-service attributes for all modes are unavailable for the city of Halle, the dataset is narrowed to trips contained within Karlsruhe. Unlinked trips are aggregated into home-based work and non-work tours, following an approach similar to Cirillo et al. (2006). For each trip, the data contain the modal chain (including access and egress modes for transit). A “main mode” for a tour is defined to be the mode used to cover the greatest motorized distance, tacitly assuming that mode choice is dictated by the longest leg of the tour. Four main modes are defined: auto, transit, bike, and walk. Trips taken as car passengers are counted under auto, as are trips made by motorcycle (less than 2 percent). Though car passengers are expected to be different behaviorally than car drivers, data didn’t allow us to treat the two as independent travel modes. Consequently, individuals belonging to households with no cars were specified to have access to “auto,” since they could potentially get a ride from a neighbor or a friend.

In the case of work tours we consider both simple work-only tours without any additional stops, and tours on which the individual made additional stops on the way to work, on the way back, or both. However, for work tours with additional stops, the same level-of-service attributes are used in the mode choice models as those for the usual tour from home to work and back, and the presence or absence of intermediate stops is represented by a binary variable. This is the typical practice with activity-based models, where destination choice for intermediate stops is often predicated on mode choice (see, for example, Bradley et al., 2010). For instance, the location an individual chooses to stop on his way back from work to buy groceries might depend on whether he’s walking, on the bus, or in a car. Therefore, it is suspect to compare different modes for the specific tour route, since a different tour might have been undertaken had a different mode been chosen. For these same reasons, in the case of non-work tours we limit our attention to only those tours with two constituent trips, one each to and from the main destination, with no additional stops along the way. Consideration of tours with intermediate stops would call for a model that predicts destination choice as well. The absence of land use data and level-of-service attributes for all possible tours, and not merely the tour that was made, preclude estimation of such a model, and therefore non-work tours with multiple stops have been excluded. These restrictions reduce the dataset to 1445 work tours and 3359 non-work tours made between 117 individuals over the six-week observation period.

### 3.1.2 Descriptive analysis

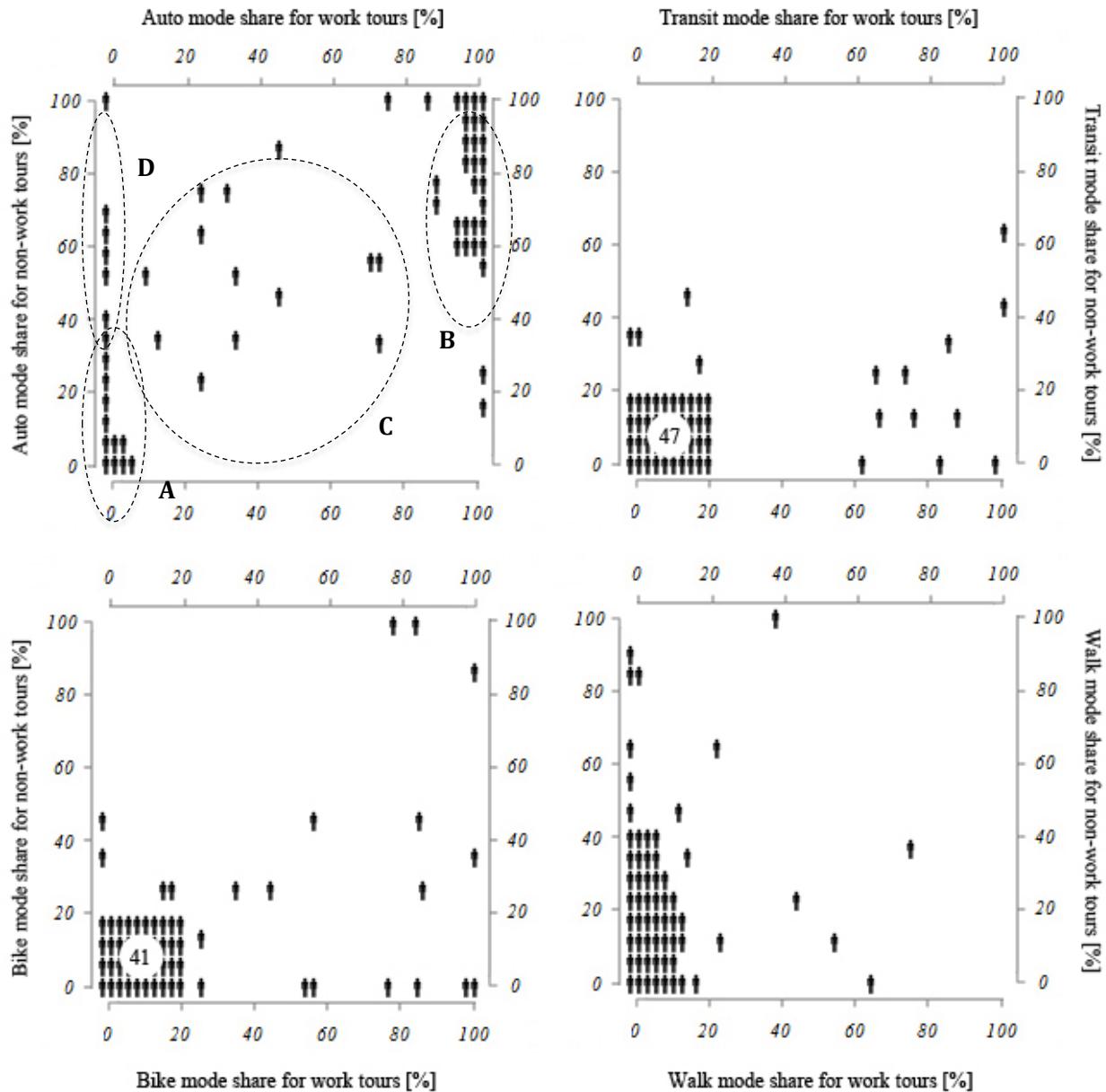
Individuals might have different modality styles for work and non-work tours. To investigate the relationship between work and non-work modality styles for a given individual, we select 62 individuals with five or more work and non-work tours each in the reduced dataset, and we plot the mode shares for the four main modes against each other. The results are presented in Figure 3.1 as a scatter plot, where each stick figure represents one individual. For the transit and bike plots, there is a substantial number of individuals that have a near zero mode share for both work and non-work tours; these are plotted in the lower left portion of the graph and we write the total number of individuals inside a white circle.

When comparing the distributions, one sees a strong polarization in the automobile data points. There are many people who use the car either very little, labeled “A”, or very much, labeled “B”, with a lower concentration of individuals elsewhere on the spectrum. The individuals in cluster B are heavily automobile-oriented, using the car for almost all tours, regardless of purpose. Of the 27 individuals with more than 90% mode share on work tours, only 2 individuals reported less than 50% auto mode share on non-work tours. On the other hand, 18 individuals take the car on less than 10% of their work tours, but their car use for non-work tours is somewhat more evenly spread out between 0% and 100%.

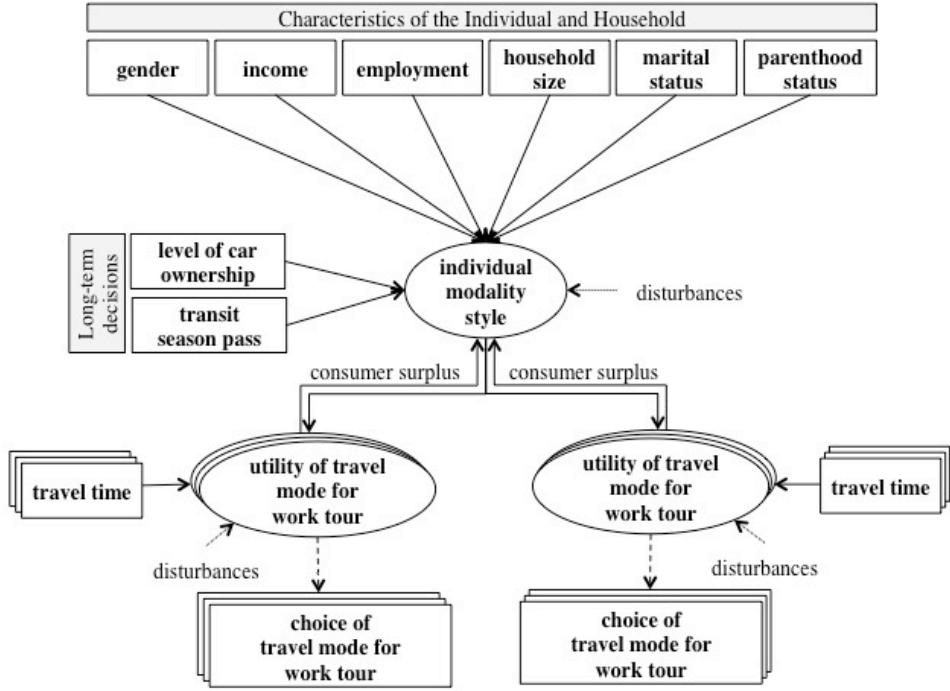
The plot also suggests a correlation between work and non-work auto use: Practically all individuals who occasionally or regularly used the automobile for work tours use it for 50% or more of their non-work tours: Of the 17 individuals in the center of the scatter plot, between approximately 10% and 90% auto use for work tours, 11 have an auto use for non-work tours of more than 50% (cluster “C”), lending reason to believe that somebody who is multimodal is likely to be so both in work and non-work tours. The scatter plot supports the concept of three or four modality classes for auto use: “quasi-unimodal auto” users (B), who are almost entirely reliant on the automobile for all of their mobility requirements, “multimodal all” users (C), who use a combination of auto and other modes, and two groups with low auto use: one that uses the automobile mostly for non-work tours (D), and one that makes little use of the automobile altogether (A).

In the case of transit, bike and walk, the most distinctive feature of the plots is the large number of people in the bottom left corner who make piddling use of either mode (so large in fact that it prevented representation by individual stick figures). Apart from that block of individuals, we find a small cluster that does not use these modes for work tours, and uses them occasionally for non-work tours. There exists a small group of individuals that relies predominantly on transit and/or bike for work tours, whereas walking enjoys greater popularity for non-work tours, reflecting the specialized nature of these modes. Compared to the scatter plot for auto, fewer data points appear in the unimodal areas.

Averaged across all tours, more than a third of the individuals in our sample population use the same mode for more than 80% of their tours during the six-week observation period. However, consistency in choices does not necessarily imply that a choice is not being made at all, and evidence of modal predispositions can also be attributed to modal availability, and temporal stability of other controlling factors, such as level-of-service attributes. At the other extreme, one



**Figure 3.1:** Scatter plot of individual mode shares for work and non-work tours. Each stick figure on the plot represents one individual in our sample, except for the transit and bike plots where the number of individuals in the lowest quintile is written into the plot inside the white circle



**Figure 3.2:** Model of travel mode choice for the MOBIDRIVE dataset

should also not conclude that observed multimodality is always due to an objective optimization of mode choice. In some cases, the multimodal all group may include individuals from households where a car is shared among drivers in the household, and therefore multimodality is more an issue of availability. These limitations are addressed more thoroughly in the following subsections on model development, where both availability and level-of-service attributes are explicitly accounted for.

### 3.1.3 Model Specification

The model specification is illustrated in Figure 3.2. Consistent with the usual notation, ellipses denote unobservable variables and rectangles denote observable variables, while dashed arrows represent measurement equations and solid arrows represent structural equations. As mentioned before, LCCMs comprise two components: a class membership model and a class-specific choice model. Individual modality styles are represented as latent classes. Class membership is hypothesized to be a function of observable household and individual characteristics, medium and long-term travel and activity decisions, and the consumer surplus offered by different modality styles. The disturbances denote unobserved factors that influence class membership, assumed to be i.i.d. Extreme Value across individuals. The class-specific choice model depicts the influence exerted by a single overarching modality style on an individual's travel mode choices across multiple work and non-work tours over time (denoted by the stacked shapes in the figure). The two choice dimensions are correlated through the modality styles construct. Travel

**Table 3.1:** Summary statistics for different model specifications

Model	Parameters	LL	$\bar{p}^2$	BIC	AIC
Mixed logit with error components					
Taste homogeneity	22	-3,732	0.390	7,650	7,508
Systematic taste heterogeneity	70	-3,596	0.404	7,786	7,332
Random taste heterogeneity	72	-3,548	0.411	7,707	7,240
Two-class LCCM					
With uniform choice sets	54	-3,539	0.416	7,536	7,186
With heterogeneous choice sets	42	-3,652	0.399	7,660	7,388
With feedback (uniform)	56	-3,537	0.416	7,548	7,185
Three-class LCCM					
With uniform choice sets	86	-3,400	0.433	7,530	6,973
With heterogeneous choice sets	66	-3,471	0.425	7,501	7,073
With feedback (heterogeneous)	69	-3,457	0.427	7,499	7,052
Four-class LCCM					
With uniform choice sets	118	-3,335	0.439	7,670	6,906
With heterogeneous choice sets	104	-3,419	0.427	7,720	7,047
With feedback (heterogeneous)	108	-3,413	0.427	7,741	7,042

mode choices are further conditioned on the travel times of the different travel mode alternatives. Unfortunately, cost data isn't available for any of the travel modes, and so no price parameters could be estimated for the model. In terms of the error structure, the class-specific choice models are mixed logit models where error components are introduced to capture serial correlation that accrues from the panel nature of our data (the stochastic component of the utility of each alternative is correlated across different travel mode choice decisions for the same individual). Heterogeneity across modality styles includes the travel modes considered, alternative-specific constants, sensitivity to travel times and the parameters corresponding to the error structure. The class membership and class-specific choice models together explicitly integrate the modality style construct with mode choice models.

### 3.1.4 Estimation Results

In determining a final model specification for the sample population, we estimated numerous models where we varied the utility specification, number of classes and choice set assumptions. Here we briefly summarize this process and present key results in Table 3.1 for twelve different models. To facilitate comparison, Table 3.1 enumerates for each model the number of parameters estimated (#), the log-likelihood of the training data, the adjusted rho-bar-squared ( $\bar{p}^2$ ), the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC). Since the adjusted rho-bar-squared and the AIC are equivalent measures of fit, we will be restricting our attention to the BIC and the AIC when comparing different models.

First, we compare the mixed logit model with error components and homogenous taste parameters to the LCCMs. As is apparent from Table 3.1, each of the nine LCCMs, with and

without heterogeneous choice sets and feedback, outperforms the mixed logit model in terms of both the BIC and the AIC. Next, we compare the LCCMs against other ways of incorporating heterogeneity in choice models (without the use of attitudinal indicators), namely a mixed logit model with systematic taste heterogeneity and a mixed logit model with random taste heterogeneity. Both are single class models that build on the framework of the mixed logit model with error components and homogenous taste parameters. The mixed logit model with systematic taste heterogeneity differs from the mixed logit model with error components in that the variables denoting individual and household characteristics and long-term travel decisions that enter each of the LCCMs through the class-membership model enter the utility specification here as alternative-specific variables. As with the mixed logit model with error components and homogenous taste parameters, each of the nine LCCMs performs better than the model specification in terms of both the BIC and the AIC. The mixed logit model with random taste heterogeneity builds on this model by specifying a lognormal distribution on the coefficient of travel time. Here, the comparison isn't as clear-cut, with the model performing better than the four-class LCCM with heterogeneous choice sets in terms of the BIC and the two-class LCCM with heterogeneous choice sets in terms of the AIC. However, by and large, the LCCMs appear to outperform the mixed logit model with random taste heterogeneity.

Finally, we limit our comparison to the nine LCCM specifications, which doesn't reveal as clear a trend. For the two-class model, the LCCM with uniform choice sets and without feedback is better in terms of the BIC and the LCCM with uniform choice sets and with feedback is better in terms of the AIC. For the four-class model, the LCCM with uniform choice sets is better than the LCCM with heterogeneous choice sets, with and without feedback, in terms of both the BIC and the AIC. For the three-class model, the LCCM with uniform choice sets is better in terms of the AIC and the LCCM with heterogeneous choice sets and feedback is better in terms of the BIC. Based both on statistical measures of fit and behavioral interpretation, we select the three-class LCCM with heterogeneous choice sets and feedback as the preferred model specification. In terms of fit, the model has the lowest BIC (and the fifth-lowest AIC, which might admittedly be somewhat high) across all specifications. In terms of the signs and relative magnitudes of the different model parameters and the accompanying behavioral interpretation of each of the latent classes, results for the three-class LCCM with heterogeneous choice sets and with feedback proved to be the most satisfying.

Tables 3.2 and 3.3 list estimation results for the class-specific travel mode choice model for work and non-work tours, respectively, and Table 3.4 lists estimation results for the class membership model for the three-class LCCM with heterogeneous choice sets (the utility of each alternative for the class-specific choice models and the class membership model was specified linear in the variables listed in each table). Since travel time is the only level-of-service variable in the class-specific choice models, in order to facilitate a comparison between estimates we enumerate in Table 3.5 the aggregate elasticity of demand with regards to travel time for each of the four travel modes across the three classes for work and non-work tours, where the measure is defined as the change in the percentage of tours made by a particular travel mode in response to a one percent increase in travel times for that travel mode.

**Table 3.2:** Class-specific travel mode choice model for work tours

<b>Variable</b>	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>
<i>Alternative-specific constants</i>			
Auto	0.000	0.000	0.000
Transit	-	-1.265	0.305*
Bike	-	-0.502 *	-0.932
Walk	1.607	2.127	0.175*
<i>Level-of-service attributes</i>			
Travel time (minutes)	-0.126	-0.100	-0.001*
<i>Heteroskedastic error components</i>			
Auto	0.000	0.726	2.252
Transit	-	1.207	1.000
Bike	-	2.541	2.981
Walk	0.881 *	2.115	1.274

\* Insignificant at the 10% level

**Table 3.3:** Class-specific travel mode choice model for non-work tours

<b>Variable</b>	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>
<i>Alternative-specific constants</i>			
Auto	0.000	0.000	0.000
Transit	-	-1.384	-1.011
Bike	-	-1.013	-1.116
Walk	3.184	2.476	0.926
<i>Level-of-service attributes</i>			
Travel time (minutes)	-0.288	-0.080	-0.014
<i>Heteroskedastic error components</i>			
Auto	0.000	0.786	1.873
Transit	-	1.904	2.291
Bike	-	2.343	1.968
Walk	4.037	1.968	0.208*

\* Insignificant at the 10% level

**Table 3.4:** Class membership model

<b>Variable</b>	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>
Class specific constant	0.000	2.818	1.810*
<i>Consumer surplus</i>			
Work tours (utils)	0.062 *	0.192 *	0.042*
Non-work tours (utils)	0.624	0.345 *	1.292
<i>Individual and household characteristics</i>			
Male	0.000	-1.820	-0.968*
Married	0.000	0.636*	-1.002*
Parent	0.000	-0.537*	-0.424*
Employed	0.000	-1.832	-0.781*
Single adult	0.000	-1.099*	-1.859
Household income (1,000 DM)	0.000	0.167*	0.152*
<i>Long-term travel decisions</i>			
Number of cars	0.000	-0.046*	-0.483*
Transit season pass	0.000	1.947	2.130

\* Insignificant at the 10% level

Over the course of the following paragraphs, we rely on the results presented in Tables 3.2-3.5 to describe in greater detail each of the three classes identified by the model (class labels are descriptive and not definitive). Estimation results for the class membership model and the class specific choice models provide information on how the classes differ from one another in terms of the kinds of decision-makers that belong to each class and the relative importance that they attach to each of the level-of-service attributes, respectively. To further underscore behavioral differences between the three classes, a sample enumeration is carried out, and the results are incorporated in our description of the three classes. The class membership probabilities for each individual are summed to arrive at the expected size of the three modality style segments. The class-specific probability of choosing an alternative on a tour is weighted by the class membership probability for the respective individual, and the product is summed over all tours to arrive at the expected modal split for each of the three modality styles. A similar procedure is used to calculate the socioeconomic composition of each class. Before we describe the classes in greater detail, it is worth reemphasizing that the estimation process is exploratory in that the number of classes and the behavior of each class are uncovered in the course of testing different model specifications. The class labels are assigned based on what the estimation results imply regarding behavior.

**1. Auto-oriented Individuals:** Comprising 16.5% of the sample population, auto-oriented individuals only consider auto and walk when deciding how to travel, with the mode split being roughly four to one in favor of auto for both work and non-work tours. In terms of the elasticity of demand with respect to travel times, relative to the other two modality styles auto-oriented individuals fall in the middle. The high elasticity of demand for walk suggests that auto-oriented individuals are willing to walk for short-distance tours. As one would expect, they have the

**Table 3.5:** Aggregate elasticity of demand with respect to travel times

Travel Mode	Class 1	Class 2	Class 3
<i>Work tours</i>			
Auto	-0.065	-0.363	-0.008
Transit	0.000	-1.353	-0.025
Bike	0.000	-1.113	-0.017
Walk	-1.199	-1.183	-0.053
<i>Non-work tours</i>			
Auto	-0.056	-0.291	-0.074
Transit	0.000	-0.876	-0.148
Bike	0.000	-1.049	-0.256
Walk	-1.093	-0.877	-0.388

highest average auto ownership rate of 1.43 cars per household. The segment largely comprises working men: 81% of the segment is employed and 76% is male. Single adults have a higher propensity to be auto-oriented as well.

**2. Choice Multimodals:** At 44.2% of the sample population, the segment considers all modes of travel and has the highest elasticity of demand with regards to travel time, though the elasticities for not all of the travel modes are greater than one. However, a higher elasticity of demand indicates that choice multimodals make an objective trade-off between different travel modes based on the travel time of each mode. Roughly one in two tours, work or non-work, is made by auto. Choice multimodals have a median auto ownership rate of 1.24 cars per household. Females are more likely to belong to the segment than men, and employment has a negative and significant influence on class membership. In some ways, choice multimodals represent one of the two halves that make a traditional household structure, the other being working men belonging to the auto-oriented segment.

**3. Captive Multimodals:** Comprising 39.3% of the sample population, this segment too considers all modes of travel but the demand for any of the modes is inelastic with regards to travel time. The high inelasticity of demand suggests that the segment comprises captive travelers who do not choose how to travel based on a trade-off between travel times for each mode. Moreover, captive multimodals will continue to act as they did before regardless of future changes in travel times for one or more travel modes. Roughly one in three tours, work or non-work, is made by auto, and captive multimodal have the lowest average auto ownership rate at 1.15 cars per household. Transit season pass possession has a positive and significant influence on class membership. In terms of sociodemographic composition, captive multimodals are evenly split between the two genders and have a high employment rate of 71%.

Individuals with different modality styles likely respond differently to policies aimed at changing travel behavior. When considering different options, it is important to have an understanding of the distribution of modality styles in the population and of the possible responses. The class membership model is rich in terms of interpretation, and serves as a harbinger of some of these potentially important implications for policy-makers. As an example, it is likely that policies

aimed at achieving a mode shift from automobile to transit (e.g., through financial incentives to buy transit passes) might bear greater fruit if targeted specifically at choice multimodals. The influence of any such policy on auto-oriented individuals is expected to be little, even though that segment of the population might account for a majority of automobile trips. However, land use and residential zoning policies could take advantage of the fact that auto-oriented individuals display a strong willingness to walk for short distance tours through the design of more walkable auto-oriented suburban neighborhoods. Alternatively, attention could be given to initiatives that promote less driving through changes in destination or travel time, but do not aim at forcing a shift in travel modes.

### **3.1.5 Conclusions**

Results from the case study are encouraging in that both the descriptive and the econometric analysis reveal that there are different groups in the population that are clearly distinguishable by their travel mode choices. Modality styles can not only be observed directly from a person's travel mode choice behavior over a period of time, but that they can also be inferred through the means of travel mode choice models using the model framework developed in Section 2. Findings indicate the presence of three broad modality styles within the sample population that differ from one another in terms of their socioeconomic make-up, their relative dependence on the automobile and their sensitivity to travel times. Results indicate the presence of auto-oriented individuals who display a strong bias for using the automobile and multimodal individuals who appear to be more open to alternative modes of travel. Multimodal individuals can further be decomposed into those who appear to be multimodal by choice and those who appear to be captive. Though the focus of the case study, and the chapter at large, is on travel mode choice, estimation results show that modality styles are strongly correlated with more long-term travel decisions and life-cycle characteristics. Finally, the model framework was tested using travel diary data collected over a six-week observation period from a sample of individuals from Karlsruhe, Germany. It would be interesting to compare the performance of the model framework using datasets from different socioeconomic, cultural and geographic contexts and different observation periods.

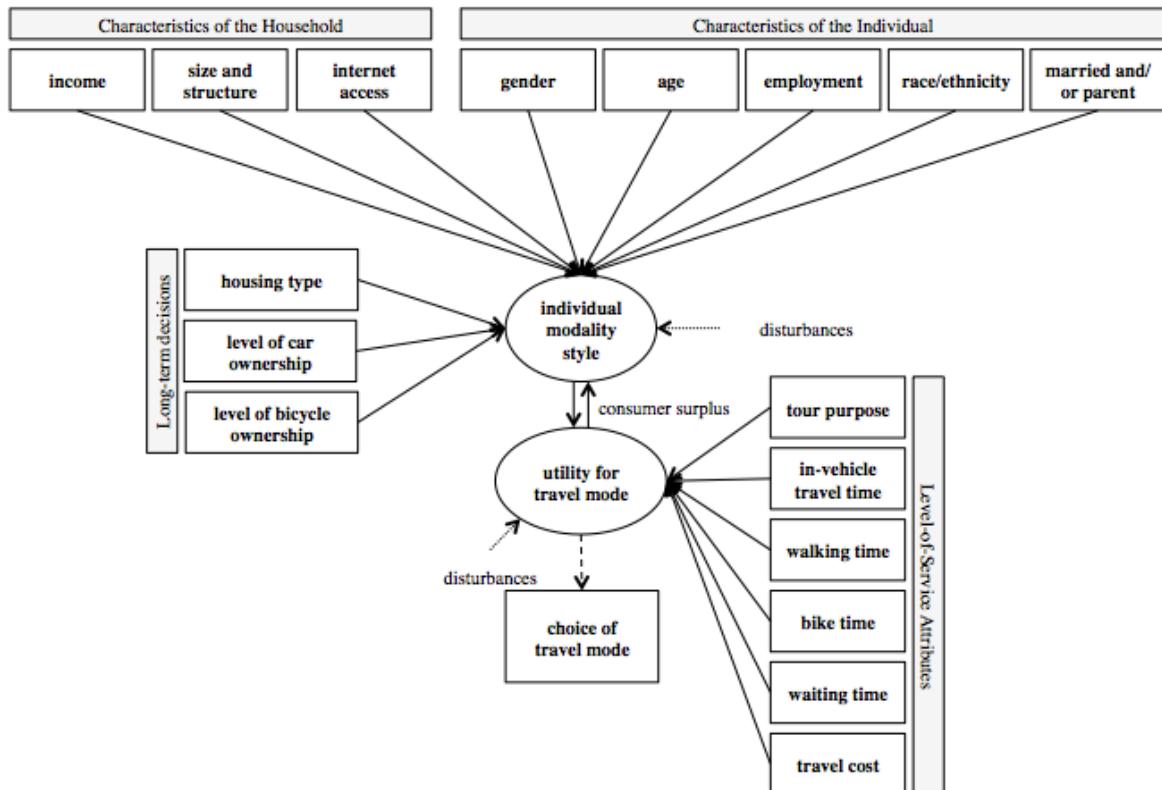
## **3.2 Case Study II: San Francisco Bay Area, United States**

Metropolitan Planning Organizations usually employ cross-sectional travel diary data recorded over one or two days in order to estimate travel demand models. It is therefore important to test if the modality styles construct can be operationalized using these more readily available one or two day travel diary datasets. Towards this end, we apply the model framework to the Bay Area Travel Survey (BATS) 2000, a household travel survey conducted in the nine county San Francisco Bay Area of California that surveyed respondents over a two-day period. The focus of this case study will be on capturing the influence of individual modality styles on two specific dimensions of travel behavior: travel mode choice for mandatory tours and travel mode choice for non-mandatory tours, where mandatory tours include trips to work and/or school and non-mandatory tours include all other tours.

The Section is structured as follows: Section 3.2.1 describes the dataset and model specification in greater detail; Section 3.2.2 presents the model specification; and Section 3.2.3 discusses estimation results for a six-class model of travel mode choice with feedback from the class-specific models to the class membership model; and Section 3.2.4 concludes the section with a comparison between findings from the two case studies.

### 3.2.1 The Dataset

Data for our analysis comes from BATS 2000, a large-scale regional household travel survey conducted in the nine county San Francisco Bay Area of California. The San Francisco Metropolitan Transportation Commission (SFMTA) has periodically sponsored BATS to provide data to support travel modeling and analysis of regional travel behavior. The target data collection period for BATS 2000 was of course the 2000 calendar year. The survey consisted of an activity-based travel diary that requested information on all in-home and out-of-home activities over a two-day period, including weekday and weekend pursuits. In all, more than 15,000 households participated in the survey. Travel diary data for 29,964 mandatory home-based tours and 29,889 non-mandatory home-based tours made by 26,350 individuals from 12,634 households was used for model estimation, and travel diary data for an additional 3,642



**Figure 3.3:** Model of travel mode choice for the BATS 2000 dataset

mandatory home-based tours and 3,283 non-mandatory home-based tours made by 2,931 individuals from 1,405 households was used for model validation. More information on the raw data can be found in Morpace International, Inc. (2002). We employed San Francisco County Transportation Authority's version of the dataset that used travel skims in order to construct for each choice situation the feasible choice set and the level-of-service attributes of each of the travel modes contained within the choice set.

### **3.2.2 Model Specification**

Figure 3.3 shows the model framework. Individual modality styles are represented as latent classes. Class membership is hypothesized to be a function of observable household and individual characteristics, medium and long-term travel and activity decisions, and the consumer surplus offered by different modality styles. The disturbances denote unobserved factors that influence class membership, assumed to be i.i.d. Extreme Value across individuals. Travel mode choices are conditioned on individual modality styles and on observable attributes of the different modal alternatives. Consistent with travel demand models employed by the SFMTA, six modal alternatives are defined: drive alone, shared ride, walk, bike, walk to transit and drive to transit. Heterogeneity across modality styles includes both the travel modes considered (the choice set) and the sensitivity to different alternative attributes (the taste parameters). The disturbances reflect unobserved factors that influence mode choice, assumed to be i.i.d. Extreme Value across individuals and observations. As before, separate class-specific models are estimated for mandatory and non-mandatory tours, but for the sake of visual interpretability we do not show them as separate models in the figure.

### **3.2.3 Estimation Results**

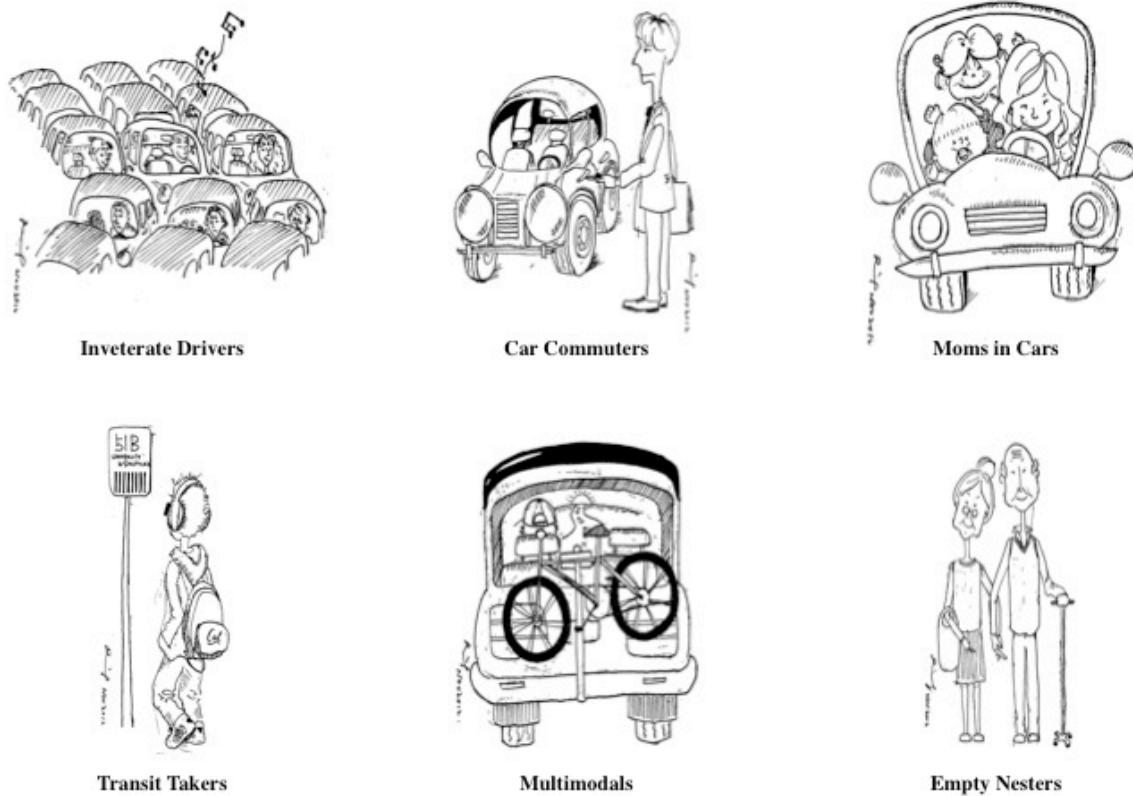
In determining a final model specification for the sample population, we estimated numerous models where we varied the utility specification, number of classes and choice set assumptions. Here we briefly summarize this process and present key results in Table 3.6 for fourteen different models. To facilitate comparison, Table 3.6 enumerates for each model the number of parameters estimated, the log-likelihood of the training data, the adjusted rho-bar-squared ( $\bar{\rho}^2$ ), the Bayesian Information Criterion (BIC), the Akaike Information Criterion (AIC) and the log-likelihood of the holdout data. Since the adjusted rho-bar-squared and the AIC are equivalent measures of fit, we will be restricting our attention to the BIC, the AIC and the log-likelihood of the holdout data when comparing different models. There are two key trends to notice: (1) each of the twelve LCCMs (with and without feedback) outperforms the multinomial logit model (with and without sociodemographic variables, where for the former the sociodemographic variables included in the LCCMs through the class membership model were included directly in the travel mode choice model as alternative-specific variables) on all three measures of goodness of fit; and (2) each of the six LCCMs with feedback through consumer surplus outperforms its counterpart without feedback on a majority of the three measures of goodness of fit, though a lower log-likelihood for the holdout data for two and three class models indicates a risk of overfitting.

**Table 3.6:** Summary statistics for different model specifications

<b>Model</b>	<b>Parameters</b>	<b>LL<sup>a</sup></b>	<b><math>\bar{p}^2</math></b>	<b>BIC</b>	<b>AIC</b>	<b>LL<sup>b</sup></b>
Multinomial Logit						
Without sociodemographics	20	-50,215	0.357	100,650	100,470	-5,613
With sociodemographics	260	-44,888	0.422	92,636	90,296	-4,982
Two Class LCCM						
Without feedback	70	-43,927	0.437	88,623	87,993	-4,909
With feedback	72	-43,802	0.439	88,396	87,758	-4,929
Three Class LCCM						
Without feedback	118	-41,803	0.464	84,903	83,841	-4,631
With feedback	121	-41,624	0.466	84,759	83,490	-4,634
Four Class LCCM						
Without feedback	144	-41,286	0.470	84,156	82,860	-4,575
With feedback	150	-41,220	0.471	84,090	82,740	-4,567
Five Class LCCM						
Without feedback	176	-40,668	0.477	83,272	81,688	-4,529
With feedback	182	-40,315	0.482	82,632	80,994	-4,468
Six Class LCCM						
Without feedback	196	-40,387	0.481	82,931	81,167	-4,478
With feedback	205	-40,194	0.483	82,643	80,798	-4,452
Seven Class LCCM						
Without feedback	250	-40,158	0.483	83,066	80,816	-4,458
With feedback	257	-40,103	0.484	83,033	80,720	-4,423

<sup>a</sup> Log-likelihood for training data<sup>b</sup> Log-likelihood for holdout data

Based both on statistical measures of fit and behavioral interpretation, we select the six-class LCCM with feedback as the preferred model specification. In terms of fit, the six-class LCCM with feedback has the second lowest BIC after the five-class LCCM with feedback, and the second lowest AIC and log-likelihood for the holdout data after the seven-class LCCM with feedback. In terms of the signs and relative magnitudes of the different model parameters and the accompanying behavioral interpretation of each of the latent classes, results for the six-class LCCM with feedback proved to be the most satisfying. Tables 3.7 and 3.8 list estimation results for the class-specific travel mode choice model for mandatory and non-mandatory tours, respectively, and Table 3.9 lists estimation results for the class membership model for the six-class LCCM with feedback (the utility of each alternative for the two class-specific choice models and the class membership model was specified linear in the variables listed in each table). Over the course of the following paragraphs, we describe in greater detail each of the six classes identified by the model (class labels are descriptive and not definitive). The descriptions are accompanied by illustrations, shown in Figure 3.4, that capture both the kinds of individuals that belong to each of the six classes and their latent preferences for different travel modes.



**Figure 3.4:** Illustrations showing the kinds of individuals that belong to each of the six modality styles and their latent preferences for different travel modes (illustrations by Rui Wang)

**1. Inveterate drivers:** Comprising 12.8% of the sample population, inveterate drivers do not consider travel modes other than drive alone or shared ride. They have a very low value of in-vehicle travel time of 0.55 \$/hr for mandatory tours and are insensitive to in-vehicle travel times for non-mandatory tours. They are sensitive to travel costs though. Household income, car ownership and house ownership each exert positive and significant influence on class membership. Individuals belonging to smaller households are more likely to be inveterate drivers.

**2. Car commuters:** Comprising 16.9% of the sample population, car commuters too do not consider travel modes other than drive alone and shared ride. They have a value of in-vehicle travel time of 6.95 \$/hr for mandatory tours and are insensitive to travel costs for non-mandatory tours, indicating a very high value of in-vehicle travel time for the same. Employment has a positive and significant influence on class membership. Men are more likely than women to be car commuters. Marriage and the presence of young kids at home positively influence class membership as well (76% of the segment is married and 63% of the segment comprises individuals from households with pre-school or school going kids). 86% of mandatory tours are drive alone, whereas 82% of non-mandatory tours are shared rides. Car commuters represent working members in a traditional household structure.

**Table 3.7:** Class-specific travel mode choice model for mandatory tours

<b>Variable</b>	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Class 5</b>	<b>Class 6</b>
<i>Alternative specific constants</i>						
Drive Alone	0.000	0.000	0.000	0.000	0.000	0.000
Shared Ride	-1.912	-2.738	1.459	-2.569	-1.194	-
Walk	-	-	1.332	2.300	1.626	-
Bike	-	-	-	-	-0.182*	-
Walk to Transit	-	-	1.533	4.125	1.078	3.528
Drive to Transit	-	-	0.542 *	0.476	1.053	0.450*
<i>Level-of-service variables</i>						
In-vehicle time (min)	-0.004 *	-0.011	-0.048	-0.010	-0.016	-0.084
Walking time (min)	-	-	-0.040	-0.002	-0.023	-0.084
Bike time (min)	-	-	-	-	-0.013	-
Waiting time (min)	-	-	-0.028	-0.008	-0.001*	-0.094
Cost (\$)	-0.434	-0.095	-0.167	-0.377	-0.026*	0.000

\* Insignificant at the 5% level

**Table 3.8:** Class-specific travel mode choice model for non-mandatory tours

<b>Variable</b>	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Class 5</b>	<b>Class 6</b>
<i>Alternative specific constants</i>						
Drive Alone	0.000	0.000	0.000	0.000	0.000	0.000
Shared Ride	-0.891	1.260	1.112	0.343	0.209	-1.890
Walk	-	-	1.826	3.667	-1.224	1.461
Bike	-	-	-	-	0.116	-
Walk to Transit	-	-	4.033	3.296	-2.152	-0.354
Drive to Transit	-	-	2.734	0.731	-2.683	-0.705
<i>Level-of-service variables</i>						
In-vehicle time (min)	0.000	-0.017	-0.012	-0.011	-0.005	-0.021
Walking time (min)	-	-	-0.030	-0.025	0.000	-0.041
Bike time (min)	-	-	-	-	-0.021	-
Waiting time (min)	-	-	-0.105	-0.016	-0.000*	-0.009*
Cost (\$)	-0.423	0.000	-0.987	-0.067	-0.013*	-0.118

\* Insignificant at the 5% level

**Table 3.9:** Class membership model

<b>Variable</b>	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>	<b>Class 5</b>	<b>Class 6</b>
Class specific constant	0.000	-2.480	-0.386*	0.587*	-0.552*	1.988
<i>Consumer surplus</i>						
Mandatory tours (utils)	0.117 *	0.343	0.024*	1.366	0.105*	0.503
Non-mandatory tours (utils)	0.275	0.642	0.062	0.761	0.163	0.098
<i>Household characteristics</i>						
Household income <sup>a</sup>	0.000	-0.099	-0.032*	-0.080	-0.112	-0.060
Household size	0.000	1.293	1.354	1.685	1.154	-0.012*
Internet access	0.000	-0.093*	-0.091*	-0.266*	-0.003*	-0.115*
Presence of pre-school kids	0.000	1.390	1.492	0.653*	1.026	-0.569*
Presence of school kids	0.000	-0.077*	-0.081*	-0.946	-0.830	-1.042
<i>Individual characteristics</i>						
Male	0.000	0.493	-0.327	0.251	0.736	0.254
Full time worker	0.000	0.692	-0.644	-1.135	-0.080*	-0.042*
Part time worker	0.000	0.230*	-0.524	-0.948	-0.300*	-0.003*
College student	0.000	-0.082*	-0.408	-0.097	0.085*	-0.061*
Disabled	0.000	-0.256*	0.480*	0.031*	0.064*	-0.917
Caucasian	0.000	0.157*	0.037*	0.018*	0.216*	0.320
Hispanic	0.000	-0.430*	-0.344*	0.167*	-0.510*	0.079*
African American	0.000	-0.448*	-0.317*	-0.183*	-0.781	-0.236*
Married	0.000	0.978	0.084*	0.035*	-0.069*	-0.525
Parent	0.000	-0.425*	0.152*	-0.899	-0.591	0.167*
Ages 12 and under	0.000	1.241	0.901*	-0.760*	-0.455*	-2.574
Ages 13-17	0.000	-1.911*	1.963	1.241*	1.036*	-0.958
Ages 18-24	0.000	-0.203*	0.628	-0.134*	-0.398*	-0.625
Ages 25-44	0.000	0.832	1.084	0.691	0.665	0.326
Ages 65 and over	0.000	-0.756	0.620	-0.625	-0.423	-0.350*
<i>Long-term travel and activity decisions</i>						
Number of household cars	0.000	-0.667	-1.511	-2.024	-1.213	-0.226
Number of household bicycles	0.000	0.074*	0.083*	0.052*	0.324	0.021
Single family detached house	0.000	0.208*	0.040*	0.036*	-0.299*	0.023*
Probability that house is owned	0.000	-0.466	-0.364	-0.727	-0.313*	-0.077*

\* Insignificant at the 5% level

<sup>a</sup> Household income is a categorical variable between 1 and 15, where 1 represents households whose annual income is less than \$10,000 and 15 represents households whose annual income is more than \$150,000. The income range between successive categories increases from \$5,000 for the first 9 categories to \$25,000 for the last 3 categories. By incorporating household income as a continuous variable in the class membership model, it is assumed that the marginal utility of any class decreases with increasing income.

**3. Moms in cars:** Comprising 31.1% of the sample population, moms in cars consider all travel modes except bicycling. However, a high value of time inherently biases their observed choices towards the car. 94% of all tours, mandatory and non-mandatory, are made by car, of which 92% are shared rides. For mandatory tours, moms in cars have a value of in-vehicle travel time of 17.25 \$/hr, a value of walking time of 14.37 \$/hr and a value of waiting time of 10.06 \$/hr. For non-mandatory tours, moms in cars have a value of in-vehicle travel time of 0.73 \$/hr, a value of walking time of 1.82 \$/hr and a value of waiting time of 6.38 \$/hr, indicating a greater sensitivity to the last mile. The presence of pre-school kids has a positive and significant effect on class membership (77% of the segment comprises individuals from households with pre-school or school going kids), whereas employment (full-time or part-time) has a negative and significant effect on class membership. Adult women and children ages 12 and under are most likely to belong to this class. Moms in cars represent stay-at-home mothers and young children in a traditional household structure.

**4. Transit takers:** Comprising a small 7.4% of the sample population, transit takers consider all travel modes except bicycling when deciding how to travel but harbor a positive predisposition towards transit: 43% of mandatory tours and 19% of non-mandatory tours are made by transit by the segment. Transit takers have a value of in-vehicle travel time of 1.59 \$/hr for mandatory tours and 9.85 \$/hr for non-mandatory tours. The last mile is far more important for non-mandatory tours than it is for mandatory tours. Unemployed individuals in their teens and twenties belonging to large low-income households with low car ownership rates are most likely to belong to this segment.

**5. Multimodals:** At 9.1% of the sample population, multimodals comprise the only class uncovered by the model that considers bicycling when deciding how to travel. 18% of mandatory tours and 13% of non-mandatory tours are made by bike by the segment. Men are more likely to be multimodal than women, and African Americans are less likely to be multimodal than other races and ethnicities. Multimodals have a high value of in-vehicle travel time of 36.92 \$/hr for mandatory tours and 23.08 \$/hr for non-mandatory tours.

**6. Empty nesters:** Comprising 22.7% of the sample population, empty nesters consider drive alone, walk to transit and drive to transit for mandatory tours, and all modes except bicycling for non-mandatory tours. For mandatory tours, they are insensitive to travel costs and more or less equally sensitive to in-vehicle travel times, walking times and waiting times, indicating a very high value of time for the same. For non-mandatory tours, they have a value of in-vehicle travel time of 10.68 \$/hr, a value of walking time of 20.85 \$/hr and a value of waiting time of 4.57 \$/hr. Empty nesters have the highest average age of the six classes at 49 years. As indicated by the label, small households with no children have a positive and significant influence on class membership. Less than 6% of the segment is composed of individuals from households with pre-school or school going kids and 70% of the segment is aged 40 and over.

### 3.2.4 Conclusions

Contrary to our prior expectations, the model framework developed in Chapter 2 appears to work just as well with cross-sectional datasets. In fact, in many ways, due largely to the greater number of individuals in the sample population, estimation results for the BATS 2000 dataset are

far richer than those for the MOBIDRIVE dataset. That the framework worked equally well for both datasets, despite distinct differences in sample size, observation period, the local transportation system and cultural context, attests to its robustness. LCCMs with feedback represent an important first step in the development of discrete choice models that allow for both preference heterogeneity and preference endogeneity. Future research should look to expand upon the framework developed in Chapter 2 to include the influence of other external stimuli, such as social norms and past experiences, on preference formation and change.

Findings from the BATS 2000 dataset reveal six broad modality styles that differ substantially from one another in terms of both the kinds of individuals that belong to each group and the relative importance that they attach to different level-of-service attributes of the transportation system. For example, 30 percent of the sample population belonging to two of the six modality styles only considers the automobile when deciding how to travel. Even worse, only nine percent of the sampled population belonging to one of the six modality styles even thinks about bicycling as a mode of transportation when deciding how to travel. Value of time was found to vary across the six modality styles from as little as 0.5 \$/hr (indicating near insensitivity to travel times) to as high as 37 \$/hr. How do these findings compare with those from Karlsruhe, Germany? The proportion of the sample population that displays a strong predisposition towards the automobile is much smaller at 16%, and even these auto-oriented individuals are willing to walk for short-distance tours. The contrast grows starker in terms of public transit and bicycling: 84% of the sample population belonging to the remaining two modality styles actively consider the two modes as alternatives when deciding how to travel. In terms of usage, these two modality styles use public transit and bicycling for 14% and 20% of their work tours, respectively, and 13% and 14% of their non-work tours, respectively.

To put things in perspective, the San Francisco Bay Area has one of the more extensive public transit networks of metropolitan regions in the United States. It is a poly-nucleated metropolitan region with central business districts in San Francisco, Oakland and San Jose. San Francisco city itself is compact and walkable, and is connected both to downtown Oakland and major bedroom communities in the Bay Area such as Fremont and Pleasanton by the Bay Area Rapid Transit (BART), a heavy-rail commuter service. San Jose, on the other hand, famously excluded itself from the BART project in 1957, opting instead like much else of the country to build expressways. However, the city is connected to major centers along the San Francisco peninsula, including San Francisco city, by Caltrain which operates commuter trains at headways of 20 minutes during peak hours, and to towns that lie to the east of the San Francisco Bay by bus service. Even exurban towns like Tracy, 80% of whose workforce is employed in the Bay Area, is connected to the BART system and to job centers in the South Bay by Amtrak buses and commuter lines operated by the San Joaquin Regional Transit District. And yet, nearly a third of the sample population uncovered within the BATS 2000 dataset does not even consider public transit when deciding how to travel. And the remaining two-thirds that do consider public transit only actually use it for 13% of their mandatory tours and 4% of their non-mandatory tours. As we demonstrate in the following chapter, these findings hold important implications for policies meaning to change existing patterns of travel mode choice behavior.

## **Chapter 4**

# **Policy Implications**

Forecasts from travel demand models are regularly employed by Metropolitan Planning Organizations to determine the required capacity that new infrastructure must satisfy, and to facilitate the economic, environmental and social impact assessments that usually accompany the debate on how to allocate funds between competing initiatives. A greater comprehension of the many factors that shape behavior is essential to the successful design of systems that serve the immediate needs of the population while satisfying long-term societal objectives. In general, transport policies and infrastructural initiatives seeking to change travel behavior will elicit different responses from individuals with differing modality styles in the short-term and force changes in the distribution of modality styles across the population in the long-term. Traditional travel mode choice model forms such as the multinomial logit or mixed logit model assume that all individuals are multimodal, i.e. they are aware of the full range of alternatives at their disposal and they make a considered decision regarding their mode of travel each time they step out of the house. LCCMs without feedback assume that different individual modality styles exist, that they are a function of individual and household characteristics and long-term travel and activity decisions but are immune to changes to the transportation system. LCCMs with feedback through consumer surplus allow the distribution of individual modality styles to be subject to the influence of changes to the transportation system. As a consequence, traditional travel mode choice models are insensitive to both short and long-term effects, LCCMs without feedback are sensitive to short-term effects but insensitive to long-term effects, and LCCMs with feedback through consumer surplus are sensitive to both short and long-term effects.

The objectives of this chapter are two-fold: (1) to show how travel mode choice models that overlook preference heterogeneity and/or preference endogeneity can result in systematic biases in forecasts; and (2) to demonstrate that improvements to the transportation network must necessarily be attended by commensurate changes in individual modality styles for the proposed improvements to have a substantive effect on existing travel patterns. We compare forecasts for two different scenarios using the BATS 2000 dataset. Sections 4.1 and 4.2 compare the impact of increased auto congestion and improvements to the public transit system on travel mode choice behavior, respectively, as predicted by a multinomial logit model, the six-class LCCM without feedback and the six-class LCCM with feedback presented in Chapter 3. Section 4.3 concludes the chapter with a comparison of our findings with those in the literature.

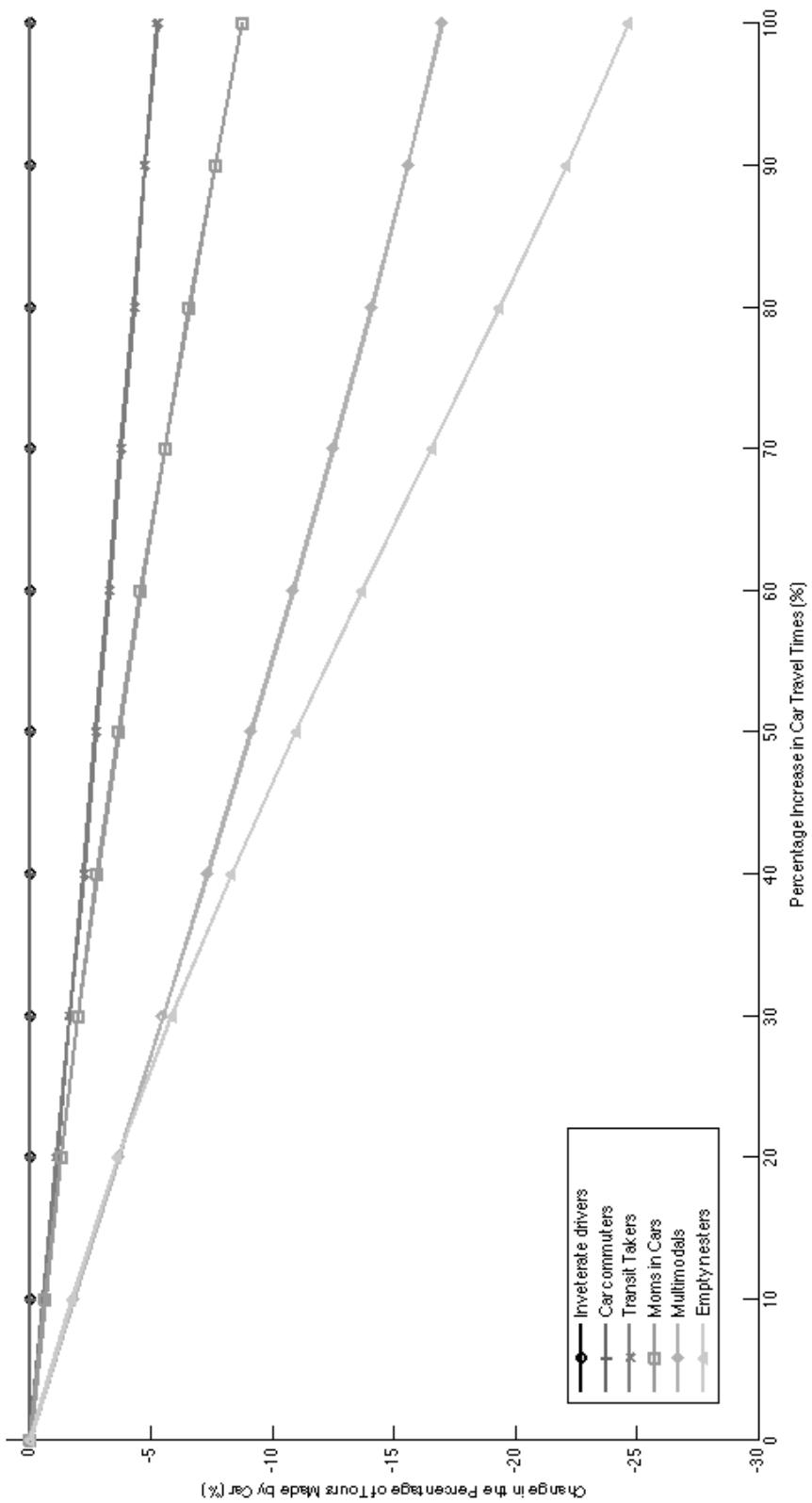
### **4.1 Scenario I: Increased Auto Congestion**

The first scenario increases in-vehicle travel times for drive alone and shared ride, without changing any of the other level-of-service attributes for these and the four remaining travel modes. The scenario is clearly not realistic: increased congestion will have spillover effects on the performance of other travel modes that share road space with cars. However, our purpose

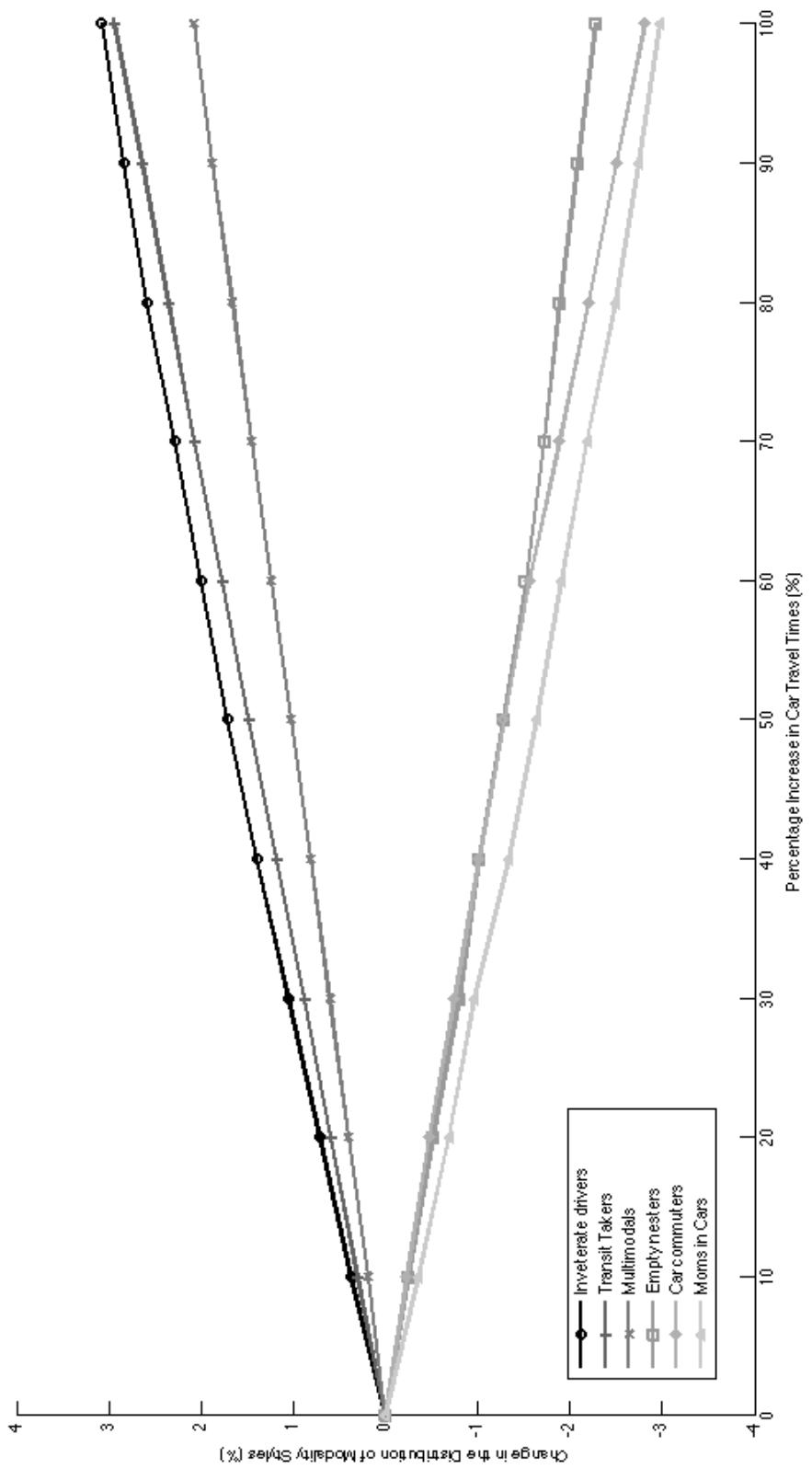
here is not to offer definite predictions but merely to illustrate the comparative benefits that can be expected from the model framework developed in Chapter 2 over simpler model forms.

Figure 4.1 shows how different modality styles respond differently in the short-term to a change to the transportation system through a plot of the aggregate change in the percentage of mandatory and non-mandatory tours made by individuals belonging to different modality styles within the BATS 2000 dataset by car (drive alone and shared ride) as a function of increasing in-vehicle travel times for car, as predicted by the six-class LCCM with feedback through consumer surplus presented in the Chapter 3. Since inveterate drivers and car commuters only consider drive alone and shared ride when deciding how to travel, the percentage of tours made by car by the two segments is insensitive to increases in in-vehicle travel times and the change in the percentage of tours made by car is zero throughout. Empty nesters and multimodals have the highest values of in-vehicle travel time and consequently show the greatest change in travel mode choice. For moms in cars, the change in the percentage of tours made by car isn't as dramatic as the other two classes because the increase in in-vehicle travel times for drive alone and shared ride is partially offset by the high disutility of walking and waiting times associated with other travel modes, particularly for non-mandatory tours. Transit takers show a smaller change compared to these three classes in part due to their lower value of in-vehicle travel time and in part due to a lower baseline percentage of tours made by car (38% of existing tours made by transit takers within the dataset are by car, compared to 58% for multimodals, 90% for empty nesters and 93% for moms in cars). A key takeaway from Figure 4.1 is the wide disparity in responses to increasing car in-vehicle travel times across the six modality styles that would otherwise be missed by traditional models of travel mode choice that assume all individuals to be multimodal.

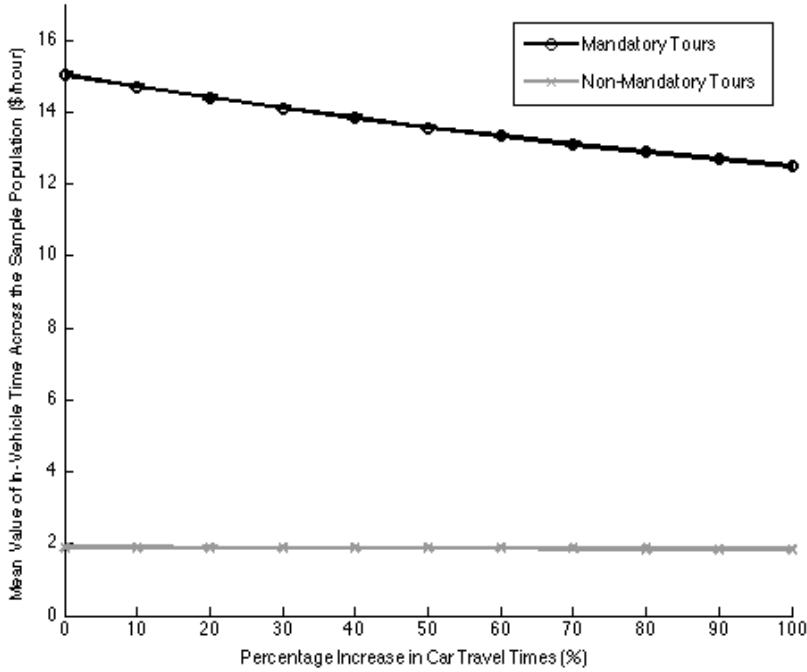
Figure 4.2 shows how changes in the transportation system can force changes in the distribution of modality styles across the population in the long-term through a plot of the aggregate change in the distribution of the modality styles across the BATS 2000 dataset as a function of increasing in-vehicle travel times for car, as predicted by the same six-class LCCM with feedback through consumer surplus as before. As in-vehicle travel times for drive alone and shared ride increase, individuals move away from car commuters, moms in cars and empty nesters and towards inveterate drivers, transit takers and multimodals. The migration pattern suggests two distinct coping mechanisms at play: (1) as the car becomes progressively less attractive as a form of transportation, some individuals react by expanding their choice set to include other travel modes, explaining the increase in the percentage of transit takers and multimodals; and (2) individuals unwilling to change lifestyles built around the use of the car respond to the increase in in-vehicle travel times by lowering their value of time and continuing to drive as before, explaining the increase in the percentage of inveterate drivers. The relatively small changes in the distribution of modality styles across the sample population further suggests that socioeconomic factors are more powerful predictors of class membership than the level of service offered by different travel modes. Figure 4.3 plots the mean value of in-vehicle travel time across the sample population as a function of the change in car travel times. For mandatory tours, the value of in-vehicle travel time decreases from 15.0 \$/hr for the baseline case to 12.5 \$/hr when car travel times have doubled. For non-mandatory tours, since waiting and walking times are stronger determinants of travel mode choice across most modality styles than in-vehicle travel times, the change in the value of in-vehicle travel time is marginal at best, from 1.9



**Figure 4.1:** Change in the percentage of tours, mandatory and non-mandatory, made by car (drive alone and shared ride) across different modality styles as a function of increasing car travel times, as predicted by the six-class LCCM with feedback



**Figure 4.2:** Change in the distribution of modality styles across the sample population as a function of increasing car travel times, as predicted by the six-class LCCM with feedback



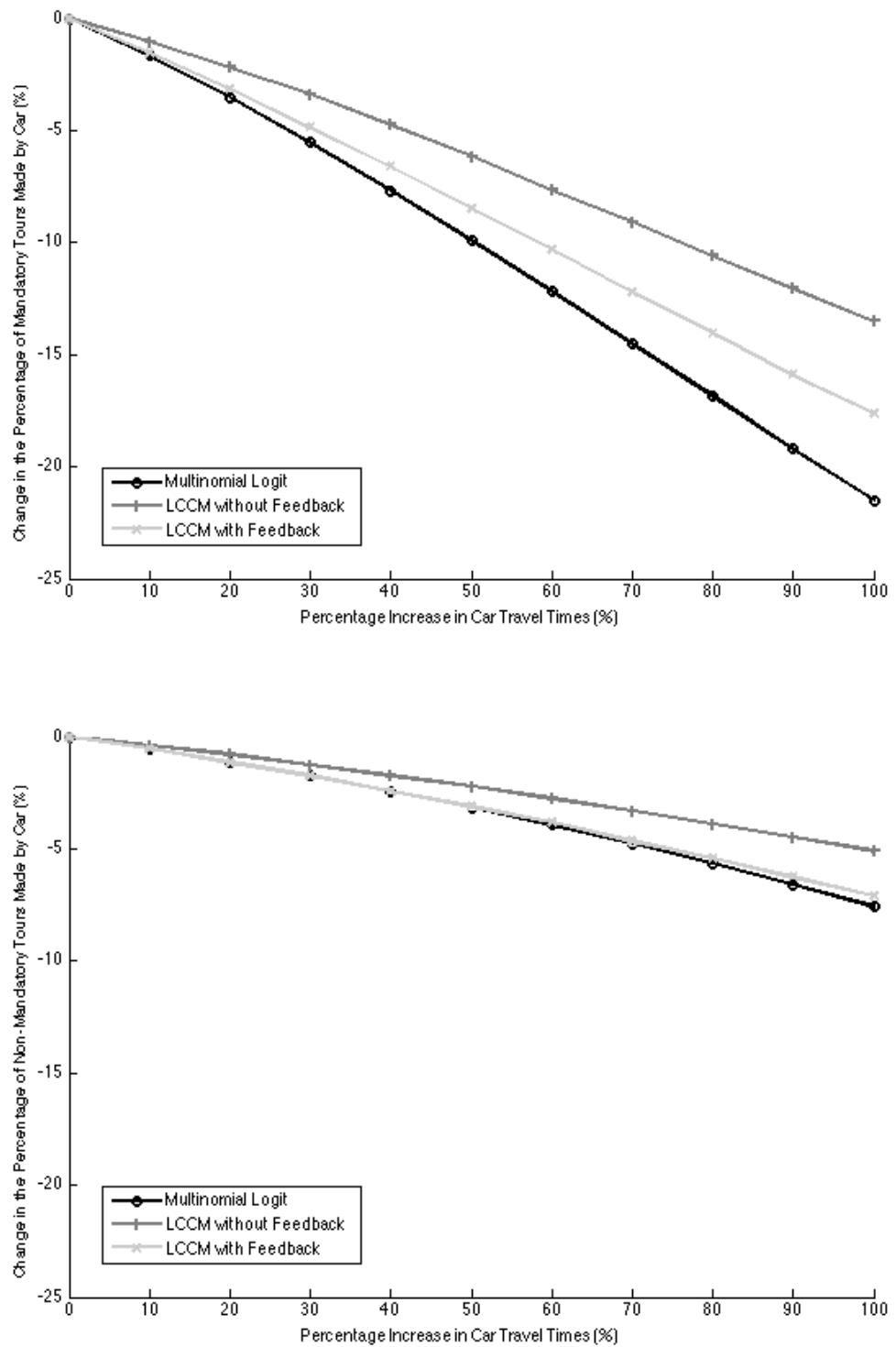
**Figure 4.3:** The mean value of in-vehicle travel time across the sample population for mandatory tours and non-mandatory tours as a function of increasing car travel times, as predicted by the six-class LCCM with feedback through consumer surplus

\$/hr for the baseline case to 1.8 \$/hr when car travel times have doubled. Finally, it merits mentioning here that an LCCM without feedback through consumer surplus would be oblivious to the influence of either coping mechanism on existing travel patterns.

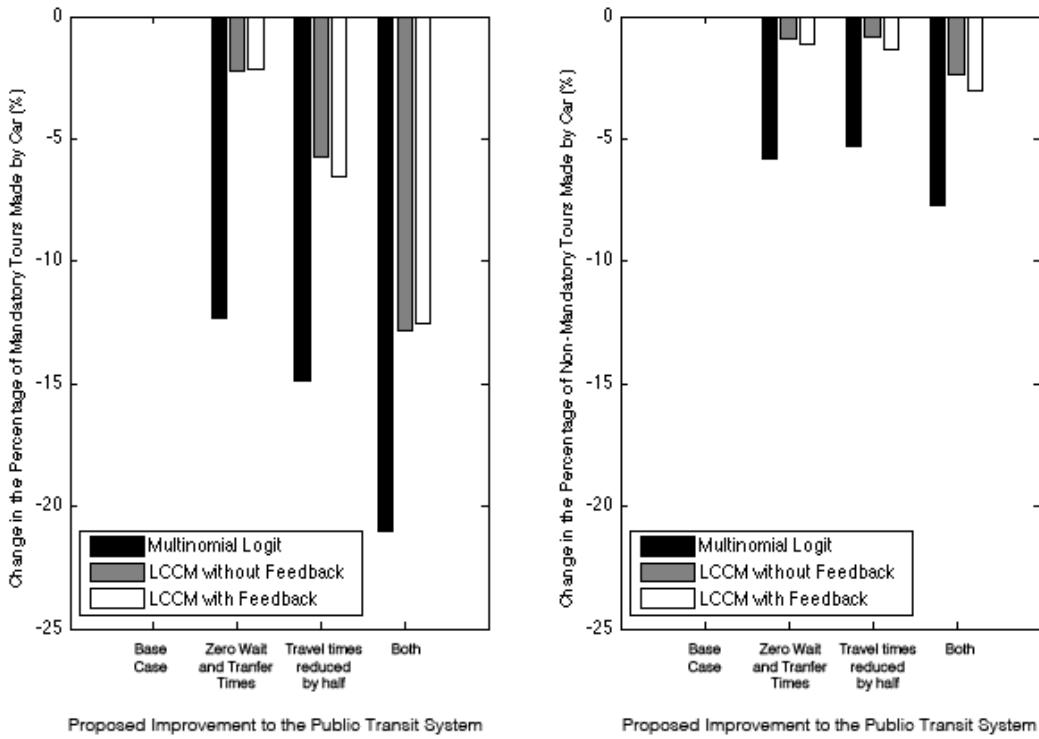
What do these findings imply for aggregate forecasts? Figure 4.4 plots the aggregate change in the percentage of mandatory and non-mandatory tours made by all individuals within the BATS 2000 dataset by car (drive alone and shared ride) as a function of increasing in-vehicle travel times for car, as predicted by the multinomial logit model with sociodemographic variables, the six-class LCCM without feedback and the six-class LCCM with feedback through consumer surplus. For mandatory tours, the changes predicted by the three models are starkly divergent. For non-mandatory tours, the multinomial logit model and the LCCM with feedback through consumer surplus are in greater agreement and the LCCM without feedback is somewhat of an outlier. Relative to the LCCM with feedback, for both mandatory and non-mandatory tours the multinomial logit model overestimates the change in the percentage of tours made by car and the LCCM without feedback underestimates the change in the percentage of tours made by car.

## 4.2 Scenario II: Improvements to the Public Transit System

The second scenario evaluates how three rather dramatic improvements to the public transit system might impact travel mode choice behavior. First, we eliminate waiting and transfer times



**Figure 4.4:** Change in the percentage of mandatory tours (top) and non-mandatory tours (bottom) made by car (drive alone and shared ride) as a function of increasing travel times, as predicted by the multinomial logit model with sociodemographic variables, the six-class latent class choice model without feedback, and the six-class latent class choice model with feedback through consumer surplus

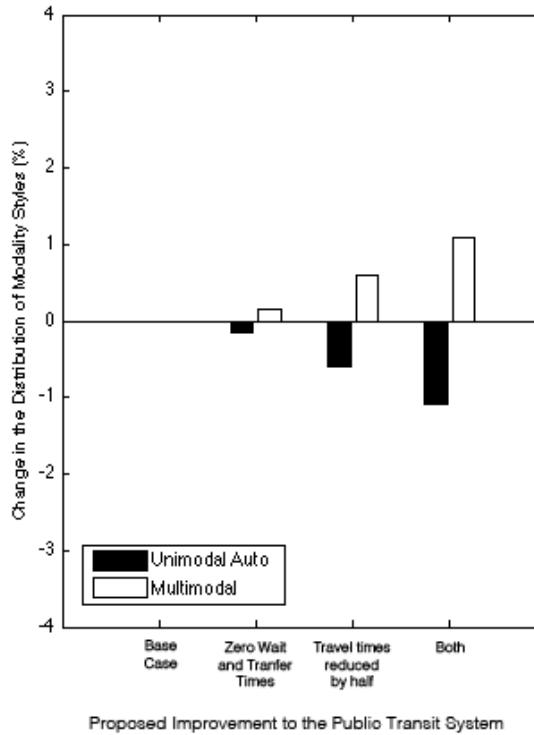


**Figure 4.5:** Change in the percentage of mandatory tours (left) and non-mandatory tours (right) made by car (drive alone and shared ride) as a function of proposed improvement to the public transit system, as predicted by the multinomial logit model with sociodemographic variables, the six-class latent class choice model without feedback, and the six-class latent class choice model with feedback through consumer surplus

for public transit. Next, we reduce in-vehicle travel times for public transit by half. And finally, we both eliminate waiting and transfer times and reduce in-vehicle travel times by half for public transit. In all three cases, the level-of-service for other travel modes is left unchanged.

Figure 4.5 plots the change in the percentage of mandatory and non-mandatory tours made by car (drive alone and shared ride) in response to each of the three proposed improvements to the public transit system, as predicted by the multinomial logit model with sociodemographic variables, the six-class LCCM without feedback and the six-class LCCM with feedback. For the sake of clarity, we do not plot the change in mode shares for any of the other four travel modes. The shift is almost entirely from car towards public transit, and the mode shares for bike and walk remain more or less the same across both mandatory and non-mandatory tours, the three sub-scenarios and the three model specifications. For both mandatory and non-mandatory tours, the change in car mode shares predicted by the LCCM without feedback closely mirrors the change predicted by the LCCM with feedback, and in this case it is the multinomial logit model that is the outlier.

Reasons for this are made clear by Figure 4.6, which plots the change in the distribution of individuals across modality styles in response to each of the three proposed improvements to the

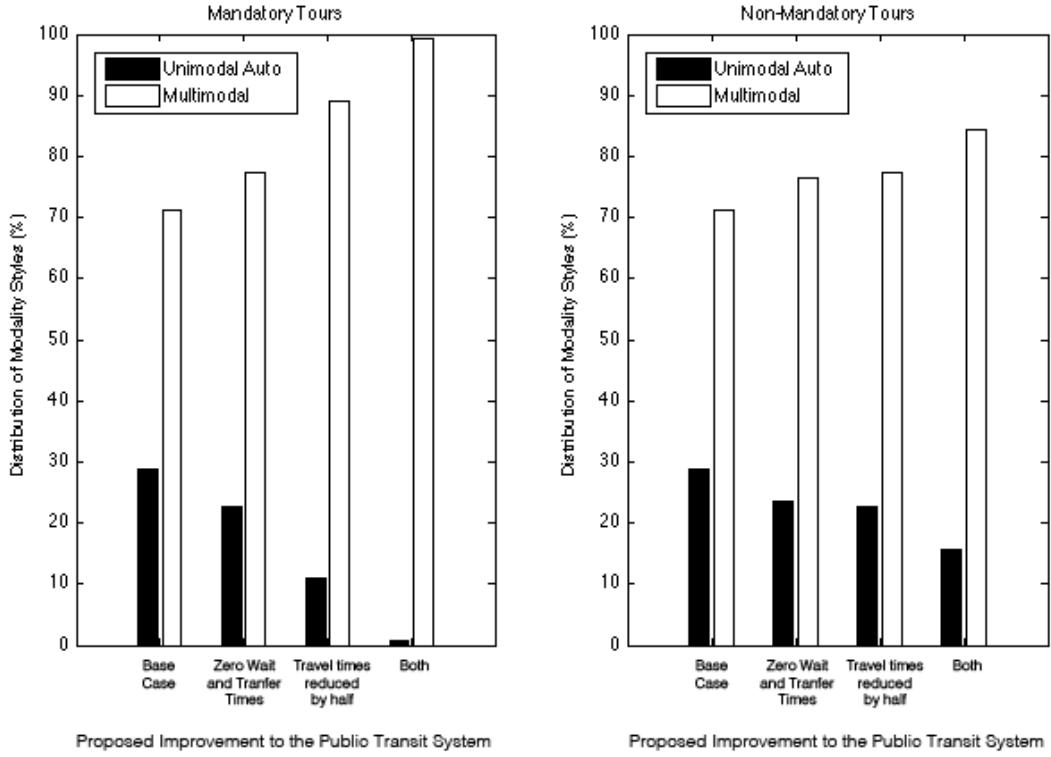


**Figure 4.6:** Change in the distribution of individuals across modality styles as a function of proposed improvement to the public transit system, as predicted by the six-class latent class choice model with feedback through consumer surplus

public transit system, as predicted by the six-class LCCM with feedback. To ease interpretation, we've grouped the two auto-oriented modality styles – inveterate driver and car commuters – under the label “unimodal auto” and the remaining four modality styles that consider other travel modes under “multimodal”. As is apparent, there is almost no change in the distribution of individuals, and as a consequence, predicted mode shifts from the six-class LCCM without feedback and the six-class LCCM with feedback are nearly identical.

What these results are further saying is that incremental changes to the level-of-service of alternative modes, unless attended by corresponding shifts in the distribution of modality styles, will have very little effect on existing travel mode shares. The three scenarios were chosen deliberately to illustrate the point that you can altogether eliminate waiting and transfer times and cut travel times by half for public transit, and still 73% of mandatory tours and 87% of non-mandatory tours continue to be made by the car.

Figure 4.7 shows how an equivalent shift in mode shares away from the automobile could be achieved in place of each of the proposed improvements to the public transit system by a change in the proportion of multimodals in the sample population. For example, if 77% of the sample population were multimodal, as opposed to the 71% that already is, the expected decrease in mode share for auto for mandatory tours would be the same as that achieved by eliminating waiting and transfer times for public transit. Similarly, if almost everybody in the sample



**Figure 4.7:** The required distribution of individuals across modality styles that would produce an equivalent shift in travel mode shares away from the automobile as each of the three proposed public transit improvements listed along the horizontal axis

population were multimodal, the expected decrease in mode share for auto for mandatory tours would be the same as that attained by both eliminating waiting and transfer times and reducing in-vehicle travel times by half for public transit. The purpose of the graph is to demonstrate that in some cases, it might be in society's best interest to change existing travel mode choice patterns through improvements to the transportation system, but in others it might be more feasible to persuade individuals to be more multimodal in order to meet the same objectives. The question then is: how?

Incremental changes in the level-of-service of alternative modes aimed at inducing a shift in travel modes often come unstuck in the face of firmly rooted daily patterns that revolve around the use of the automobile. Human beings are creatures of habit. When an action has been repeated frequently in stable contexts in the past, only minimal, sporadic thought is required to initiate, implement, and terminate it (Wood et al., 2002). Any attempt to influence choices will fail if the choices are non-deliberate (Gärling and Axhausen, 2003). For example, an increase in bus frequencies or the introduction of bike lanes is of little to no consequence to individuals who drive because they have always driven; such individuals will continue to drive even when new information has changed the contextual environment in which the original decision to drive might have been made (Aarts et al., 1997; Axhausen et al., 2001; Simma and Axhausen, 2003; Thøgersen, 2005). It is ironic then that what first attracts many individuals to the automobile are

the ideas of free will and self-determination, but the behavior itself is sustained over time by automatic, unconscious mental processes (Bargh and Chatrand, 1999).

So how else do we get individuals to leave the car at home? Changes in lifestyles and modality styles as characterized by corresponding changes in individual values, attitudes, and behavioral orientations will take time (Kitamura, 2009). However, more immediate changes can indeed be forced by exogenous influences, such as the effects of past experiences, altered personal circumstances and changes to the transportation system (NCTR, 2008; Verplanken et al., 2008). One bad bus ride can potentially put an individual off public transit forever. Major life events such as the birth of a child can trigger commensurate changes in lifestyles, and consequently modality styles. From a policy standpoint what is needed is a jolt to the system, an irremediable change in the transportation network that forces individuals to reconsider how they travel.

If the goal is to persuade individuals to drive less than classical economic theory mandates that driving be priced accordingly, thereby internalizing any externalities associated with traffic congestion and pollution, and attaining a socially optimum level of automobile use. For instance, the London Congestion Charge resulted in a 33% decrease in the number of automobiles entering or leaving the congestion zone during charging hours, and a corresponding increase of 29,000 (or 38%) in bus patronage within the Central London area (Transport for London, 2004). Though the success of the congestion charge demonstrates that major changes in lifestyles and modality styles can indeed be brought about by appropriately designed pricing schemes, the failed experiment to levy the same in New York City shows that political opposition to pricing schemes can often prove insurmountable. The chorus of cries within the academic community beseeching an increase in the gas tax in the United States has grown louder with each passing decade. Studies have variously pegged the optimal level at anywhere between \$1.01 per gallon (Parry and Small, 2005) to \$0.34 per vehicle mile travelled (Levinson and Gillen, 1998), which adjusting for inflation and assuming an average mileage of 23.8 miles per gallon (BTS, 2012) is equivalent to \$10.71 per gallon. The unfortunate political reality is that the gas tax in the United States continues to languish at a state average of 49 cents per gallon.

From the perspective of psychology and behavioral theory, an alternative way to coax individuals to consider alternative modes of travel might be through the use of incentives that promote societally efficient behavior (Smith et al., 2003). Travel demand management schemes have employed incentives in the past to reward commuters for changing travel modes (Meyer, 1997) or avoiding the rush hour (Ben-Elia and Ettema, 2009; Merugu et al., 2009), and to persuade individuals to walk more (Gomes et al., 2012). It has been argued that rewards are more likely to foster learning and internalization of the socially desirable behavior absent the unpleasant memories and issues of avoidance that result from similarly intentioned punishment schemes (Rescorla, 1987). And of course, the use of incentives is a far easier political sell.

Irrespective of whichever approach or combination of approaches is adopted, success or failure will ultimately hinge upon whether the planned policy or infrastructural initiative can force a change in the distribution of modality styles in the population of interest. As demonstrated by Figure 4.7, changes in the distribution of individuals across modality styles can prompt equivalent gains in travel mode shifts. But if existing modality styles persist, then even the most ambitious of initiatives will accrue modest benefits at best.

### **4.3 Conclusions**

The automobile's continued preeminence in much of the developed world, and its more recent proliferation in many developing countries, is a source of grave concern to the health of our cities and the global environment at large. The growing social costs imposed by the automobile through its impacts on congestion and safety, and the increased relevance of issues of equity and livability, have together contributed to a renewed interest in the United States in alternative modes of travel, such as public transit and bicycling, and their potential to offer a more sustainable solution to our mobility requirements. However, policies meaning to effect a change in travel behavior often come unstuck against long ingrained lifestyles and deeply entrenched habits built around the use of the automobile.

A comparison between travel demand forecasts from LCCMs with feedback and other model forms that do not account for preference heterogeneity and/or preference endogeneity found that the latter can bias forecasts by factors of between one-and-a-half and three. How do these numbers compare with findings elsewhere in the literature? A retrospective study by Parthasarathi and Levinson (2010) comparing the accuracy of traffic forecasts for 108 recently completed roadway projects in Minnesota discovered a general trend of underestimation, with 65% of the critical links showing underestimated traffic forecasts. Similar retrospective studies by Pickrell (1992) and Flyvbjerg et al. (2007) comparing the accuracy of forecasts regarding rail investment with actual observed market shares uncovered systematic biases as well. In seven of the eight cases investigated by Pickrell, the actual demand was less than half of the forecasted demand. Similarly, nine of the ten projects surveyed by Flyvbjerg et al. overestimated ridership by an average of 106%. We argue that the gap between predicted and observed travel mode shares is at least fractionally attributable to the fact that most models either overlook the influence of lifestyles and modality styles on travel mode choice or capture it in a manner that is behaviorally incomplete, and this gap could partially be bridged by adopting the methodological framework presented in this dissertation.

# **Chapter 5**

## **Dynamics of Modality Styles**

The model framework developed in Chapter 2 was used to examine the relationship between individual modality styles and travel mode choice behavior in a static context. However, a static framework is guilty of making two assumptions: (1) that an individual's preferences in the present are independent of her preferences in the past and, by extension, her preferences in the future; and (2) that decisions are made in a bubble without any consideration for past experiences or future expectations. The objective of this chapter is to translate the framework developed in Chapter 2 to a dynamic context. Section 5.1 reviews past literature on dynamic models of decision-making that have sought to relax these assumptions. Section 5.2 develops and estimates a dynamic model of travel mode choice behavior that relaxes the first of these two assumptions.

### **5.1 Literature Review**

Dynamic discrete choice models have typically accounted for the influence of past experiences on present choices by specifying utility in the present time periods as some function of observed variables from past time periods. For example, Cantillo et al. (2007) and Yáñez et al. (2009) include in the utility specification in the present time period for any alternative the difference in utilities in the preceding time period between that alternative and the alternative that was chosen in the preceding time period. But if past choices affect present choices, then present choices must necessarily be understood to affect future choices. For example, the decision to purchase a car holds long-term implications for both the amount of resources that an individual can expend on future activities and the amount of resources required to engage in each of these activities. Whether the individual chooses to buy a car or not will depend upon, among other things, how she foresees using the car over her period of ownership. Most random utility maximization models of decision-making that account for the influence of future expectations on current choices assume that individuals are aware that their current choices affect the alternatives at their disposal in the future and know that they will maximize utility among these alternatives in the future just as they maximize utility among the alternatives at their disposal currently. As a consequence, individuals are hypothesized to choose that alternative in the current time period that maximizes their expected utility over current and future time periods. The framework was first proposed by Rust (1987). For a comprehensive review of more recent developments in the field, the reader is referred to Train (2009).

The approach adopted in this chapter does not concern itself with the dynamics underlying choices as much as it does with the dynamics underlying preferences, as represented by modality styles. An individual's modality style in the current time period is hypothesized to be some function of the choices that she made in previous time periods, but conditioned on the individual's modality style in the current time period her choices in that time period are assumed

to be independent of both past experiences and future expectations. In the context of travel mode choice behavior, the assumption is not entirely unreasonable. For example, an individual's decision to take the bus on a particular day should have no bearing on the travel modes that she can take on subsequent days. However, one bad bus ride could force a change in the individual's preferences such that she never gets on a bus again. Previous studies on travel behavior by Ben-Akiva and Abou Zeid (2007) and Choudhury et al. (2010) have used Hidden Markov Models (HMM) to model the evolution of latent variables and observed choices over time but have overlooked the influence of external conditions on the same. Our work builds on these studies through the inclusion of changes in external conditions on modality styles and, consequently, travel mode choice behavior.

## 5.2 Proof of Concept

In this section we use travel diary data from a multi-wave panel survey to develop a preliminary framework within which to study modality styles in a dynamic context. Section 5.2.1 describes the dataset; Section 5.2.2 introduces the proposed preliminary methodological framework; and Section 5.2.3 concludes the section with a discussion of the estimation results. The work presented in this section is intended to serve as a proof-of-concept that future studies can build upon in moving towards our stated objective of fully operationalizing the modality styles construct within a dynamic context. We discuss some of the limitations of the approach presented here and how they might be overcome at the end of Section 5.2.3.

### 5.2.1 The Dataset

In February of 2007, Santiago, Chile introduced Transantiago, a complete redesign of the public transit system in the city. Before the introduction of Transantiago, public transport in Santiago comprised a privately operated and uncoordinated system of buses and shared taxis, and the publicly run underground Metro lines. The old bus system was characterized by a large and inefficient fleet of 8,000 buses operating 380 lines, competition among buses on streets to gain passengers, higher than required frequencies along the busiest corridors and inadequate service along the less travelled ones, low quality vehicles, high accident rates, rude drivers, high levels of air and noise pollution, fractured ownership, and many empty buses circulating during off peak hours (Yáñez et al., 2010). The Metro system, though considerably safer, faster and more reliable than the bus system, only accounted for 8% of the city's trips under the old system, due largely to sparser network coverage and the high cost of transfers between buses and the Metro.

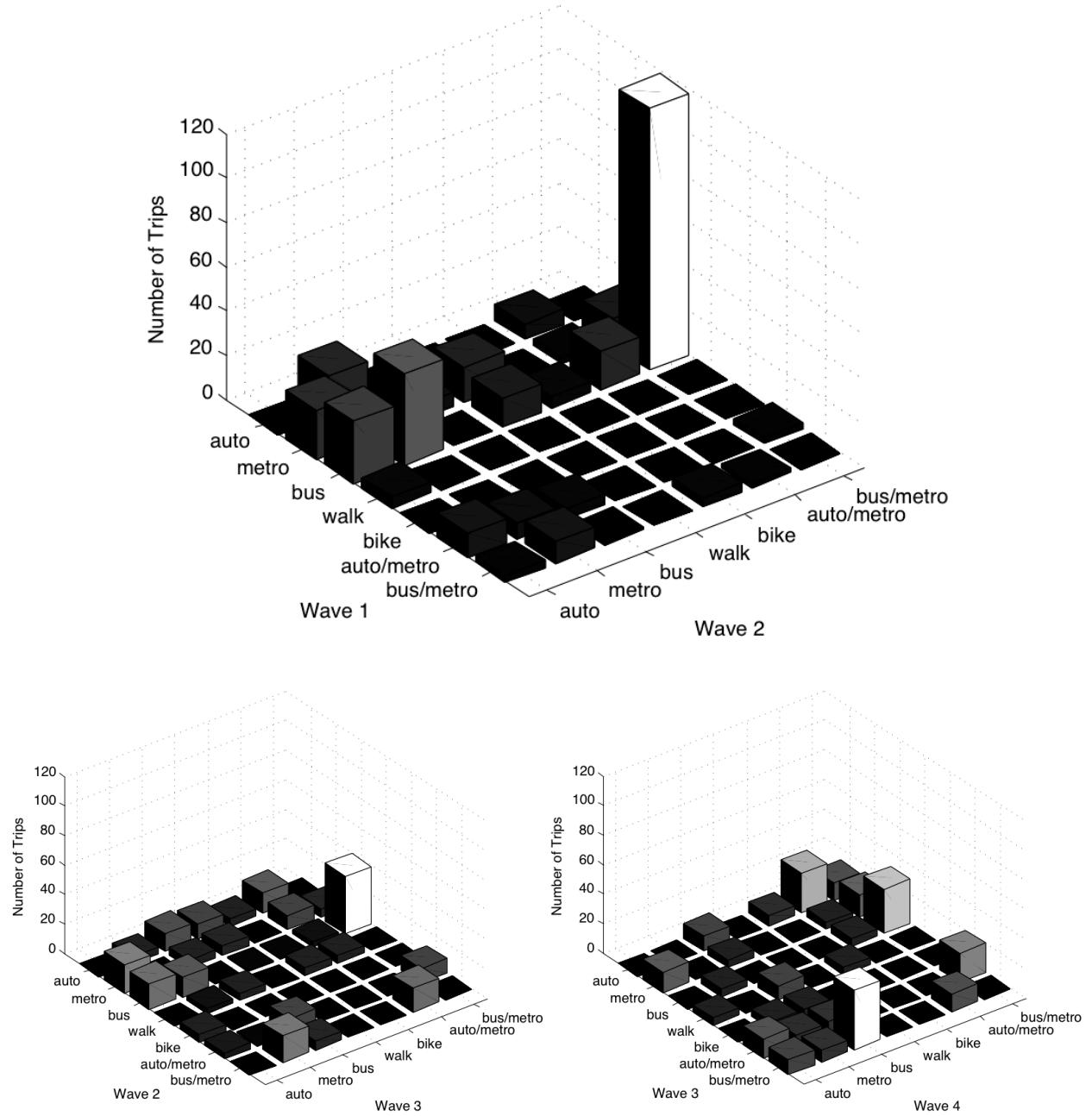
With the aim of addressing these problems and stemming the decline in the public transportation system, the city assembled a team of Chilean specialists and consultants in 2005 to come up with a design for Transantiago (Fernández et al., 2008). Under the new system, the metropolitan region in and around Santiago was divided into ten zones and operations were taken over by a group of ten new companies. Bus routes were consolidated into a hierarchical system of trunk and feeder routes. The feeder routes connected each of these zones to the Metro lines, which served as the backbone of the new system. The trunk routes complemented the Metro lines by connecting different zones of the city. Benefits envisaged under Transantiago included the elimination of route redundancies, increased safety through the introduction of new low-floor

buses, approximately half of them articulated, an integrated fare collection system through the means of a contactless smart card, lower travel times, a smaller fleet size, and reduced levels of air and noise pollution.

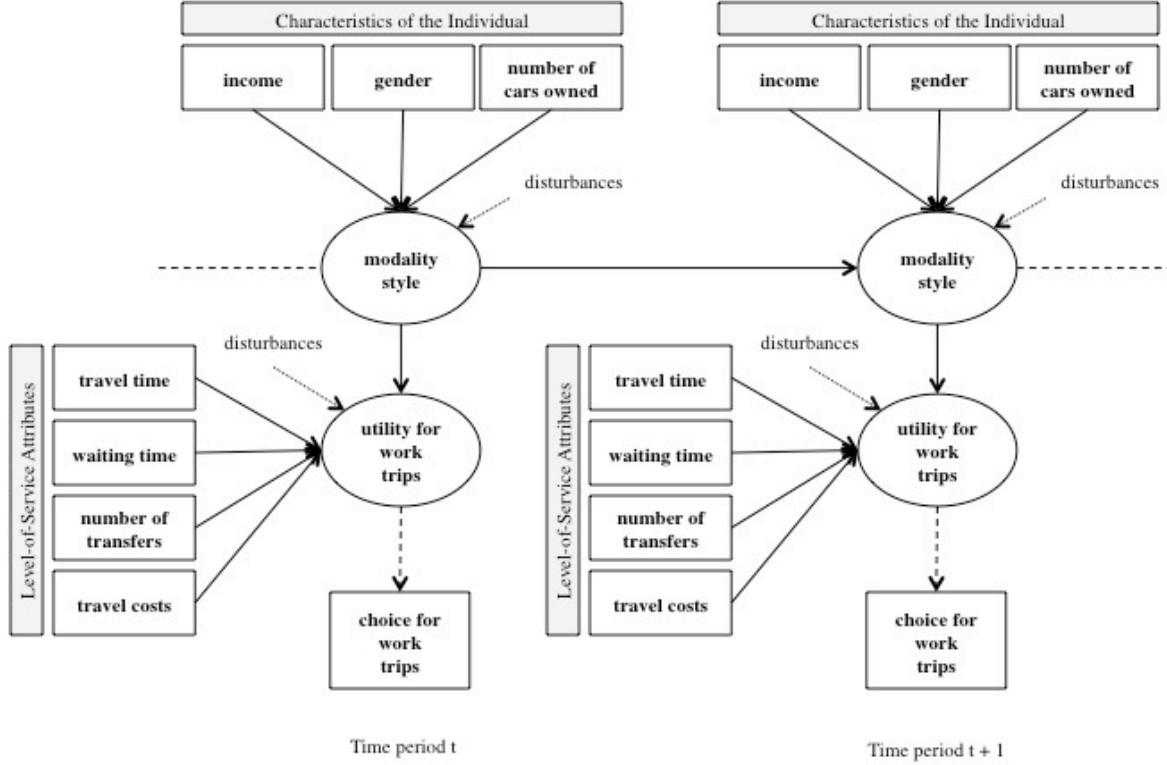
Though the system succeeded in achieving many of these goals, as a result of poor implementation it inadvertently created several new problems. First, the system was introduced in a ‘big-bang’ fashion with no pilot studies or public information campaigns leading up to the change. As a consequence, the first few weeks following the change resulted in great chaos and confusion among users of the city’s public transportation system. Second, the system was designed under the assumption that by the time of its introduction, certain critical bus-only lanes would have been constructed and all buses in the public transit fleet would have been fitted with on-board GPS tracking systems. Neither of these goals was achieved in time, and as a consequence buses ran well below design speeds, introducing significant unreliability into the system. Third, most new bus routes were confined to run along major arterials, increasing the access and egress distances to bus stops, particularly in the suburban corners of the city. And finally, given the hierarchical nature of the new bus system, most bus routes were limited to run within the boundary of a single zone, increasing the number of transfers for trips that required traversing multiple zones. These four factors combined drove a number of passengers to alternative modes of travel, most notably the Metro, which, unlike the bus system, ran at least as reliably as before, resulting in extreme overcrowding on Metro trains, with average occupancy levels during peak hours on certain routes of 5-6 passengers per square meter. As one can imagine, Transantiago generated considerable ill will among city residents, some of which has persisted to this day.

The dataset for the study comes from the Santiago Panel, comprising four one-week waves of pseudo travel-diary data collected over a span of twenty-two months that extends both before and after the introduction of Transantiago. The first wave was conducted in December 2006, three months before the Transantiago was introduced, and the next three waves were implemented in May 2007, December 2007 and October 2008, respectively. Survey respondents were drawn from full-time employees working at one of six campuses of Pontificia Universidad Católica de Chile spread across Santiago. The Panel restricted its attention to trips made to work during the morning peak. Though this limits the number of destinations to just these six campuses, the panel was fortunate in that the distribution of origins was well spread across the city. In all, the Panel interviewed 303 individuals during the first wave, 286 individuals during the second wave, 279 individuals during the third wave, and 258 individuals during the final wave. Considering that the four waves were spread across nearly two years, the Panel has a comparatively low attrition rate. Each of the respondents was asked questions regarding their socioeconomic characteristics; attributes of their morning trip to work; additional activities before, during and after work and their influence, if any, on the respondent’s choice of travel mode; subjective perceptions about the performance of the new system (collected only during the second and third waves); and their level of agreement with attitudinal statements about different aspects of the transportation system, such as safety, reliability and accessibility (collected only during the fourth wave). For more details about the dataset, the reader is referred to Yáñez et al. (2010).

The dataset offers a unique opportunity to investigate the effects of systemic changes in the transportation network on the evolution and persistence of individual preferences. For the



**Figure 5.1:** Shifts across travel modes between subsequent waves of the Panel



**Figure 5.2:** An HHMM showing the influence of individual modality styles on travel mode choice behavior over time

purpose of our analysis, we will be restricting our attention to 220 respondents, each of whom has at least one recorded observations in each of the four waves that constitute the Panel. We aggregate the modal alternatives into seven travel modes: auto, metro, bus, walk, bike, drive to metro, and bus to metro. Figure 5.1 plots the number of trips where individuals switched travel modes between any two subsequent waves of the Panel. The scale of the vertical axes for each of the three plots is the same, to make for easy comparison. As one would expect, the majority of the shift occurs from wave 1 to wave 2, immediately in the wake of the introduction of Transantiago, and most of it away from “bus” and towards “bus to metro”. However, as the system stabilizes over time, so does the behavior of its users, with significantly less movement across travel modes between waves 2 and 3 and waves 3 and 4. Given the nature of the differences between the old and the new system, this is hardly surprising. The more interesting question is: does the shift in observable travel mode choice behavior indicate a corresponding shift in latent individual modal preferences? And does this latter shift, if any, persist beyond the first wave? The next section introduces the methodological framework that we employ to address these related questions.

### 5.2.2 Model Framework

Building on the LCCM framework employed by Chapters 2-4 that examined individual modality styles within a static context, and past research on dynamic discrete choice models, it is proposed

that travel mode choice can be represented as the Heterogeneous Hidden Markov Model (HHMM) illustrated in Figure 5.2. The unobserved states denote different individual modality styles, and the evolutionary path is hypothesized to be a Markov process such that an individual's modality style during the current wave is dependent only on her characteristics and her modality style during the previous wave. The Markov process is heterogeneous in that the transition probabilities between subsequent waves differ from wave to wave. It is further proposed that, conditioned on the individual's modality style during the current wave, her observed mode choices during that wave are independent of the mode choices observed in previous waves. Travel mode choice is additionally some function of observed level-of-service attributes of each of the modal alternatives, such as travel times and travel costs. Individual modality styles, in turn, are some function of sociodemographic characteristics such as gender and household income.

There are three pieces to the model framework. Over the course of the following paragraphs, we introduce the functional form for each of the three sub-components, beginning with the class-specific travel mode choice model, which predicts the probability that individual  $n$  over wave  $t$  and choice situation  $k$  chooses alternative  $j$ , conditional on the individual having modality style  $s$ , denoted as follows:

$$P(y_{ntkj} = 1 | q_{nts} = 1) \quad (1)$$

, where  $y_{ntkj}$  equals one if individual  $n$  over wave  $t$  and choice situation  $k$  chose alternative  $j$ , and zero otherwise; and  $q_{nts}$  equals one if individual  $n$  over wave  $t$  has modality style  $s$ , and zero otherwise. Let  $u_{ntkj|s}$  be the utility of alternative  $j$  over choice situation  $k$  and wave  $t$  for individual  $n$ , given that the individual belongs to modality style  $s$ , expressed as follows:

$$u_{ntkj|s} = \mathbf{x}'_{ntkj} \boldsymbol{\beta}_s + \varepsilon_{ntkj|s} \quad (2)$$

, where  $\mathbf{x}_{ntkj}$  is a vector of attributes of alternative  $j$  over choice situation  $k$  and wave  $t$  for individual  $n$ ;  $\boldsymbol{\beta}_s$  is a vector of parameters specific to modality style  $s$ ; and  $\varepsilon_{ntkj|s}$  is the stochastic component of the utility specification, assumed to be i.i.d. Extreme Value across individuals, waves, choice situations, alternatives and classes with mean zero and variance  $\pi^2/6$ . Assuming that all individuals are utility maximizers, the class-specific choice model may be formulated as the familiar multinomial logit model:

$$P(y_{ntkj} = 1 | q_{nts} = 1) = \frac{\exp(\mathbf{x}'_{ntkj} \boldsymbol{\beta}_s)}{\sum_{j' \in C_{ntk|s}} \exp(\mathbf{x}'_{ntkj'} \boldsymbol{\beta}_s)} \quad (3)$$

, where  $C_{ntk|s}$  is the choice set faced by individual  $n$  over wave  $t$  and choice situation  $k$ , given that the individual belongs to modality style  $s$ . Equation (3) may be combined iteratively over alternatives, choice situations and waves to yield the following conditional probability of observing the vector of choices  $\mathbf{y}_n$  for individual  $n$ :

$$P(\mathbf{y}_n | q_{n1s_1} = 1, \dots, q_{nTs_T} = 1) = \prod_{t=1}^T \prod_{k=1}^{K_{nt}} \prod_{j \in C_{ntk|s}} [P(y_{ntkj} = 1 | q_{nts_t} = 1)]^{y_{ntkj}} \quad (4)$$

, where  $K_{nt}$  is the number of choice situations faced by individual  $n$  over wave  $t$ ; and  $T$  is the number of waves, equal to four in our case. The second piece is the initialization sub-model, which predicts the probability that individual  $n$  has modality style  $s$  during the first wave, expressed as a multinomial logit model:

$$P(q_{n1s} = 1 | \mathbf{Z}_{n1}; \boldsymbol{\alpha}) = \frac{\exp(\mathbf{z}'_{n1} \boldsymbol{\alpha}_s)}{\sum_{s'=1}^S \exp(\mathbf{z}'_{n1} \boldsymbol{\alpha}_{s'})} \quad (5)$$

, where  $\mathbf{z}_{n1}$  is a vector of observable socioeconomic characteristics of individual  $n$  over the first wave;  $\boldsymbol{\alpha}_s$  is a vector of class-specific parameters; and  $S$  is the number of distinct modality styles in the sample population. The number of modality styles is determined by estimating models with different number of classes and using a combination of goodness-of-fit measures and behavioral interpretation to select the most appropriate model. An issue with using HMMs is that the initialization condition must be specified appropriately, or else the model might result in inconsistent estimates.

The third and final piece to the model is the transition sub-model, which predicts the probability that individual  $n$  has modality style  $s$  during wave  $t$ , for  $t > 1$ , conditional on the individual having modality style  $r$  during the previous wave  $t - 1$ , given as follows:

$$P(q_{nts} = 1 | q_{n(t-1)r} = 1) = \frac{\exp(\mathbf{z}'_{nt} \boldsymbol{\gamma}_{tsr})}{\sum_{s'=1}^S \exp(\mathbf{z}'_{nt} \boldsymbol{\gamma}_{ts'r})} \quad (6)$$

, where  $\mathbf{z}_{nt}$  is a vector of observable socioeconomic characteristics of individual  $n$  over wave  $t$ ; and  $\boldsymbol{\gamma}_{tsr}$  is a vector of parameters specific to wave  $t$  and modality style  $s$ , given that the individual has modality style  $r$  over wave  $t - 1$ . Again, equations (5) and (6) may be combined iteratively to yield the probability that individual  $n$  has the vector of modality styles given by  $\{q_{n1s_1}, \dots, q_{nTs_T}\}$ , as follows:

$$P(q_{n1s_1} = 1, \dots, q_{nTs_T} = 1) = P(q_{n1s_1} = 1) \prod_{t=2}^T P(q_{nts_t} = 1 | q_{n(t-1)s_{t-1}} = 1) \quad (7)$$

Equations (4) and (7) may be combined to give the unconditional probability of observing the vector of choices  $\mathbf{y}_n$ , as follows:

$$P(\mathbf{y}_n) = \sum_{s_1=1}^S \dots \sum_{s_T=1}^S P(\mathbf{y}_n | q_{n1s_1} = 1, \dots, q_{nTs_T} = 1) P(q_{n1s_1} = 1, \dots, q_{nTs_T} = 1) \quad (8)$$

Equation (8) may be combined iteratively over all individuals to yield the unconditional likelihood function for the sample population as follows:

$$L(\alpha, \beta, \gamma | y, x, z) = \prod_{n=1}^N P(y_n) \quad (9)$$

, where N denotes the number of individuals in the sample population. Equation (9) has to be computed recursively over each of the time periods. The unknown parameters  $\{\alpha, \beta, \gamma\}$  may be estimated by maximizing equation (9). Since equation (9) is not globally concave, locating the global maximum using gradient-descent based optimization algorithms can prove troublesome, leading us instead to employ the Expectation-Maximization (EM) algorithm, also known in this specific case as the Baum-Welch algorithm (Welch, 2003).

### 5.2.3 Estimation Results

In determining the final model specification, we estimated numerous models where we varied the utility specification, number of classes and choice set assumptions. Here we briefly summarize this process and present key results in Table 5.1 for 3 different models. We found that the relatively small dataset could not support models with more than four classes, and so we report here results for two, three and four latent classes. To facilitate comparison, Table 5.1 enumerates for each model its log-likelihood, the number of parameters estimated, and the corresponding values for the adjusted rho-bar-squared ( $\bar{\rho}^2$ ), the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC). Table 5.1 shows that the four-class model outperforms all other models across all measures of statistical fit. However, in terms of behavioral interpretation, results for the three-class model proved to be more satisfying. Therefore, the three-class model is our preferred model.

Tables 5.2 and 5.3 present detailed parameter estimates for the class-specific model corresponding to travel mode choice and the class membership model. The class membership model includes parameters estimates corresponding to the influence of the sociodemographic characteristics on class membership (assumed to be the same across the initialization sub-model and all transition sub-models) and the class-specific constant corresponding to the initialization and transition sub-models. Over the course of the following paragraphs, we rely on results from each of the constituent sub-models to describe in greater detail the three modality styles identified by the model (class labels are descriptive and not definitive).

**1. Unimodal auto users:** This segment only considers auto when deciding how to travel. As a consequence, the class-specific choice model for the segment is deterministic in that decision-makers belonging to the segment chooses auto, regardless of the level-of-service of any of the other travel modes. The class membership models reveals that women and high-income groups are most likely to be unimodal auto. As one would expect, the segment has the highest average auto ownership rate of 1.46 cars per household.

**Table 5.1:** Summary statistics for different model specifications

Model	Parameters	LL	$\bar{p}^2$	BIC	AIC
Two Class HHMM	31	-2,250	0.543	4,761	4,563
Three Class HHMM	42	-1,971	0.597	4,294	4,026
Four Class HHMM	74	-1,661	0.652	3,943	3,470

**2. Unimodal transit users:** The choice set for unimodal transit users comprises bus, metro and bus to metro. Men and low-income groups are most likely to belong to the segment. Unimodal transit users have the lowest average auto ownership rate of the three classes at 0.49 cars per household. They have a low average value of travel time at 0.4\$/hr and a comparatively higher average value of waiting time at 3.9\$/hr.

**3. Multimodal users:** This segment considers all travel modes when deciding how to travel. Men are more likely to belong to the segment. Income has a negative effect on class membership relative to unimodal auto users but a positive effect relative to unimodal transit users. The segment has a median auto ownership rate of 0.61 cars per household. The average value of travel time is unusually, and maybe unreasonably, high at 30\$/hr; the average value of waiting time is more acceptable at 1.6\$/hr.

What do these estimation results imply in terms of the evolution and persistence of modality styles following the introduction of Transantiago? Figure 5.3 plots results from a sample enumeration showing the number of individuals in the sample population belonging to each of the three modality styles across the four waves of data collection as predicted by the model. Prior to the introduction of Transantiago, roughly half of the sample is unimodal transit, and the other half is more or less evenly split between unimodal auto users and multimodal users. Following the introduction of Transantiago, the number of unimodal transit users in the sample population decreases by nearly 50%, and the number of multimodal users and unimodal auto users increases by 40% and 15%, respectively. The third wave witnesses a marginal rebound in the number of unimodal transit users and a commensurate reduction in the number of multimodal users. The fourth wave heralds a slight decline in the number of unimodal transit users and multimodal users. In contrast, the number of unimodal auto users continues to grow, albeit slowly, over waves three and four. Judging by the parameter estimates for the transition matrices between these two periods, the growth in the segment appears to be driven mostly by rising incomes and increasing vehicle ownership rates. However, the latter could also be attributed to Transantiago and its unintended impact on travel patterns within the city.

Though these findings are promising, they should be interpreted with some caution. Despite the panel nature of the dataset, none of the models estimated in this chapter accounted for serial correlation across observations taken over the same wave and individual. Given that observations for any one individual in the dataset correspond to the same trip to work, the assumption that the stochastic component of the utility specification is independent across trips for the same individual over the same wave is hard to justify. A second shortcoming to the HHMM model framework as it stands now is that the transition matrices are black boxes in that the cause for the

**Table 5.2:** Class-specific travel mode choice model

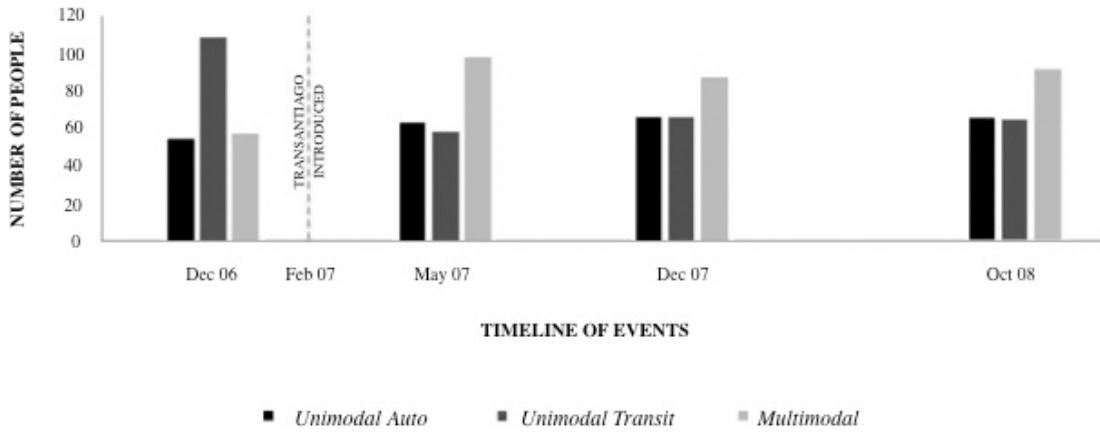
Variable	Class 1	Class 2	Class 3
<i>Alternative-specific constants</i>			
Auto	-	-	0.000
Metro	-	0.000	1.690
Bus	-	5.349	0.540
Walk	-	-	3.931
Bike	-	-	2.037
Drive to metro	-	-	1.905
Bus to metro	-	-2.904	2.019
<i>Level-of-service attributes</i>			
Travel time (minutes)	-	-1.110	-0.004
Waiting time (minutes)	-	-10.767	-0.002
Number of transfers	-	-5.183	-0.004
Travel cost/wage rate (CLP/CLP per minute)	-	-16.109	-0.007

\* Insignificant at the 5% level

**Table 5.3:** Class membership model

Variable	Class 1	Class 2	Class 3
<i>Sociodemographic variables</i>			
Monthly income (1000s CLP)	0.000	-0.227	-0.118
Male	0.000	1.481	1.169
Number of household cars	0.000	-1.271	-1.098
<i>Constants for the initialization model</i>			
Constants	0.000	2.737	1.656
<i>Constants for the transition model from wave 1 to 2</i>			
Given class 1 in wave 1	0.000	-1.480*	-1.471*
Given class 2 in wave 1	0.000	3.583	3.293
Given class 3 in wave 1	0.000	1.907	3.169
<i>Constants for the transition model from wave 2 to 3</i>			
Given class 1 in wave 2	0.000	-3.015	-0.503*
Given class 2 in wave 2	0.000	3.713	2.797
Given class 3 in wave 2	0.000	3.227	3.546
<i>Constants for the transition model from wave 3 to 4</i>			
Given class 1 in wave 3	0.000	-1.262*	0.078*
Given class 2 in wave 3	0.000	4.970	3.988
Given class 3 in wave 3	0.000	2.504	3.227

\* Insignificant at the 5% level



**Figure 5.3:** Results from a sample enumeration illustrating the distribution of survey respondents across the three modality styles at each of the four waves comprising the Santiago Panel

change in preferences is not readily apparent. For example, is the decrease in the number of unimodal transit users between waves 1 and 2 a consequence of increased transfers, greater access and egress times, changes in the fare structure or fluctuations in travel times? The answer hides behind the constants in the transition matrices, and as analysts we can only conjecture based on what we know about the system and how it changed between waves 1 and 2. For the model to be truly general, we need to parameterize the transition matrix as some function of the decision-making environment, just as we parameterized the class membership model when working with LCCMs in a static context. An analogous solution would be to define the transition probabilities as some function of the consumer surplus that a decision-maker would derive from different modality styles between consecutive time periods.

To conclude, in Chapter 4 we argued that incremental improvements in the transportation system, unless accompanied by corresponding shifts in individual modality styles, will result in far smaller changes in travel behavior than would otherwise be predicted by a traditional travel demand model, and that what is needed is a shock to the system, an irremediable change that forces individuals to reconsider how they travel. Findings from this chapter demonstrate that a shock to the system along the lines of Transantiago can and did indeed force individuals to reconsider how they travel. In the case of Transantiago, bad design and poor implementation might have perpetuated a change in modality styles for the worse. But evidence from elsewhere indicates that good design and efficient implementation can just as easily foster a change for the better. For instance, the London Congestion Charge resulted in a 33% decrease in the number of automobiles entering or leaving the congestion zone during charging hours, and a corresponding increase of 29,000 (or 38%) in bus patronage within the Central London area (Transport for London, 2004). TransMilenio, Bogota's bus rapid transit system, was first opened to the public in 2000. Five years later, the system was moving 900,000 passengers per day, 11% of whom were reported to be former automobile drivers (Wright and Fjellström, 2005). The framework developed in this chapter should serve as a foundation on which to build future model forms that can better predict the impact of major systemic changes such as these on both latent modal preferences and observable travel behavior.

## **Chapter 6**

# **Household Modality Styles**

Our discussion thus far has been limited to the individual as the decision-making unit, with attention centered on the influence of individual modality styles on travel mode choice behavior. However, the influence of modality styles must extend beyond travel mode choice behavior and include other dimensions as well, such as, for example, vehicle ownership and transit season pass possession. Estimation results from both Karlsruhe, Germany and the San Francisco Bay Area, United States found that individual modality styles are strongly correlated with vehicle ownership and transit season pass possession. In examining the influence of modality styles on travel mode choice, the model framework developed and applied in Chapters 2 to 5 took these medium-term decisions as exogenous inputs. However, such a causal representation risks endogeneity. Individuals are not auto-oriented because they own more cars or multimodal because they possess a transit season pass. Rather, auto-oriented individuals are more likely to own more cars, and multimodal individuals to possess a transit season pass. This chapter attempts to reverse the causal representation to reflect the influence of modality styles on these additional dimensions of individual and household travel and activity behavior.

Level of vehicle ownership, along with many other dimensions of travel and activity behavior, is a decision that is not made by individuals in isolation from other members of their household. An individual's preferences and choices are strongly shaped by the opinions and behaviors of the people around her (Thorndike, 1938; Davis, 1976; Rose and Hensher, 2004; Zhang et al., 2009), particularly when a choice is made collectively by a group of individuals, as in the case of a household. Past studies that have explored household interdependencies in the context of travel and activity behavior have concerned themselves with the generation and allocation of different household activities between household members and the division of shared resources, such as a car, among household members. Given the collective nature of decision-making that accompanies choices such as level of vehicle ownership, we argue that interaction between household members must also be understood to influence attitudes and beliefs towards, among other things, individual travel and activity behavior, or modality styles.

To reflect the influence of intra-household interactions on individual travel and activity behavior, we introduce the household modality styles construct, characterized by the modality styles of the respective individuals that make up the household. The objective of this chapter is to demonstrate the potential value of the household modality styles construct to the study of travel and activity behavior. Section 6.1 undertakes a literature review of past work on intra-household interactions in the context of individual and household travel and activity behavior, identifying key shortcomings, some of which we attempt to address in this chapter and some that we leave to future work. Section 6.2 develops and estimates a relatively simple model framework that

captures the influence of household modality styles on multiple dimensions of individual and household travel and activity behavior using the MOBIDRIVE dataset.

## 6.1 Literature Review

Traditional travel demand models have tended to focus on the individual as the decision-making unit, and the influence of household interdependencies has only been indirectly captured through the use of household characteristics as explanatory variables (see, for example, Bowman, 1998). Though sociodemographic variables denoting household structure and individual characteristics can adequately represent the role different individuals play within a household, they potentially overlook the direct effect of intra-household interactions and group dynamics on individual travel behavior (Bhat and Pendyala, 2005). Households make a number of short-term and long-term travel and activity decisions collectively, and it is important that household behavior modeling be approached from the viewpoint of group decision-making theory for a behaviorally realistic representation of intra-household interactions.

Past studies on group decision-making have adopted one of two broad approaches. Early representations of the effect of household interactions on observable behavior have built on Becker's seminal work on time allocation theory (Becker, 1965). According to Becker's original formulation, the household may be abstracted as a single unit of decision-making with a common preference function. While the simplicity of the approach is appealing, it overlooks differences across household members in terms of individual preferences and relative influence, and the accompanying process of bargaining and compromise between household members that results in a common preference function for the household as a whole. It has been argued that the use of models that disregard the decision-making mechanism underlying household behavior can lead to incorrect inferences regarding the impact of policies seeking to influence behavior, and conversely, a greater understanding of how households make decisions can strengthen the design of the same policies (see, for example, Alderman et al., 1995; Lundberg et al., 1997; Vermeulen, 2002; Adamowicz et al., 2005; Munro, 2009).

In an attempt to overcome some of these drawbacks, studies on travel demand analysis in the last decade have sought to model explicitly the dynamic interplay between each of the different members that make up a household. The focus of most of these studies has been on the generation and allocation of household activities between household members (see, for example, Gliebe and Koppelman, 2005; Meister et al., 2005; Srinivasan and Bhat, 2005; Kato and Matsumoto, 2009; Roorda et al., 2009; Wang and Li, 2009; and Zhang et al., 2009) and the division of shared resources, such as a car, among household members (see, for example, Golob et al., 1996; Arentze and Timmermans, 2004; Petersen and Vovsha, 2006; and Roorda et al., 2009). These studies have repeatedly confirmed the presence of significant interaction effects between household heads, and the persistence of strong gender, income, employment and life-cycle effects on the pattern of allocation of activities and resources among household members.

However, the focus of this chapter is not on the generation and allocation of different household activities between members or the division of shared resources among members. Rather, we wish to explore the effect of intra-household interaction on individual attitudes and beliefs towards travel and activity behavior, and their subsequent influence on lifestyles and modality styles.

Towards this end, we review studies on group decision-making within both travel behavior and the marketing sciences that have sought to decompose household preferences as some function of the individual preferences of the household members and their relative influence on the decision-making process. Early work by Rao and Steckel (1991) formulates the utility derived by the group as a linear combination of the utility derived by each of its constituent members, weighted by the relative influence exerted by each member. Arora and Allenby (1999) extend their framework to allow the weights to vary by product attribute, thereby capturing individual influence on specific aspects of the shared preference function of the group. Work by Aribarg et al. (2002), Rose and Hensher (2004) and Hensher et al. (2008) extend this framework further to account explicitly for the bargaining process through which individuals may revise their preferences and/or concede them to another member's.

Our work builds upon past research on group decision-making within the context of individual and household travel and activity behavior through the introduction of the household modality style construct. As mentioned previously, many of the dimensions of travel and activity behavior studied by travel demand analysts involve choices made at the level of the household. The preferences of the household are the outcome of a process of negotiation between the individuals that comprise the household and their respective preferences (Corfman, 1991; Lee and Beatty, 2002). In turn, the preferences of the individuals themselves are shaped by the preferences of other household members, and are therefore some reflection of the preferences of the household as a whole (Davis, 1973; Menasco and Curry, 1989). A comprehensive travel demand model must recognize the dialogue between individual and household preferences, or modality styles, that underlies observable behavior. However, unlike some of the studies cited in previous paragraphs that have been very detailed in their representation of the dynamics underlying group decision-making, we won't be as explicit. That being said, we will be relying on findings from these studies to develop a simpler framework that captures the reciprocal influence of individual and household modality styles on each other and concurrently on different dimensions of observable travel and activity behavior.

## 6.2 Proof of Concept

In this section we demonstrate how the household modality styles construct can be integrated into the framework of travel demand models. Section 6.2.1 describes the dataset. Section 6.2.2 presents the methodological framework that captures the influence exerted by household modality styles on three dimensions of individual travel behavior – mode choice for work tours, mode choice for non-work tours, and transit season pass possession – and one dimension of household travel behavior – vehicle ownership. In developing the framework, we ignore the construct of individual modality styles, assuming for now that a household's modality style implicitly determines the modality style of each of its members. Section 6.2.3 discusses estimation results. The work presented in this section is in some ways a rough sketch that serves to demonstrate the value of the household modality styles construct to travel demand analysis. In Chapter 7, we discuss how the framework might be developed further so as to integrate it with both our past work on individual modality styles and the extensive body of literature on group decision-making.

### 6.2.1 The Dataset

The methodological framework that we develop subsequently in Section 6.2.2 will be used to test the influence of household modality styles on three dimensions of individual travel behavior – mode choice for work tours, mode choice for non-work tours, and transit season pass possession – and one dimension of household travel behavior – vehicle ownership. The dataset that we use consists of six-week travel diary surveys administered as part of the MOBIDRIVE research project (Axhausen, 2002). The survey was conducted in the two German cities of Karlsruhe and Halle in the fall of 1999. A total of 317 persons over 6 years of age in 139 households participated in the study. The survey consisted of a face-to-face interview in which socio-demographic characteristics and household information were collected. This was followed by a self-administered travel diary survey in which participants recorded for each trip during the six-week study period the day the trip was made, trip purpose, modes used, departure and arrival times, accompanying individuals, etc. During post-processing, the level-of-service for all modes (walk, bike, auto, and transit) was generated from transportation network data for the city of Karlsruhe. More details on the survey and the resulting dataset can be found in Axhausen (2002).

Since level-of-service attributes for all modes are unavailable for the city of Halle, the dataset is narrowed to trips contained within Karlsruhe. Unlinked trips are aggregated into home-based work and non-work tours, following an approach similar to Cirillo et al. (2006). For each trip, the data contain the modal chain (including access and egress modes for transit). A “main mode” for a tour is defined to be the mode used to cover the greatest motorized distance, tacitly assuming that mode choice is dictated by the longest leg of the tour. Four main modes are defined: auto, transit, bike, and walk. Trips taken as car passengers are counted under auto, as are trips made by motorcycle (less than 2 percent). Though car passengers are expected to be different behaviorally than car drivers, data didn’t allow us to treat the two as independent travel modes. Consequently, individuals belonging to households with no cars were specified to have access to “auto,” since they could potentially get a ride from a neighbor or a friend.

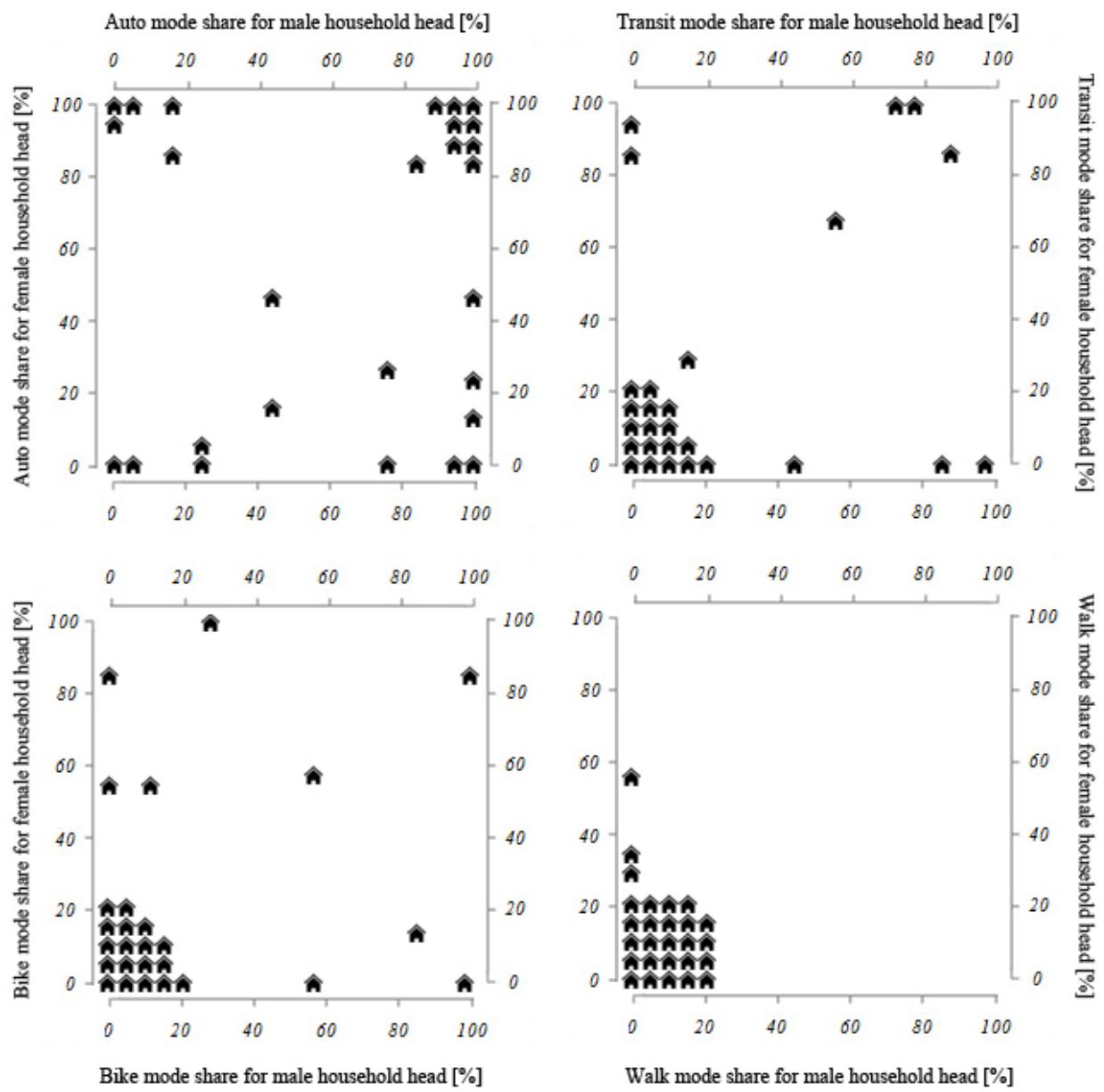
For the purposes of model development, we narrow our dataset to tours contained within Karlsruhe, since level-of-service attributes for all modes are unavailable for the city of Halle. In the case of work tours we consider both simple work-only tours without any additional stops, and tours on which the individual made additional stops on the way to work, on the way back, or both. However, for work tours with additional stops, we use the same level-of-service attributes in our mode choice models as those for the usual tour from home to work and back, and the presence or absence of intermediate stops is represented by a binary variable. This is the typical practice with activity-based models, where destination choice for intermediate stops is often predicated on mode choice (see, for example, Bradley et al., 2010). For instance, the location an individual chooses to stop on his way back from work to buy groceries might depend on whether he’s walking, on the bus, or in a car. Therefore, it is suspect to compare different modes for the specific tour route, since a different tour might have been undertaken had a different mode been chosen. For these same reasons, in the case of non-work tours we limit our attention to only those tours with two constituent trips, one each to and from the main destination, with no additional stops along the way. Consideration of tours with intermediate stops would call for a model that predicts destination choice as well. The absence of land use data and level-of-service attributes for all possible tours, and not merely the tour that was made, preclude estimation of such a model, and therefore we exclude non-work tours with multiple stops. We further limit our

attention to the two household heads from any household. The two household heads were identified by removing adult children and elderly family members from the dataset. In all cases, the two household heads were one female and one male. These restrictions reduce the dataset to 1235 work tours and 2576 non-work tours made between 96 individuals from 48 households over the six-week observation period.

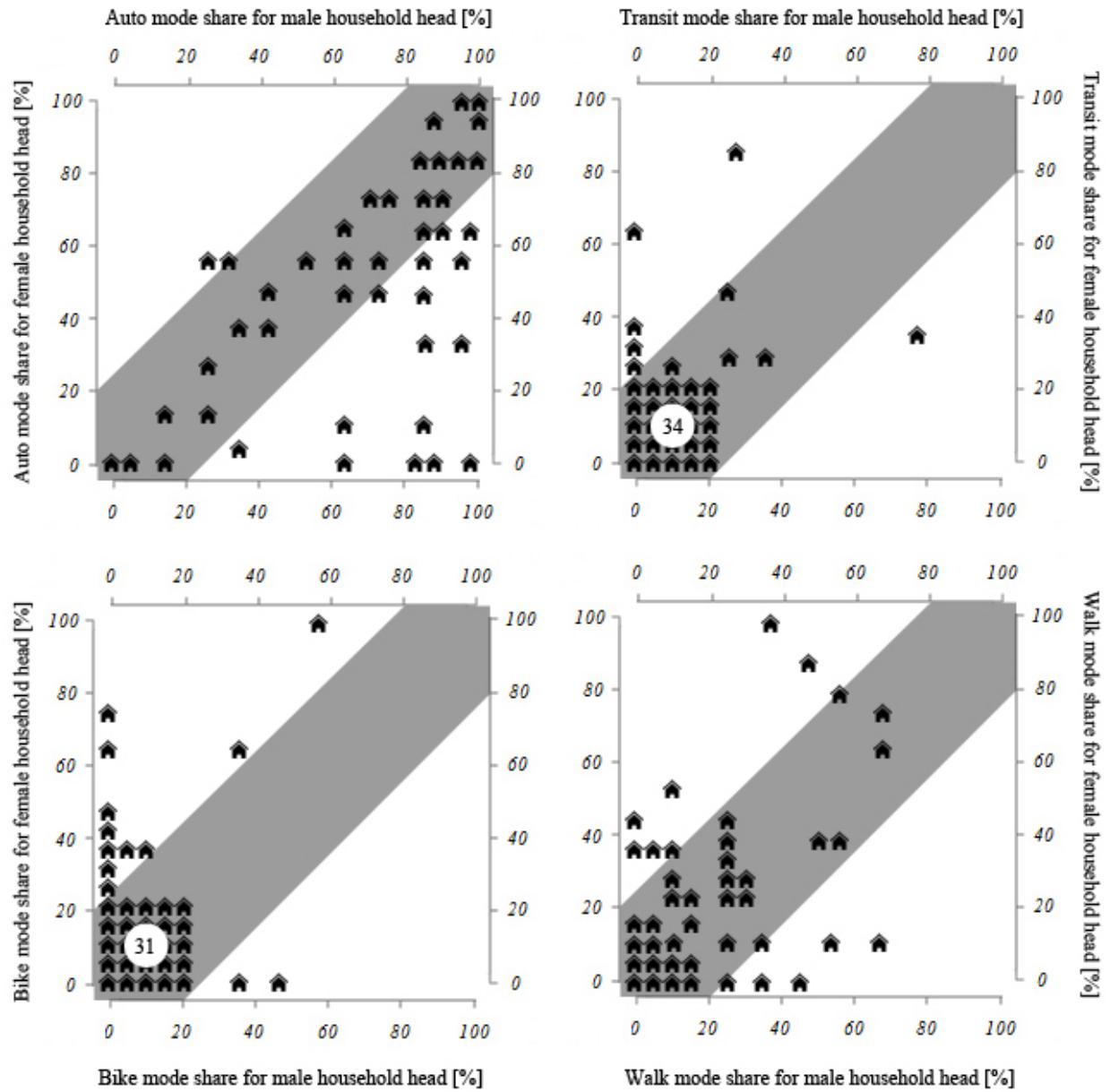
To explore potential correlation between latent modal preferences of different individuals from the same household, Figures 6.1 and 6.2 plot mode shares for the male and female household heads for each of the four modes across work and non-work tours, respectively. For the purpose of the plots, we limit our attention to households where each of the two households head had made five or more tours belonging to that activity type. Consequently, the sample size stands at 27 households for Figure 6.1 and 44 households for Figure 6.2. Each house figure on the plot denotes one household in the sample population. For certain mode and activity type combinations, there is a substantial number of households that have a near zero mode share for both household heads; as before, these are plotted in the lower left portion of the graph and we write the total number of individuals inside a white circle.

For work tours, it's interesting to note the large concentration of households near the corners for each of the four travel modes, suggesting that most individuals optimize their choice once for commuting and then use that mode on a more permanent basis. In general, the male household head is more reliant on the auto for his commute needs than his female counterpart, who seems more inclined to take transit or walk. Bicycling appears to be equally popular with both genders. The scatter plots for non-work tours is considerably different from the plot for work tours. There are two broad trends to be noted. First, a significant number of households are accumulated along the diagonals across all four modes and both activity types, most perceptibly for auto and walk, and somewhat less so for transit and bike. Though partially attributable to joint activity participation, this does appear to suggest some degree of correlation between the modality styles of the two household heads for non-work tours. Second, many households can be seen stretched along the bottom and left axes, particularly in the plots for transit and bike, indicating the presence of households where one member is especially inclined to using that mode of travel, and the other equally averse. A closer inspection reveals that most of these households own a single car. Therefore, their distribution is a likely consequence of the allocation of limited shared resources between the two household heads such that one head enjoys access to the car all the time. For non-work tours, there are only 2 households above the diagonal for auto, as opposed to 15 households below it, suggesting that, more often than not, it is the male household head that gets preference over use of the car.

The descriptive analysis presented here serves to indicate that intra-household interactions, whether through their direct effect on the allocation of shared resources or joint activity participation, or more subtly in the way that they influence individual attitudes and beliefs towards different travel modes, do appear to exercise some control over latent individual modal preferences, and the hypothesis deserves to be explored in fuller detail.



**Figure 6.1:** Scatter plot of mode shares for work tours for the male and female household head; each house figure on the plot represents one household in our sample.

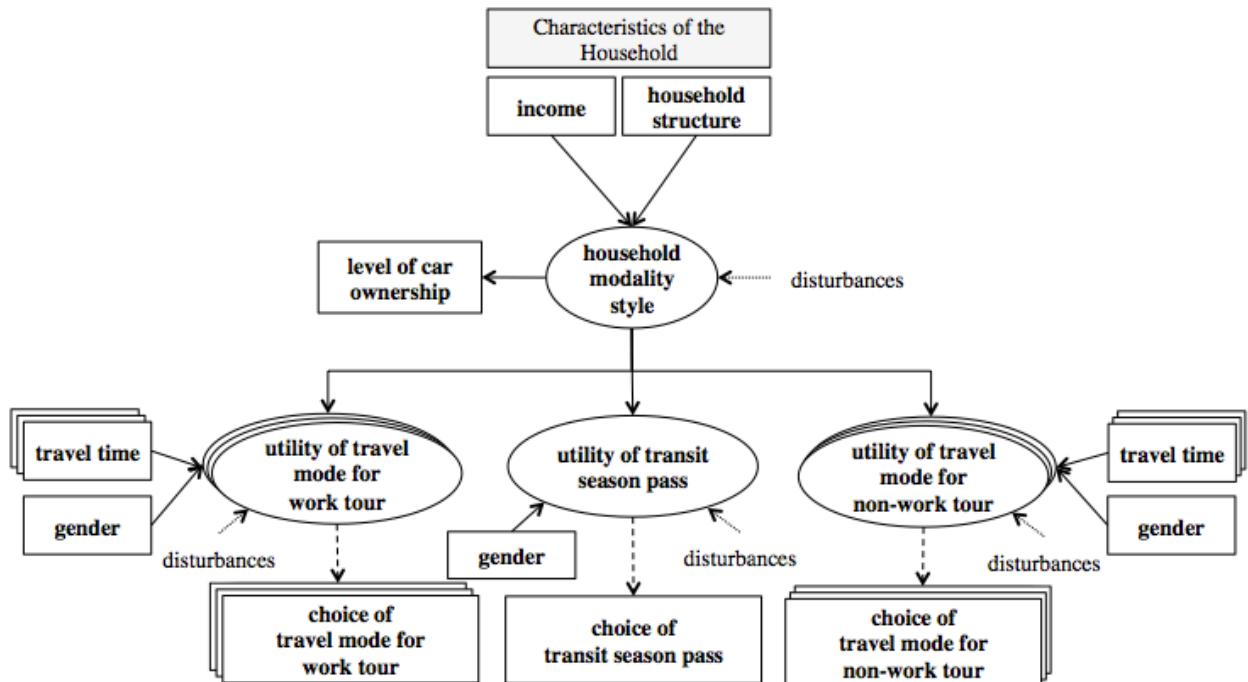


**Figure 6.2:** Scatter plot of mode shares for non-work tours for the male and female household head; each house figure on the plot represents one household in our sample, except for the transit and bike plots where the number of individuals in the lowest quintile is written into the plot inside the white circle. The grey diagonal band is to help identify households where the two household heads have relatively similar mode shares.

## 6.2.2 Model Framework

In developing a framework for our model, we use LCCMs without feedback as a starting point. We argue that discrete household modality styles exist, that these modality styles are indicative of higher-level household orientations that influence individual and household choices across multiple dimensions, and consequentially households with different modality styles exhibit different travel behavior. The LCCM framework is particularly appropriate given the discrete nature of heterogeneity hypothesized here (Gopinath, 1995).

The model framework is illustrated in Figure 6.3. Consistent with the usual notation, ellipses denote unobservable variables and rectangles denote observable variables, while dashed arrows represent measurement equations and solid arrows represent structural equations. As mentioned before, LCCMs comprise two components: a class membership model and a class-specific choice model. The latent classes represent different household modality styles, and conditioned on a household's modality style we have separate class-specific models for choices made at both the household level and the individual level. Class membership is hypothesized to be a function of observable household characteristics. The disturbances denote unobserved factors that influence class membership, assumed to be i.i.d. Extreme Value across households. The class-specific choice model depicts the influence exerted by a single overarching household modality style on three dimensions of individual travel behavior – mode choice for work tours, mode choice for non-work tours, and transit season pass possession – and one dimension of household



**Figure 6.3:** Model of travel mode choice, transit season pass possession and level of vehicle ownership for the MOBIDRIVE dataset

travel behavior – level of vehicle ownership. The four choice dimensions are correlated through the modality styles construct. Travel mode choices are further conditioned on the travel times of the different travel mode alternatives and the gender of the household head under consideration. Unfortunately, cost data isn't available for any of the travel modes, and so no price parameters could be estimated for the model. In terms of the error structure, the class-specific choice models are multinomial logit models. Given an observation period of six weeks, a multinomial logit specification for the travel mode choice models overlooks potential serial correlation across choices made by the same individual, and future work must remedy this limitation. Heterogeneity across modality styles includes alternative-specific constants, sensitivity to travel times and the parameters corresponding to gender of the household head. Transit season pass possession is conditioned on the gender of the household head. Level of vehicle ownership is specified as a constants-only model. Note that the effect of household-level sociodemographic variables on the level of vehicle ownership is indirectly captured through the sociodemographic variables that enter the class-membership model. Both class-specific choice models are multinomial logit as well. The class membership and class-specific choice models together explicitly integrate the household modality styles construct with multiple dimensions of individual and household travel and activity behavior.

Over the course of the following paragraphs, we introduce the framework in greater detail. We begin with the class membership model, which predicts the probability that household  $h$  belongs to latent class  $s$ , written as:

$$P(q_{hs} = 1) \quad (1)$$

, where  $q_{hs}$  equals one if household  $h$  belongs to latent class  $s$ , and zero otherwise. The class membership model can take a wide variety of functional forms, the most common being the multinomial logit model:

$$P(q_{hs} = 1) = \frac{\exp(\mathbf{z}'_h \boldsymbol{\gamma}_s)}{\sum_{s'=1}^S \exp(\mathbf{z}'_h \boldsymbol{\gamma}_{s'})} \quad (2)$$

, where  $\mathbf{z}_h$  is a vector of characteristics of household  $h$ ; and  $\boldsymbol{\gamma}_s$  is a vector of parameters associated with the household's characteristics.

The second piece to the LCCM is the class-specific choice model, which might be decomposed further into choices that are made at the level of the household and choices that are made at the level of the individual. We first look at choices made by the household as a whole, namely level of vehicle ownership. Each household chooses from among three alternatives: zero cars, one car or two and more cars. The class-specific probability that household  $h$  chooses level of vehicle ownership  $j$ , conditional on the household belonging to latent class  $s$ , may be written as follows:

$$P(l_{hj} = 1 | q_{hs} = 1) \quad (3)$$

, where  $l_{hj}$  equals one if household  $h$  has level of vehicle ownership  $j$ , and zero otherwise. Let  $u_{hj|s}$  be the utility of level of vehicle ownership  $j$  for household  $h$  given that the household

belongs to latent class  $s$ . Since the class-specific choice model for level of vehicle ownership is constants-only,  $u_{hj|s}$  may be expressed as follows:

$$u_{hj|s} = \alpha_{js} + \varepsilon_{hj|s} \quad (4)$$

, where  $\alpha_{js}$  is a parameter specific to level of vehicle ownership  $j$  and modality style  $s$ ; and  $\varepsilon_{hj|s}$  is the stochastic component of the utility specification, assumed to be i.i.d. Extreme Value across households, levels of vehicle ownership and classes with mean zero and variance  $\pi^2/6$ . Assuming that all households are utility maximizers, the class-specific choice model may be then be formulated as the familiar multinomial logit expression:

$$P(l_{hj} = 1 | q_{hs} = 1) = \frac{\exp(\alpha_{js})}{\sum_{j'=1}^J \exp(\alpha_{j's})} \quad (5)$$

, where  $J$  denotes the number of alternatives, equal to three in this case. Equation (5) may be combined iteratively over all levels of vehicle ownership  $J$  to yield the following conditional probability of observing the vector of level of vehicle ownership  $\mathbf{l}_h$  for household  $h$ :

$$P(\mathbf{l}_h | q_{hs} = 1) = \prod_{j=1}^J [P(l_{hj} = 1 | q_{hs} = 1)]^{l_{hj}} \quad (6)$$

Moving on to choices made at the individual level, the class-specific probability that individual  $n$  belonging to household  $h$  over choice dimension  $d$  and choice situation  $t$  chooses alternative  $j$ , conditional on the household belonging to latent class  $s$ , may be specified as follows:

$$P(y_{hndtj} = 1 | q_{hs} = 1) \quad (7)$$

, where  $y_{hndtj}$  equals one if individual  $n$  belonging to household  $h$  over choice dimension  $d$  and choice situation  $t$  chose alternative  $j$ , and zero otherwise. Let  $u_{hndtj|s}$  be the utility of alternative  $j$  over choice situation  $t$  for choice dimension  $d$  and individual  $n$  belonging to household  $h$  given that the household belongs to latent class  $s$ , which may be expressed as follows:

$$u_{hndtj|s} = \mathbf{x}'_{hndtj} \boldsymbol{\beta}_{ds} + \varepsilon_{hndtj|s} \quad (8)$$

, where  $\mathbf{x}_{hndtj}$  is a vector of attributes of alternative  $j$  over choice situation  $t$  for choice dimension  $d$  and individual  $n$  belonging to household  $h$ ;  $\boldsymbol{\beta}_{ds}$  is a vector of parameters for choice dimension  $d$  specific to the class  $s$ ; and  $\varepsilon_{hndtj|s}$  is the stochastic component of the utility specification, assumed to be i.i.d. Extreme Value across households, individuals, choice dimensions, choice situations, alternatives and classes with mean zero and variance  $\pi^2/6$ . Assuming that all individuals are utility maximizers, the class-specific choice model may be formulated as follows:

$$P(y_{hndtj} = 1 | q_{hs} = 1) = \frac{\exp(x'_{hndtj} \beta_{ds})}{\sum_{j' \in C_{hndt|s}} \exp(x'_{hndtj'} \beta_{ds})} \quad (9)$$

, where  $C_{hndt|s}$  is the choice set for choice situation t for choice dimension d and individual n belonging to household h given that the household belongs to latent class s. Equation (9) may be combined iteratively over alternatives, choice situations and choice dimensions to yield the following conditional probability of observing the vector of choices  $\mathbf{y}_{hn}$  for decision-maker n belonging to household h:

$$P(\mathbf{y}_{hn} | q_{hs} = 1) = \prod_{d=1}^{D_{hn}} \prod_{t=1}^{T_{hnd}} \prod_{j \in C_{hndt|s}} [P(y_{hndtj} = 1 | q_{hs} = 1)]^{y_{hndtj}} \quad (10)$$

, where  $T_{hnd}$  denotes the number of distinct choice situations observed for choice dimension d and individual n belonging to household h; and  $D_{hn}$  denotes the number of choice dimensions at the individual level for individual n belonging to household h, equal to three in this case. Equations (6) and (10) may be combined iteratively over individuals to yield the following conditional probability of observing the vector of choices  $\mathbf{y}_h = \{\mathbf{l}_h, \mathbf{y}_{hn}\}$  for household h:

$$P(\mathbf{y}_h | q_{hs} = 1) = P(\mathbf{l}_h | q_{hs} = 1) \prod_{n=1}^{N_h} P(\mathbf{y}_{hn} | q_{hs} = 1) \quad (11)$$

, where  $N_h$  denotes the number of individuals belonging to household h, equal to two in our case (the male and female household heads). Equation (11) may be combined iteratively with equation (1) over all households and modality styles to yield the unconditional likelihood function for the sample population as follows:

$$L(\alpha, \beta, \gamma; \mathbf{y}, \mathbf{x}, \mathbf{z}) = \prod_{h=1}^H \sum_{s=1}^S P(q_{hs} = 1) P(\mathbf{y}_h | q_{hs} = 1) \quad (12)$$

, where H denotes the number of households in the sample population and S denotes the number of modality styles. The unknown parameters  $\{\alpha, \beta, \gamma\}$  may be estimated by maximizing the likelihood function given by equation (12). The number of modality styles is determined endogenously through a comparison across models with different numbers of classes in terms of both statistical measures of fit and behavioral interpretation.

### 6.2.3 Estimation Results

In determining the final model specification, we estimated numerous models where we varied the utility specification and the number of classes. Here we briefly summarize this process and present key results in Table 6.1 for 4 different models. We found that the relatively small dataset could not support models with more than four classes, and so we report here results for one, two,

**Table 6.1:** Summary statistics for different model specifications

Model	Parameters	LL	$\bar{p}^2$	BIC	AIC
Multinomial Logit	18	-3,701	0.243	7,551	7,439
Two Class LCCM	40	-3,172	0.347	6,674	6,424
Three Class LCCM	56	-2,843	0.410	6,147	5,853
Four Class LCCM	77	-2,659	0.444	5,951	5,547

three and four latent classes. To facilitate comparison, Table 6.1 enumerates for each model its log-likelihood, the number of parameters estimated, and the corresponding values for the adjusted rho-bar-squared ( $\bar{p}^2$ ), the Bayesian Information Criterion (BIC) and the Akaike Information Criterion (AIC). Since the AIC and  $\bar{p}^2$  are equivalent measures of fit, we will be restricting our attention to the AIC and the BIC when comparing different models. Table 6.1 shows that the four-class model outperforms all other models across all measures of statistical fit. In terms of behavioral interpretation as well, results for the four-class model proved to be the most satisfying. Therefore, the four-class model is our preferred model.

Tables 6.2, 6.3, 6.4 and 6.5 present detailed parameter estimates for the class-specific models corresponding to travel mode choice for work tours, travel mode choice for non-work tours, transit season pass possession and level of vehicle ownership. Table 6.6 presents detailed parameter estimates for the class membership model. Note that an alternative-specific constant of -100 denotes an alternative that has been ruled out of the choice set for a household/individual belonging to that particular class. For example, the results for the auto ownership choice model, listed in Table 6.2, state that the probability that a household belonging to Class 1 owns no cars is zero, and the same holds for households belonging to Classes 3 and 4 as well. Regarding estimates for the class membership model, shown in Table 6.6, though none of the lifecycle variables are significant, it merits repeating that there are only 48 households in our sample, and that the large p-values are attributable to the small sample size of our dataset. Despite that, there are some interesting results to be had from Table 6.6. Over the course of the following paragraphs, we rely on results from each of the constituent sub-models to describe in greater detail the four household modality styles identified by the model (class labels are descriptive and not definitive).

To further underscore behavioral differences between the four classes, a sample enumeration is carried out, and the results are illustrated in Figure 6.4. The class membership probabilities for each household are summed to arrive at the expected size of the four household modality style segments. The class-specific probability of choosing an alternative on a tour is weighed by the class membership probability for the respective household to which the individual belongs, and the product is summed over all tours to arrive at the expected modal split for each of the four modality styles. A similar procedure is used to calculate the socioeconomic composition of each class. Before we describe the classes in greater detail, it is worth reemphasizing that the estimation process is exploratory in that the number of classes and the behavior of each class are uncovered in the course of testing different model specifications. The class labels are assigned based on what the estimation results imply regarding behavior.

**Table 6.2:** Class-specific travel mode choice model for work tours

<b>Variable</b>	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>
<i>Alternative-specific constants</i>				
Auto	0	0	0	0
Transit	0.199*	2.514	-5.387	-3.494
Bike	-0.729	0.217*	-5.022	0.089*
Walk	0.129*	1.122*	-5.294	-3.057
<i>Level-of-service attributes</i>				
Travel time (minutes)	-4.380	-0.937*	-0.098*	-0.437*
<i>Male</i>				
Auto	0	0	0	0
Transit	-0.031*	0.313*	-100	0.037*
Bike	-2.230	0.498*	-0.355*	0.434
Walk	-0.871*	-100	1.726	1.608

\* Insignificant at the 10% level

**Table 6.3:** Class-specific travel mode choice model for non-work tours

<b>Variable</b>	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>
<i>Alternative-specific constants</i>				
Auto	0	0	0	0
Transit	-0.973	2.120	-3.967	-1.352
Bike	0.575	3.013	-2.470	-1.240
Walk	1.973	1.879	-1.953	1.106
<i>Level-of-service attributes</i>				
Travel time (minutes)	-9.052	-0.566	-0.364	-1.557
<i>Male</i>				
Auto	0	0	0	0
Transit	-0.256*	-1.190	-0.153*	-0.166*
Bike	-3.518	-1.784	-0.557	-1.240
Walk	1.146	-0.170*	0.447	-0.174*

\* Insignificant at the 10% level

**Table 6.4:** Class-specific choice model for transit season pass possession

Variable	Class 1	Class 2	Class 3	Class 4
<i>Alternative-specific constants</i>				
No season pass	0	0	0	0
Season pass	-0.982*	3.000*	-0.600*	-100
<i>Male</i>				
No season pass	0	0	0	0
Season pass	-0.904*	2.831*	-100	-100

\* Insignificant at the 10% level

**Table 6.5:** Class-specific choice model for level of vehicle ownership

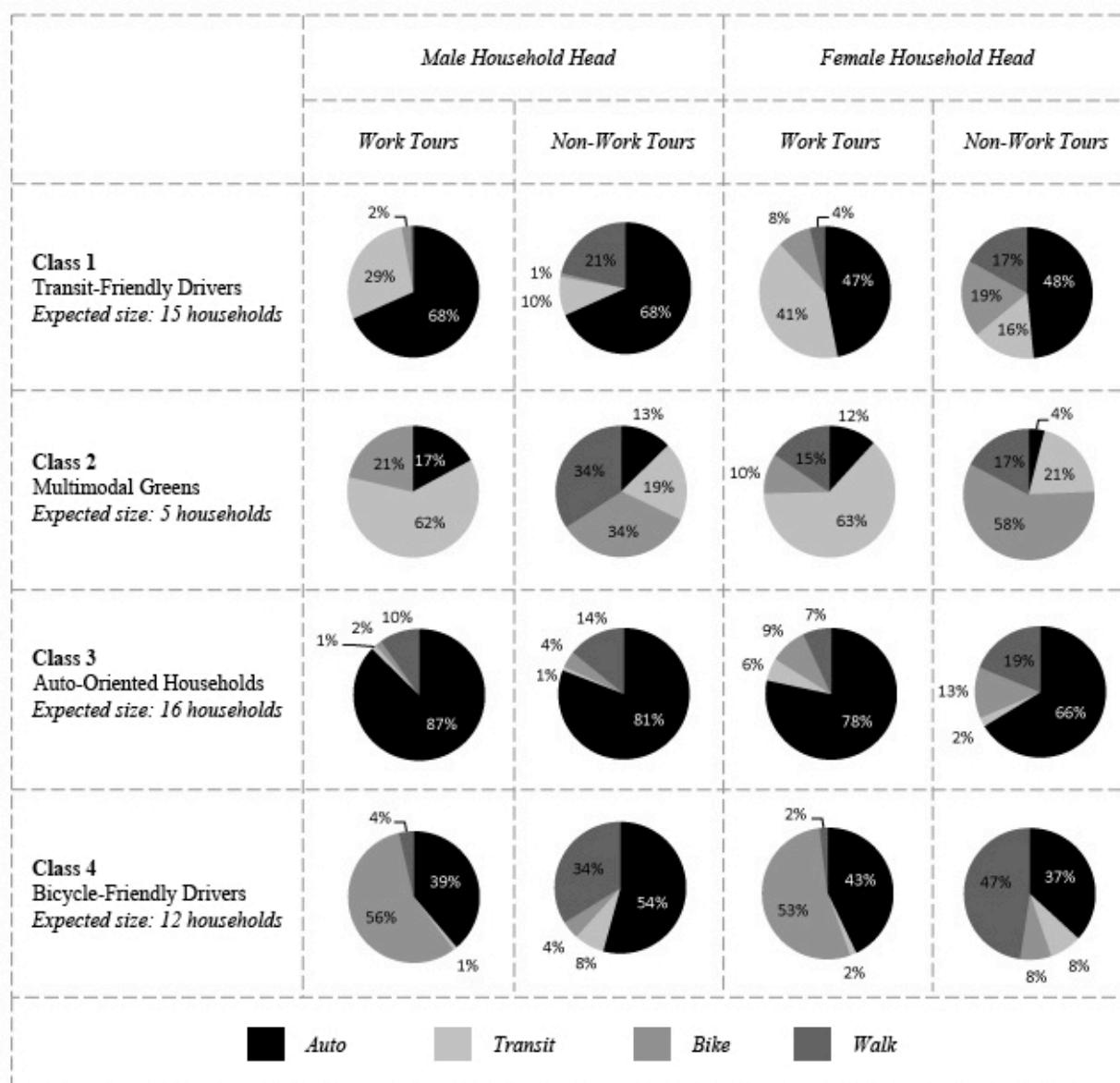
Variable	Class 1	Class 2	Class 3	Class 4
<i>Alternative-specific constants</i>				
Zero cars	-100	0	-100	-100
One car	0	0.405*	0	0
Two or more cars	-0.634*	-100	0.569*	-1.605

\* Insignificant at the 10% level

**Table 6.6:** Class membership model

Variable	Class 1	Class 2	Class 3	Class 4
Class-specific constant	0	-0.289*	-0.675*	-0.519*
High income (binary)	0	-0.533*	1.015*	0.776*
Young children (binary)	0	-100	0.473*	-1.058*
Adult children (binary)	0	0.398*	0.007*	0.044*
Empty nester (binary)	0	-0.217*	0.531*	0.768*

\* Insignificant at the 10% level



**Figure 6.4:** Sample enumeration results for travel mode choice for work and non-work tours as predicted by the preferred four-class model specification

**Transit-Friendly Drivers (Class 1):** Consisting of 31% of the sample population with an expected size of 15 households, transit-friendly drivers are dependent on the car for a majority of their travel, having the second highest propensity for car ownership. However, they are open to other modes as well. In particular, both household heads display a strong willingness to take transit for work tours. For non-work tours, walking and driving are the preferred modes. The male household head appears to be especially averse to using the bicycle for both work and non-work tours, as evidenced by the large negative value for the bike-specific male binary variable. For non-work tours, female household heads demonstrate a greater willingness to bike.

**Multimodal Greens (Class 2):** Comprising 11% of the sample population with an expected size of 5 households, multimodal greens are the smallest of the four household modality styles uncovered by the model. Multimodal greens display the lowest dependence on the car for their travel needs, relying mostly on transit for work tours and a combination of transit, bike and walk for non-work tours. Both household heads are similar in their predisposition towards different travel modes. As one would expect, multimodal greens have the lowest propensity for car ownership and the highest propensity for having a transit season pass. The presence of young children automatically rules out the possibility that the household is multimodal green.

**Auto-Oriented Households (Class 3):** Consisting of 33% of the sample population with an expected size of 16 households, auto-oriented households have the strongest propensity for auto ownership and the second lowest propensity for transit season pass possession: one-in-three female household heads own a transit season pass, while the probability that the male household head owns a transit season pass is zero. Large negative and significant alternative-specific constants for transit, bike and walk confirm a strong preference for driving, so much so that transit doesn't enter the male household head's choice set for work tours. Nearly four-fifths of both work and non-work tours are made by car, with the female household head more likely to take other modes than her male counterpart. Income and the presence of young children are positively correlated with class membership.

**Bicycle-Friendly Drivers (Class 4):** Comprising 25% of the sample population with an expected size of 12 households, individuals belonging to the class are dependent on the car for less than half of their travel needs for both work and non-work tours. In fact, both household heads are very similar in their travel behavior, displaying a strong predisposition towards the bike for work tours and towards walking for non-work tours. The class displays a distinct disinclination towards transit. These findings are supported by results for the level of vehicle ownership and transit season pass possession models. Bicycle-friendly drivers have the second lowest propensity for car ownership, and the probability that either household head has a transit season pass is zero. Empty nesters are most likely to belong to the class.

These results serve as a good starting point for a more comprehensive framework that recognizes the influence of modality styles on all dimensions of individual and household travel and activity behavior. While there is significant correlation between modal preferences of heads of the same household, there are notable differences as well. In general, female household heads are found to be less reliant on the automobile for their mobility requirements than their male counterparts. The model results further find that short-term individual decisions, such as mode choice, are inextricably linked with more long-term individual and household decisions, namely level of vehicle ownership and transit season pass possession. Though certain life-cycle variables, such

as the presence of young children in the household, are found to have a significant bearing on household modality styles, most of the life-cycle variables included in the class- membership model lack significant explanatory power. While it is tempting to conclude that household modality styles might be truly latent, it merits remembering that the number of households in our sample was a mere forty-eight.

In developing a model framework in Section 6.2.2 that recognizes the influence of household modality styles on multiple dimensions of individual and household travel and activity behavior, we made two simplifying assumptions: (1) the individual modality styles construct was not represented explicitly in the model framework but was absorbed by our definition of the household modality styles construct; and (2) we overlooked the notion of preference endogeneity, electing instead to subscribe to the usual microeconomic assumption that preferences are characteristics of the individual and household that are stable over time. In Chapter 7, we discuss how the model framework might be integrated with the framework presented in Chapter 2 in an attempt to overcome these two shortcomings and develop a comprehensive travel demand model that captures the influence of modality styles on all dimensions of individual and household travel and activity behavior.

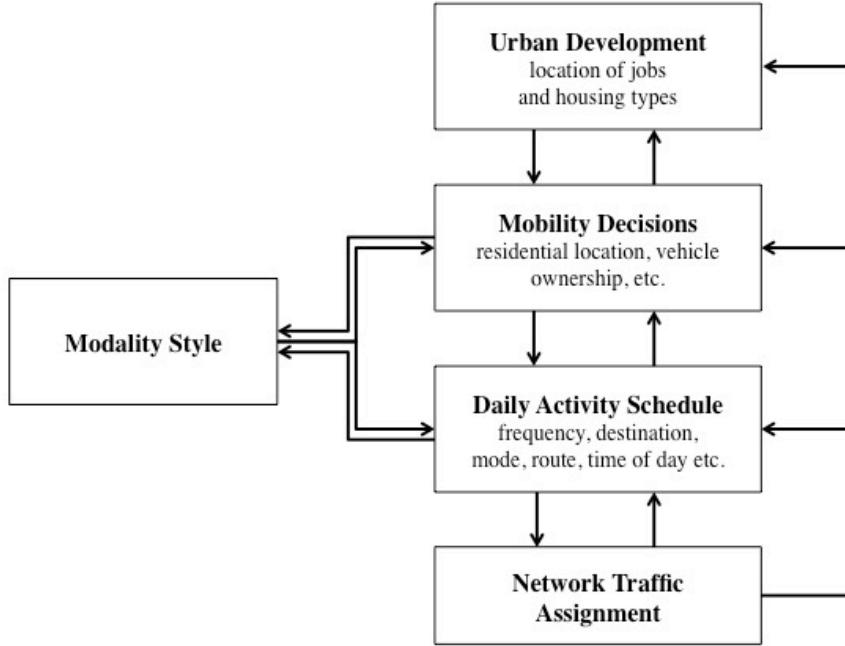
# **Chapter 7**

## **Conclusions**

Modality styles are defined as behavioral predispositions towards a certain travel mode or set of travel modes that an individual habitually uses. They are reflective of higher-level orientations, or lifestyles, that are hypothesized to influence all dimensions of an individual's travel and activity behavior. For example, in the context of travel mode choice different modality styles may be characterized by the set of travel modes that an individual might consider when deciding how to travel, her sensitivity, or lack thereof, to different level-of-service attributes of the transportation (and land use) system when making that decision, and the socioeconomic characteristics that predispose her one way or another. The objectives of this dissertation were to understand and quantify different modality styles and to demonstrate how the modality styles construct can be integrated within the framework of traditional travel demand models. This final chapter is organized as follows: Section 7.1 summarizes major findings and contributions of our work; Section 7.2 outlines a roadmap for future research in the area; and Section 7.3 concludes the chapter with closing remarks.

### **7.1 Findings and Contributions**

Modality styles are hypothesized to influence all dimensions of an individual's travel and activity behavior. Figure 7.1 adapts a schematic from Bowman (1995) to illustrate how the influence exerted by modality styles might be represented within the context of a comprehensive system for metropolitan travel forecasting and policy analysis. The focus of this dissertation has been on the relationship between modality styles, mobility decisions and the daily activity schedule. We began by developing a discrete choice model framework that captures the influence of individual modality styles on travel mode choice behavior. Different modality styles were specified as latent classes. Heterogeneity across modality styles included both the travel modes considered and the relative sensitivity to different level-of-service attributes. Class membership was hypothesized to be a function not only of household and individual characteristics and medium and long-term travel and activity decisions, but also of the consumer surplus offered by each class, which in turn is a function of alternative attributes, taste parameters and choice sets. The framework allows preferences to be both heterogeneous across decision-makers and sensitive to changes in the decision-making environment as represented by changes in alternative attributes. Correlation across multiple choice dimensions is captured through the class membership model. In developing the framework, we synthesized recent advances in discrete choice methods in the sub-domains of taste heterogeneity and choice set generation and made methodological contributions of our own to the sub-domains of preference endogeneity and simultaneous choice models.



**Figure 7.1:** The influence of modality styles in the context of a comprehensive forecasting model system  
(adapted from Bowman, 1995)

The framework was tested using two very distinct travel diary datasets from two very culturally and geographically distinct regions. For both datasets, the framework was found to outperform traditional models of travel mode choice behavior in terms of statistical measures of fit, attesting to its robustness. In terms of behavioral theory too, estimation results for the model framework proved to be more satisfying. The first dataset was collected in Karlsruhe, Germany and comprises a relatively small sample of 119 individuals surveyed over a fairly long observation period of six weeks. Estimation results indicated the presence of habitual drivers who display a strong bias for using the automobile and multimodal individuals who exhibit variation in their modal preferences. Multimodal behavior was further distinguished by those who appear to be sensitive to travel times and those who appear to be insensitive. The second dataset was collected in the San Francisco Bay Area in the United States and consists of a relatively large sample of 26,350 individuals surveyed over a fairly short observation period of two days. Estimation results uncovered six modality styles that are distinguishable from one another by the kinds of individuals that belong to each of the six modality styles, their latent preferences for different travel modes and the relative importance that they attach to different level-of-service attributes of each of the travel modes. For example, two of the six modality styles comprising 30% of the sample population only consider the car when deciding how to travel. These two modality styles, labeled inveterate drivers and car commuters, can further be distinguished from one another by their value of travel time. Inveterate drivers have a very low value of in-vehicle travel time of 0.55 \$/hr for mandatory tours and are insensitive to in-vehicle travel times for non-mandatory tours. Car commuters have a value of in-vehicle travel time of 6.95 \$/hr for mandatory tours and are insensitive to travel costs for non-mandatory tours, indicating a very high value of in-vehicle travel time for the same. Consistent with findings in the social sciences and multiple streams within economics that have shown preferences to be endogenous, the case study showed that a

decision-maker's value of time is sensitive to the level-of-service of the transportation system, and an increase in overall travel times can induce decision-makers to lower their value of time.

The framework was subsequently adapted in Chapter 5 to study the evolution and persistence of modality styles and travel mode choice behavior in a dynamic context. Individual modality styles were still represented as latent classes, but an individual was allowed to have different modality styles at different time periods. The evolutionary path was hypothesized to be a Markov process such that an individual's modality style in the current time period is dependent only on her modality style in the previous time period. As before, travel mode choices for a particular time period were conditioned on the individual's modality style for that time period. The framework was empirically tested using travel diary data collected in Santiago, Chile. The dataset comprises a sample of 220 individuals surveyed over four one-week periods that span a time period of twenty-two months that includes the introduction of Transantiago, a complete redesign of the city's public transit system. Estimation results identified three modality styles: unimodal auto users who only consider the automobile, unimodal transit users who only consider the public transit system and have a low value of time, and multimodal users who consider all travel modes and have a high value of time. The case study found that the distribution of individuals across modality styles is highly sensitive to a shock to the transportation system such as that represented by the introduction of Transantiago. Results from a sample enumeration showed that nearly a quarter of the sample population changed its modality style post-Transantiago.

In each of the three case studies, modality styles were found to be strongly correlated with more long-term travel and activity decisions such as level of vehicle ownership, transit season pass possession, housing type, etc. In examining the influence of individual modality styles on travel mode choice, we took these decisions as exogenous inputs. However, such a causal representation risks endogeneity. Individuals are not auto-oriented because they own more cars or multimodal because they possess a transit season pass. Rather, auto-oriented individuals are more likely to own more cars, and multimodal individuals to possess a transit season pass. In Chapter 6, we reversed this causal representation to reflect the influence of modality styles on these additional dimensions of individual and household travel and activity behavior within a static context. In doing so, we recognized that many of these more long-term decisions, such as level of vehicle ownership and residential location, are not made by individuals in isolation from other members of their household. An individual's preferences and choices are strongly shaped by the opinions and behaviors of the people around her, particularly when a choice is made collectively by a group of individuals, as in the case of a household. Therefore, interaction between household members must be understood to influence attitudes and beliefs towards, among other things, individual travel and activity behavior, or modality styles. To reflect this influence, we introduced the household modality styles construct, characterized by the modality styles of the respective individuals that make up the household.

Chapter 6 developed a model framework that examined the relationship between household modality styles, level of vehicle ownership, transit season pass possession and travel mode choice behavior using travel diary data from Karlsruhe, Germany. The dataset comprised a sample of 96 male and female household heads belonging to 48 households surveyed over a six-week observation period. Estimation results identified four distinct modality styles. The model uncovered both significant correlation between modal preferences of heads of the same household and notable differences as well. In general, female household heads were found to be

less reliant on the automobile for their mobility requirements than their male counterparts. Short-term individual decisions, such as mode choice, were found to be inextricably linked with more long-term individual and household decisions, namely level of vehicle ownership and transit season pass possession, both of which varied considerably across different modality styles.

Modality styles hold important implications for transportation policy and practice. Travel demand models constitute an important component of the planning and policy-making process, being widely used to make forecasts, which in turn are driven by the assumptions that these models make about how individuals arrive at decisions. The model framework developed by this dissertation offers the potential to enhance our understanding of individual and household travel and activity behavior. A greater comprehension of the many factors that shape behavior is essential to the successful design of systems that serve the immediate needs of the population while satisfying long-term societal objectives. For example, findings from this dissertation reveal that models of travel mode choice behavior that ignore the influence of modality styles can overestimate expected gains from transport policies and infrastructural initiatives seeking to reduce automobile use by factors of between one-and-a-half and three. The dissertation further demonstrates how incremental improvements in the transportation system, unless accompanied by corresponding shifts in the distribution of individuals across different modality styles, will result in far smaller changes in travel behavior than would be predicted by a traditional model of travel mode choice. This dissertation makes the case that what is needed is a dramatic change to the transportation system that forces individuals to reconsider how they travel.

## 7.2 Directions for Future Research

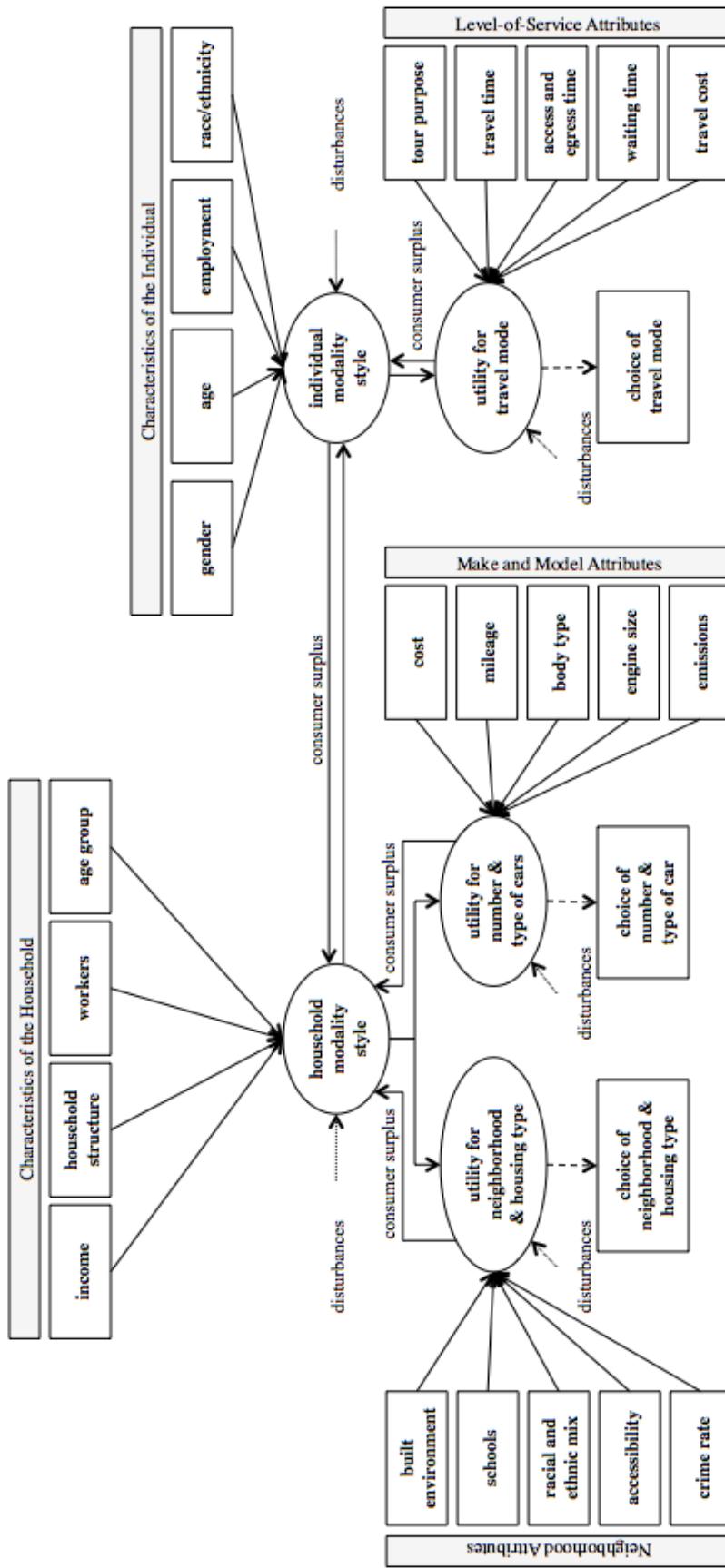
Apart from Chapter 6, where we looked at the influence of modality styles on the level of vehicle ownership and transit season pass possession, our attention thus far has centered on travel mode choice behavior. However, the influence of individual and household modality styles is expected to extend beyond travel mode choice behavior and include all dimensions of individual and household travel and activity behavior, as shown by the schematic in Figure 7.1. In the short-term, these would include dimensions such as destination choice, activity chaining and travel time choice. A habitual auto user may have a different perception of space, travel times and the activity chaining options than, for example, a habitual transit user. Moreover, the auto user may find it difficult to think of activities in ways that would be required if using a bicycle or transit, and may perceive some of the “inconveniences” associated with these modes (e.g., timing activities to transit departures, using a bike with luggage) as obstacles that drive the choice probability of those alternatives close to zero a priori. On the other hand, the regular transit user may build an activity schedule around these constraints such that they are of no or little inconvenience. From the point of view of travel demand modeling, if these interdependencies are to be introduced, it may no longer be permissible to assume that an individual chooses a mode, travel time or destination every time a trip is made, or that the full choice set is considered, but rather that habits and perceptions based on past behavior (in summary, the modality style) may influence observed behavior.

As evidenced by findings from Chapters 3, 5 and 6, the influence of modality styles extends to more medium-term decisions as well, such as level of vehicle ownership and transit season pass possession. In Chapters 3 and 5, we took these medium-term decisions as exogenous inputs to

the model framework and found them to exert significant influence on individual modality styles. However, such a causal representation risks endogeneity. Individuals are not auto-oriented because they own more cars or multimodal because they possess a transit season pass. Rather, auto-oriented individuals are more likely to own more cars, and multimodal individuals to possess a transit season pass. In Chapter 6, we reversed the causal direction to reflect the influence of modality styles on these additional dimensions of individual and household behavior. Results indicate that both level of vehicle ownership and transit season pass possession vary considerably across different modality styles.

Different modality styles are ultimately expected to manifest themselves through their effect on more long-term decisions, such as where to live. Unfortunately, extensive land use indicators were not included within the model frameworks developed and applied in Chapters 2 through 6. However, several studies that have examined the relationship between the built environment and travel behavior have uncovered the existence of higher-level orientations that contemporaneously influence attitudes towards different forms of the built environment and different dimensions of individual travel behavior. It has been argued that differences in modal choices might be a reflection of differing residential choices, and that residential self-selection might be at work (for recent reviews of the literature examining the relationship between the built environment and travel behavior, the reader is referred to Mokhtarian and Cao, 2007; and Ewing and Cervero, 2010). For instance, individuals predisposed towards the automobile are perhaps best served by moving to auto-oriented suburban environments. Similarly, transit-oriented high-density urban developments with mixed land use probably hold a greater draw for individuals with modality styles that lean towards alternative modes of travel, such as transit, bicycling or walking. Therefore, any exhaustive model of individual and household travel and activity behavior must recognize the influence of modality styles on residential location as well.

In an attempt to capture the influence of modality styles on more medium and long-term dimensions of individual and household travel and activity behavior, labeled mobility decisions in Figure 7.1, we propose the framework shown in Figure 7.2 as a possible solution. The framework captures the influence of individual and household modality styles on three specific dimensions: travel mode choice, vehicle ownership and residential location. As per the framework, a household's modality style is hypothesized to be a function of observable sociodemographic variables such as income and household structure, and the consumer surplus that the household derives from different household modality styles. Residential location and level of car ownership are subsequently conditioned on the household's modality style. Residential location is additionally a function of neighborhood characteristics such as the quality of schools, crime rate, the built environment, etc. Similarly, auto ownership is a function of make and model characteristics of different cars on the market, such as cost, mileage, emissions, etc. The modality styles of the individuals residing in a household are a function not only of the sociodemographic variables specific to the individual, such as gender, age, employment, etc., and the consumer surplus that the individual derives from different individual modality styles, but also of the modality style of the household as a whole. Finally, travel mode choice is hypothesized to be a function of the individual's modality style and attributes of the tour and the modal alternatives, namely purpose, i.e. whether the tour is work or non-work, travel time, access and egress time, waiting and transfer times, and travel cost.



**Figure 7.2:** A comprehensive travel demand model depicting the influence of household and individual modality styles on residential location, car ownership and travel mode choice behavior; other dimensions of individual and household travel and activity behavior, such as destination choice or vehicle miles travelled, can be plugged into the model framework analogously

There are a number of benefits to the model framework over more traditional representations of individual and household travel and activity behavior. First, the horizontal framework stands in stark contrast to vertical representations usually employed by activity-based travel demand models currently in practice, such as the SF-CHAMP model presented earlier in Chapter 1. The horizontal representation serves to emphasize the absence of any hierarchy between different dimensions of individual and household travel and activity behavior and the presence instead of correlation across all dimensions, induced through the modality styles construct. Second, the framework is more nuanced in its representation of the relationship between the built environment and travel behavior. It controls for residential self-selection by allowing households with different modality styles to self-select into neighborhoods that best serve their travel needs. At the same time, it also accounts for the reciprocal influence of the built environment on travel behavior. Changes in the built environment will result in unequal changes in the consumer surplus obtained by different household modality styles through residential location. This will force a change in the distribution of households across different household modality styles and subsequently, the distribution of individuals across different individual modality styles, which in turn will produce changes in existing patterns of travel mode choice behavior. And finally, the model captures the dynamic underlying group decision-making through the interplay between household and individual modality styles. By conditioning individual modality styles on household modality styles, we allow individual preferences to be some reflection of the preferences of the household as a whole. Conversely, by having feedback from individual modality styles to household modality style through the construct of consumer surplus, we allow preferences of the household to be some outcome of a process of negotiation between the individuals that comprise the household and their respective preferences.

Notwithstanding these advantages, there are some significant challenges involved with incorporating these additional dimensions within the proposed model framework. The methodological framework shown in Figure 7.2, like each of the model frameworks presented in earlier chapters, relies on a simplified version of tour-based travel demand models for its representation of travel behavior. Travel mode choices are conditioned on other short-term dimensions such as activity chaining, time of day and destination choice. In an earlier paragraph, we argued that the influence of modality styles is expected to extend to each of these additional short-term dimensions. But how to represent that influence within the travel demand model framework proposed in Figure 7.2? Analogous to how residential location and auto ownership are currently represented in the framework, the analyst could condition each of these dimensions directly on individual modality styles and stay consistent with the horizontal representation of individual and household travel and activity behavior. However, it is unclear how to forecast with a travel demand model framework that assumes no hierarchy among different dimensions of travel and activity behavior. Consider the model of residential location and travel mode choice shown in Figure 7.1, neglecting for the moment the sub-model corresponding to auto ownership. In order to predict residential location, one must know a household's likelihood of belonging to a particular modality style, which is some function of the modality styles of each of its constituent members, which in turn is determined in part by the travel mode choices that they face. But the travel mode choices faced by any individual for any home-based tour is obviously some function of where the individual resides. In the absence of a well-defined hierarchy, similar chicken-and-egg problems can be imagined between any pair of choice dimensions. There are no easy solutions to the problem. One way would be to preserve some of the hierarchy assumed by

traditional activity-based travel demand models between different dimensions of travel and activity behavior, and to integrate the modality styles construct within that hierarchy in a way that makes the model behaviorally richer. For example, the SF-CHAMP model conditions travel mode choice on destination choice, which is conditioned on time of day choice, which is conditioned on the full-day tour pattern choice, which could, in our case, be subsequently conditioned on individual modality styles, with logsums feeding upwards from the lower dimensions. And this pattern could be replicated to include both medium-term and long-term decisions and the household modality styles construct. The model framework shown in Figure 7.2 is presented as one possibility, and we leave it to future research to decide what might be the most appropriate framework.

An additional hurdle to incorporating dimensions such as destination choice, auto ownership and residential location within any model framework is that each of these dimensions involves a decision between an intractably large number of alternatives. For example, consider the case of destination choice: the nine-county San Francisco Bay Area comprises 109,223 census blocks, 1,574 census tracts and 1,454 traffic analysis zones. Depending upon the level of aggregation, each of these census blocks, census tracts or traffic analysis zones could potentially serve as an alternative in a model of destination choice. In comparison, the travel mode choice models that we've estimated in previous chapters have comprised fewer than ten alternatives! The exponential increase in the number of alternatives can often make model estimation impractical and the analyst is forced to take recourse to sampling alternatives from the full choice set. McFadden (1978) demonstrates that if the multinomial logit model is an appropriate specification for the choice process being modeled, then the analyst can obtain consistent estimates of the parameters corresponding to the utility function from a fixed or random sample of alternatives from the full choice set. Guevara and Ben-Akiva (2013) develop a methodological framework that extends McFadden's original result to the broader family of Multivariate Extreme Value (MEV) models that includes model forms such as the nested logit and cross-nested logit. Their framework achieves consistency, asymptotic normality and efficiency in the parameter estimates while sampling alternatives for any of these model forms. For conventional LCCMs without feedback, the Expectation-Maximization (EM) algorithm provides a clever way of leveraging these findings to ensure that the parameter estimates are consistent even when sampling alternatives. Under the M-step of the EM algorithm for LCCMs without feedback, each of the class-specific choice models and the class-membership model can be maximized independent of the other sub-models. Therefore, if the class-specific choice models and the class-membership model belong to the family of MEV models, sampling can still guarantee consistent parameter estimates. However, for LCCMs with feedback through consumer surplus, the class-specific choice models and the class-membership model can no longer be separated and maximized independently because the sub-models are joined through the consumer surplus construct. Consequentially, the EM algorithm is no more useful than traditional gradient-based optimization routines and sampling alternatives is no longer a viable option. Ongoing research is exploring ways in which conventional methods of model estimation can be made more efficient through the use of recent advancements in data structures, optimization routines and technological hardware.

### **7.3 Concluding Remarks**

Travel demand and land use models constitute an important component of the planning and policy-making process, being widely used to make forecasts, which in turn are driven by the assumptions that these models make about how individuals arrive at decisions. The field of travel demand and land use analysis has progressed far beyond the aggregate models that defined the profession for much of the last century. Research during the eighties and nineties on travel behavior culminated in the activity-based travel demand model, a disaggregate model of household and individual behavior which formulates the demand for travel as a function of the demand for various activities, such as work, maintenance and leisure, predicting which activities are conducted where, when, for how long, with whom and using what mode of travel. Simultaneous developments in land use analysis resulted in disaggregate agent-based models of business and household behavior in the real estate market with the objective of predicting the distribution and intensity of activities in the urban area, to be used subsequently as inputs by travel demand models. The shift towards disaggregate models of decision-making has been seen as a significant step forward, contributing to the development of a comprehensive framework that recognizes the influence of land use patterns on the demand for transportation systems through the interplay between different dimensions of travel and activity behavior.

However, travel demand and land use models currently in practice lack in one critical way: they continue to be rather rudimentary in their representation of the cognitive processes underlying the formation and persistence of choices. Focus has centered on more tangible predictors of behavior, such as travel times and costs in the context of travel mode choice, at the expense of more fundamental behavioral constructs such as habits, attitudes, values, norms and affects, reflective of more profound individual variations in modality styles, that have been shown by a rich body of work in the shared fields of psychology, sociology and behavioral economics to have a significant effect on decision-making. As mentioned previously, findings from this dissertation reveal that models of travel mode choice behavior that ignore the influence of modality styles can overestimate expected gains from transport policies and infrastructural initiatives seeking to reduce automobile use by factors of between one-and-a-half and three. Ongoing research is exploring ways in which the model framework developed in this dissertation may be extended to include additional choice dimensions. The framework shown in Figure 7.2 presents one such exciting possibility. When complete, the line of work initiated by this dissertation is expected to result in a comprehensive model of individual and household travel and activity behavior that integrates travel demand and land use analysis through the modality styles construct with the objective of offering a deeper understanding of decision-making and greater predictive power than current models in practice.

## Bibliography

- Aarts, H., Verplanken, B., van Knippenberg, A. (1997), “**Habit and information use in travel mode choices**,” *Acta Psychologica*, Vol. 96, pp. 1–14.
- Adamowicz, W., Hanemann, M., Swait, J., Johnson, R., Layton, D., Regenwetter, M., Reimer, T., and Sorkin, R. (2005), “**Decision strategy and structure in households: A ‘Groups’ perspective**,” *Marketing Letters*, Vol. 16, Nos. 3-4, pp. 387-399.
- Alderman, H., Chiappori, P. A., Haddad, L., Hoddinott, J., and Kanbur, R. (1995), “**Unitary versus collective models of the household: Is it time to shift the burden of proof?**” *The World Bank Research Observer*, Vol. 10, No. 1, pp. 1-19.
- Arentze, T. A., and Timmermans, H. J. P. (2004), “**A learning-based transportation oriented simulation system**,” *Transportation Research Part B: Methodological*, Vol. 38, No. 7, pp. 613-633.
- Aribarg, A., Arora, N., and Bodur, H. O. (2002), “**Understanding the role of preference revision and concession in group decisions**,” *Journal of Marketing Research*, Vol. 39, No. 3, pp. 336-349.
- Aribarg, A., Arora, N., and Kang, M. Y. (2010), “**Predicting joint choice using individual data**,” *Marketing Science*, Vol. 29, No. 1, pp. 139-157.
- Arora, N., and Allenby, G. M. (1999), “**Measuring the influence of individual preference structures in group decision making**,” *Journal of Marketing Research*, Vol. 36, No. 4, pp. 476-487.
- Atasoy, B., Glerum, A., and Bierlaire, M. (2011), “**Mode choice with attitudinal latent class: A Swiss case-study**,” Presented at the *2<sup>nd</sup> International Choice Modeling Conference 2011*, Leeds, United Kingdom.
- Axhausen, K.W., A. Simma, and Golob, T. (2001), “**Pre-commitment and usage: cars, season tickets and travel**,” *European Research in Regional Science*, Vol. 11, pp. 101-110.
- Axhausen, K.W., Zimmermann, A., Schönfelder, S., Rindsfüser, G., and Haupt, T. (2002), “**Observing the rhythms of daily life: A six-week travel diary**,” *Transportation*, Vol. 29, pp. 95-124.
- Bargh, J. A., and Chartrand, T. L. (1999), “**The unbearable automaticity of being**,” *American Psychologist*, Vol. 54, No. 7, pp. 462-479.
- Ben-Akiva, M. E. (1973), “**Structure of passenger travel demand models**,” *PhD dissertation*, Massachusetts Institute of Technology.

Ben-Akiva, M., D. Bolduc and M. Bradley (1993), “**Estimation of Travel Choice Models with Randomly Distributed Values of Time**,” *Transportation Research Record: Journal of the Transportation Research Board*, Washington, D.C., Vol. 1413, pp. 88-97.

Ben-Akiva, M., and Boccara, B. (1995), “**Discrete choice models with latent choice sets**,” *International Journal of Research in Marketing*, Vol. 12, pp. 9-24.

Ben-Akiva, M., and Abou-Zeid, M. (2007), “**Hybrid choice models: from static to dynamic**,” Paper presented at the proceedings of the *Oslo Workshop on Valuation Methods in Transport Planning*, Oslo, Norway, 19–20 March 2007.

Ben-Elia, E., and Ettema, D. (2009), “**Carrots versus sticks: Rewarding commuters for avoiding the rush-hour— a study of willingness to participate**,” *Transport Policy*, Vol. 16, No. 2, pp. 68-76.

Bhat, C. R. (2000), “**Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling**,” *Transportation Science*, Vol. 34, No. 2, pp. 228-238.

Bhat C. R., and Guo J (2006), “**A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels**,” In: *Presented at the 85th Annual Meeting of the Transportation Research Board*, Washington DC.

Bhat, C. R., and Pendyala, R. M. (2005), “**Modeling intra-household interactions and group decision-making**,” *Transportation*, Vol. 32, No. 5, pp. 443-448.

Bierlaire, M. (2003), “**BIOGEME: A free package for the estimation of discrete choice models**,” *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland.

Bowles, S. (1998), “**Endogenous preferences: The cultural consequences of markets and other economic institutions**,” *Journal of Economic Literature*, Vol. 36, No. 1, pp. 75-111.

Bowman, J. L. (1998), “**The day activity schedule approach to travel demand analysis**,” *PhD dissertation*, Massachusetts Institute of Technology.

Bradley, M., Bowman, J. L., and Giesenbeck, B. (2010), “**SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution**,” *Journal of Choice Modelling*, Vol. 3, No. 1, pp. 5-31.

Brown, S., Harris, M. N., and Taylor, K. (2010), “**Modelling charitable donations: A latent class panel approach**,” *Working Paper*, Department of Economics, University of Sheffield ISSN 1749-8368

Bureau of Transportation Statistics (2012), “**National Transportation Statistics**,” Table 4-23 Average Fuel Efficiency of U.S. Passenger Cars and Light Trucks (Updated January 2012).

Cambridge Systematics (2002), “**San Francisco Travel Demand Forecasting Model Development: Executive Summary**,” prepared for San Francisco County Transportation Authority.

Cantillo, V., and Ortúzar, J. de D. (2005), “**A semi-compensatory discrete choice model with explicit attribute thresholds of perception**,” *Transportation Research Part B: Methodological*, Vol. 39, No. 7, pp. 641-657.

Cantillo, V., Ortúzar, J. de D., and Williams, H. C. W. L. (2007), “**Modeling discrete choices in the presence of inertia and serial correlation**,” *Transportation Science*, Vol. 41, No. 2, pp. 195-205.

Carrasco, J. A., and Ortúzar, J. de D. (2002), “**Review and assessment of the nested logit model**,” *Transport Reviews: A Transnational Transdisciplinary Journal*, Vol. 22, No. 2, pp. 197-218.

Cascetta, E., and A. Papola (2001), “**Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand**,” *Transportation Research Part C: Emerging Technologies*, Vol. 9, No. 4, pp. 249-263.

Chamberlain, G. (1980), “**Analysis of Covariance with Qualitative Data**,” *The Review of Economic Studies*, Vol. 47, pp. 225–238.

Chandrasekharan, B., and Goulias, K. G. (1999), “**Exploratory longitudinal analysis of solo and joint trip making using the Puget Sound transportation panel**,” In *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1676, TRB, National Research Council, Washington, D.C., pp. 77–85.

Choo, S. and Mokhtarian, P. (2004), “**What type of vehicle do people drive? The role of attitude and lifestyle in influencing vehicle type choice**,” *Transportation Research Part A: Policy and Practice*, Vol. 38, No. 3, 201-222.

Choudhury, C. F., Ben-Akiva, M., and Abou-Zeid, M. (2010), “**Dynamic latent plan models**,” *Journal of Choice Modelling*, Vol. 3, No. 2, pp. 50-70.

Chung, H., Flaherty, B. P., and Schafer, J. L. (2006), “**Latent class logistic regression: application to marijuana use and attitudes among high school seniors**,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, Vol. 169, No. 4, pp. 723-743.

Cirillo, C., and Axhausen, K.W. (2006), “**Evidence on the distribution of values of travel time savings from a six-week diary**,” *Transportation Research Part A: Policy and Practice*, Vol. 50, No. 5, pp. 444-457.

Corfman, K. P., Lehmann, D. R. (1987), “**Models of cooperative group decision-making and relative influence: An experimental investigation of family purchase decisions**,” *Journal of Consumer Research*, Vol. 14, No. 1, pp. 1–13.

Corfman, K. P. (1991), “**Perceptions of relative influence: Formation and measurement,**” *Journal of Marketing Research*, Vol. 28, No. 2, pp. 125-136.

Davis, J. H. (1973), “**Group decision and social interaction: A theory of social decision schemes,**” *Psychological Review*, Vol. 80, pp. 97-125.

Davis, H.L., (1976), “**Decision making within the household,**” *Journal of Consumer Research*, Vol. 2, No. 4, pp. 241-260.

Davis, S. C., Diegel, S. W., and Boundy, R. G. (2011), “**Transportation energy data book: 30th Edition,**” Office of Energy Efficiency and Renewable Energy U.S. Department of Energy.

Eluru, N., Pinjari, A. R., Pendyala, R. M., and Bhat, C. R. (2010a), “**An econometric multi-dimensional choice model of activity-travel behavior,**” *Transportation Letters: The International Journal of Transportation Research*, Vol. 2, pp. 217-230.

Eluru, N., Bhat, C. R., Pendyala, R. M., and Konduri, K. C. (2010b), “**A joint flexible econometric model system of household residential location and vehicle fleet composition/usage choices,**” *Transportation*, Vol. 37, pp. 603–626.

Federal Highway Administration, (2011), “**National Household Travel Survey 2009,**” United States Department of Transportation.

Fernández, J. E., de Cea Ch., J., Malbran, R. H. (2008), “**Demand responsive urban public transport system design: Methodology and application,**” *Transportation Research Part A: Policy and Practice*, Vol. 42, No. 7, pp. 951–972.

Flyvbjerg, B., Holm, M. K. S., and Buhl, S. L. (2007), “**How (In)accurate Are Demand Forecasts in Public Works Projects?: The Case of Transportation,**” *Journal of the American Planning Association*, Vol. 71, No. 2, pp. 131-146.

Fosgerau, M. (2005), “**Investigating the distribution of the value of travel time savings,**” *Transportation Research Part B: Methodological*, Vol.40, No. 8, pp. 688-707.

Gärling, T., and Axhausen, K. W. (2003), “**Introduction: Habitual travel choice,**” *Transportation*, Vol. 30, No. 1, pp. 1-11.

Gaudry, M., and M. Dagenais (1979), “**The Dogit Model,**” *Transportation Research Part B: Methodological*, Vol. 13, No. 2, pp. 105-111.

Giddens, A. (1991), “**Modernity and self-identity: Self and society in the late modern age,**” Cambridge, UK: Polity Press.

Gliebe, J., and F. Koppelman (2002), “**A model of joint activity participation between household members,**” *Transportation*, Vol. 29, No. 1, pp. 49–72.

Gliebe, J., and F. Koppelman (2005), “**Modeling household activity-travel interactions as parallel constrained choices,**” *Transportation*, Vol. 32, No. 5, pp. 449-471.

Golob, T. F. (1996), “**A model of household demand for activity participation and mobility**,” UCI-ITS-WP-96-5, Institute of Transportation Studies, University of California at Irvine.

Golob, T. F., and McNally, M. (1997), “**A model of activity participation and travel interactions between household heads**,” *Transportation Research Part B: Methodological*, Vol. 31, No. 3, pp. 177–194.

Gomes, N., Merugu, D., O'Brien, G., Mandyam, C., Yue, T., Atikoglu, B., Albert, A., Fukumoto, N., Liu, H., Prabhakar, B., and Wischik, D. (2012), “**Steptacular: An incentive mechanism for promoting wellness**,” *NetHealth, Comsnets Workshop on Networked Healthcare Technology*, January 2012.

Gopinath, D. A. (1995), “**Modeling heterogeneity in discrete choice processes: application to travel demand**,” *PhD dissertation*, Massachusetts Institute of Technology.

Greene, W. H., Hensher, D. A., and Rose, J. (2006), “**Accounting for heterogeneity in the variance of unobserved effects in mixed logit models**,” *Transportation Research Part B: Methodological*, Vol. 40, No. 1, pp. 75-92.

Greene, W. H., and Hensher, D. A. (2013), “**Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model**,” *Applied Economics*, Vol. 45, No. 14, pp. 1897-1902.

Grisolía, J. M., and Willis, K. G. (2012), “**A latent class model of theatre demand**,” *Journal of Cultural Economics*, Vol. 36, No. 2, pp. 113-139.

Guevara, C. A., and Ben-Akiva, M. E. (2013), “**Sampling of alternatives in Multivariate Extreme Value (MEV) models**,” *Transportation Research Part B: Methodological*, Vol. 48, pp. 31-52.

Hensher, D.A., and Greene, W.H. (2003), “**Mixed logit models: state of practice**,” *Transportation*, Vol. 30, No. 2, pp. 133–176.

Hensher, D. A., Rose, J. M., Black, I. (2008), “**Interactive agency choice in automobile purchase decisions: The role of negotiation in determining equilibrium choice outcomes**,” *Journal of Transport Economics and Policy*, Vol. 42, No. 2, pp. 269–296.

Hess, S. (2005), “**Advanced discrete choice models with applications to transport demand**,” *PhD dissertation*, Imperial College London.

Hess, S., M. Bierlaire and J. W. Polak (2005), “**Estimation of value of travel-time savings using mixed logit models**,” *Transportation Research Part A: Policy and Practice*, Vol. 39, Nos. 2-3, pp. 221–236.

Hess, S., and Rose, J. M. (2006), “**Effects of distributional assumptions on conditional estimates from mixed logit models**,” *ITLS working paper*, Institute for Transport and Logistics Studies, University of Sydney, Sydney, Australia.

Hess, S., Ben-Akiva, M., Gopinath, D., and Walker, J. (2009), “**Taste heterogeneity, correlation and elasticities in latent class choice models**,” *Compendium of Papers, 88th Annual Meeting of the Transportation Research Board*. Washington, DC: Transportation Research Board.

Hidrue, M. K., Parsons, G. R., Kempton, W., and Gardner, M. P. (2011), “**Willingness to pay for electric vehicles and their attributes**,” *Resource and Energy Economics*, Vol. 33, No. 3, pp. 686-705.

Hirschman, A. O. (1982), “**Rival interpretations of market society: Civilizing, destructive, or feeble?**” *Journal of Economic Literature*, Vol. 20, No. 4, pp. 1463-1484.

Hollis, M. (1987), “**The cunning of reason**,” Cambridge University Press, Cambridge.

Johansson, M. V., Heldt, T., and Johansson, P. (2006), “**The effects of attitudes and personality traits on mode choice**,” *Transportation Research Part A: Policy and Practice*, Vol. 40, No. 6, pp. 507–525.

Kamakura, W. A., and Russell, G. J. (1989), “**A probabilistic choice model for market segmentation and elasticity structure**,” *Journal of Marketing Research*, Vol. 26, No. 3, pp. 379–390.

Kato, H., and Matsumoto, M. (2009), “**Intra-household interaction in a nuclear family: A utility-maximizing approach**,” *Transportation Research Part B: Methodological*, Vol. 43, No. 2, pp. 191-203.

Kitamura, R. (1988), “**Life-style and travel demand**,” *Transportation Research Board Special Report 220*, pp. 149-189. Washington, DC: Transportation Research Board.

Kitamura, R., Mokhtarian, P.L. and Laidet, L. (1997), “**A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area**,” *Transportation*, Vol. 24, pp. 125-158.

Krizek, K., and Waddell, P. (2002), “**Analysis of lifestyle choices: Neighborhood type, travel patterns and activity participation**,” *Transportation Record: Journal of the Transportation Research Board*, Washington, D.C., Vol. 1807, pp. 119-128.

Kuhnimhof, T., Chlond, B. and von der Ruhren, S. (2006), “**The Users of Transport Modes And Multimodal Travel Behavior – Steps Towards Understanding Travelers’ Options and Choices**,” *Compendium of Papers, 85th Annual Meeting of the Transportation Research Board*. Washington, DC: Transportation Research Board.

Kuhnimhof, T. (2009), “**Measuring and modeling multimodal mode use in the longitudinal section**,” *Compendium of Papers, 88th Annual Meeting of the Transportation Research Board*. Washington, DC: Transportation Research Board.

Lanzendorf, M. (2002), “**Mobility styles and travel behavior: Application of a lifestyle approach to leisure travel**,” *Transportation Record*, Vol. 1807, pp. 163-173.

Lee, C. K. C., and Beatty, S. E. (2002), “**Family structure and influence in family decision making**,” *Journal of Consumer Marketing*, Vol. 19, No. 1, pp. 24-41.

Levinson, D. M., and Gillen, D. (1998), “**The full cost of intercity highway transportation**,” *Transportation Research Part D: Transport and Environment*, Vol. 3, No. 4, pp. 207-223.

Lundberg, S. J., Pollak, R. A., and Wales, T. J. (1997), “**Do husbands and wives pool their resources? Evidence from the United Kingdom child benefit**,” *The Journal of Human Resources*, Vol. 32, No. 3, pp. 463-480.

Manski, C. (1977), “**The structure of random utility models**,” *Theory and Decision*, Vol. 8, No. 3, pp. 229–254.

McFadden, D. (1978), “**Modeling the choice of residential location**,” In: Karlquist, Lundqvist, Snickers, Weibull (Eds.), *Spatial Interaction Theory and Residential Location*, North Holland, Amsterdam, pp. 75–96.

McFadden, D. (2001), “**Economic choices**,” *The American Economic Review*, Vol. 91, No. 3, pp. 351-378.

McFadden, D., Train, K. (2000), “**Mixed MNL models for discrete response**,” *Journal of Applied Econometrics*, Vol. 15, pp. 447-470.

Menasco, M. B., and Curry, D. J. (1989), “**Utility and choice: An empirical study of wife/husband decision making**,” *Journal of Consumer Research*, Vol. 16, No. 1, pp. 87-97.

Meister, K., Frick, M., and Axhausen, K. W. (2005), “**A GA-based household scheduler**,” *Transportation*, Vol. 32, No. 5, pp. 473-494.

Merugu, D., Prabhakar, B., and Rama, N. (2009), “**An incentive mechanism for decongesting the roads: a pilot program in Bangalore**,” *Proceedings of NetEcon '09, ACM Workshop on the Economics of Networked Systems*, July 2009.

Meyer, M. D. (1997), “**Demand management as an element of transportation policy: using carrots and sticks to influence travel behavior**,” *Transportation Research Part A: Policy and Practice*, Vol. 33, Nos. 7-8, pp. 575-599.

Morpace International, Inc. (2002), “**Bay Area Travel Survey 2000: Final Report**.”

Munro, A. (2009), “**Introduction to the special issue: Things we do and don't understand about the household and the environment**,” *Environmental and Resource Economics*, Vol. 43, No. 1, pp. 1-10.

National Center for Transit Research (2008), “**Transit ridership, reliability, and retention**,” State of Florida Department of Transportation.

Olaru, D., Smith, B., and Taplin, J. H. E. (2011), “**Residential location and transit-oriented development in a new rail corridor**,” *Transportation Research Part A: Policy and Practice*, Vol. 45, No. 3, pp. 219-237.

Parry, I. W. H., and Small, K. A. (2005), “**Does Britain or the United States Have the Right Gasoline Tax?**” *The American Economic Review*, Vol. 95, No. 4, pp. 1276-1289.

Parthasarathi, P., and Levinson, D. (2010), “**Post-construction evaluation of traffic forecast accuracy**,” *Transport Policy*, Vol. 17, No. 6, pp. 428-443.

Petersen, E., and Vovsha, P. (2006), “**Intrahousehold car-type choice for different travel needs**,” In *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1985, TRB, National Research Council, Washington, D.C., 2001, pp. 207-219.

Pickrell, D. H. (1992), “**A desire named streetcar: Fantasy and fact in rail transit planning**,” *Journal of the American Planning Association*, Vol. 58, No. 2, pp. 158-176.

Pinjari, A. R., Pendyala, R. M., Bhat, C. R., and Waddell, P. A. (2007), “**Modeling residential sorting effects to understand the impact of the built environment on commute mode choice**,” *Transportation*, Vol. 34, pp. 557-573.

Pisarski, A. E. (2006), “**Commuting in America III: The Third National Report on Commuting Patterns and Trends**,” Transportation Research Board, Washington, D.C.

Pollak, R. A. (1978), “**Endogenous tastes in demand and welfare analysis**,” *The American Economic Review*, Vol. 68, No. 2, pp. 374-379.

Puckett, S. M., and Raschiute, S. (2010), “**Freight stakeholders’ sensitivities under road user charging: a latent class approach**,” *Australasian Transport Research Forum 2010 Proceedings*.

Rao, V. R., and Steckel, J. H. (1991), “**A polarization model for describing group preferences**,” *Journal of Consumer Research*, Vol. 18, No. 1, pp. 108-118.

Rescorla, R.A. (1987), “**A Pavlovian analysis of goal-directed behavior**,” *American Psychologist*, Vol. 42, pp. 119-129.

Roorda, M. J., Carrasco, J. A., and Miller, E. J. (2009), “**An integrated model of vehicle transactions, activity scheduling and mode choice**,” *Transportation Research Part B: Methodological*, Vol. 43, No.2, pp. 217-229.

Rose, J., and Hensher, D.A. (2004), “**Modeling agent interdependency in group decision making**,” *Transportation Research Part E: Logistics and Transportation Review*, Vol. 40, No. 1, pp. 63–79.

Rust, J. (1987), “**Optimal replacement of GMC buses: An empirical model of Harold Zurcher**,” *Econometrica*, Vol. 55, No. 5, pp. 999–1033.

Sen, A. (1973), “**Behaviour and the concept of preference**,” *Economica*, Vol. 40, No. 159, pp. 241-259.

Siegers, P. (2010), “**Mapping religious orientations across Europe: Church religiosity, alternative spiritualities, and unbelief. Evidence from the fourth wave of the European Values Study (2008/2009)**,” <http://www.ecprnet.eu/databases/conferences/papers/121.pdf>.

Simma, A. and Axhausen, K.W. (2003), “**Commitments and modal usage: An analysis of German and Dutch panels**,” *Transportation Research Record*, Vol. 1854, pp. 22-31.

Smith, E.E., Nolen-Hoeksema, S., Fredrickson, B., Loftus, G. (2003), “**Atkinson and Hilgard's Introduction to Psychology**,” Wadsworth/Thomson Learning, Belmont, CA.

Sobel, M. (1981), “**Lifestyle and social structure: Concepts, definitions, analyses**,” New York: Academic Press.

Srinivasan, S., and Bhat, C. R. (2005), “**Modeling household interactions in daily in-home and out-of-home maintenance activity participation**,” *Transportation*, Vol. 32, No. 5, pp. 523-544.

Srinivasan, S. and Ferreira, J. (2002), “**Travel behavior at the household level: understanding linkages with residential choice**,” *Transportation Research Part D: Transport and Environment*, Vol. 7, No. 3 , pp. 225-242.

Swait, J., and Ben-Akiva, M. (1986), “**Empirical test of a constrained discrete-choice model**,” *Behavioral Research for Transport Policy*, VNU Science Press, Utrecht, The Netherlands.

Swait, J., and Ben-Akiva, M. (1987), “**Incorporating random constraints in discrete models of choice set generation**,” *Transportation Research Part B: Methodological*, Vol. 21, No. 2, pp. 91-102.

Swait, J. (2009), “**Choice models based on mixed discrete/continuous PDFs**,” *Transportation Research Part B: Methodological*, Vol. 43, No. 7, pp. 766-783.

Tecihert, T., Shehu, E., and von Wartburg, I. (2008), “**Customer segmentation revisited: The case of the airline industry**,” *Transportation Research Part A: Policy and Practice*, Vol. 42, No. 1, pp. 227-242.

Thøgersen, J. (2006), “**Understanding repetitive travel mode choices in a stable context: A panel study approach**,” *Transportation Research Part A: Policy and Practice*, Vol. 40, No. 8, pp. 621- 638.

Thorndike, R.L. (1938), “**On what type of task will a group do well?**,” *Journal of Abnormal and Social Psychology*, Vol. 33, No. 3, pp. 409–413.

Timmermans, H. J. P., and Zhang, J. (2009), “**Modeling household activity travel behavior: Examples of state of the art modeling approaches and research agenda**,” *Transportation Research Part B: Methodological*, Vol. 43, No. 2, pp. 187-190.

Townsend, T. A. (1987), “**The effects of household characteristics on the multi-day time allocations and travel-activity patterns of households and their members**,” *Unpublished Ph.D. Dissertation*, Northwestern University, Evanston, IL.

Train, K. E. (2009), “**Discrete choice models with simulation**,” Cambridge University Press, Cambridge.

Transport for London (2004), “**Impacts monitoring – Second annual report**,” Mayor of London.

Tversky, A., and Thaler, R. H. (1990), “**Anomalies: Preference reversals**,” *The Journal of Economic Perspectives*, Vol. 4, No. 2, pp. 201-211.

van Wissen, L. (1989), “**A model of household interactions in activity patterns**,” Presented at the *International Conference on Dynamic Travel Behavior Analysis*, Kyoto University, Kyoto, Japan.

Vermeulen, F. (2002), “**Collective household models: Principles and main results**,” *Journal of Economic Surveys*, Vol. 16, No. 4, pp. 533-564.

Vermunt, J. K., and J. Magidson (2005), “**Technical Guide for Latent GOLD Choice 4.0: Basic and Advanced**,” Belmont Massachusetts: Statistical Innovations Inc.

Verplanken, B., Walker, I., Davis, A., and Jurasek, M. (2008), “**Context change and travel mode choice: Combining the habit discontinuity and self-activation hypotheses**,” *Journal of Environmental Psychology*, Vol. 28, pp. 121-171.

Vij, A., Carrel, A., and Walker, J. L. (2011), “**Latent modal preferences: Behavioral mixture models with longitudinal data**,” Presented at the *2<sup>nd</sup> International Choice Modeling Conference 2011*, Leeds, UK.

Vovsha P, Petersen, E., and Donnelly, R. (2003), “**Explicit modeling of joint travel by household members: statistical evidence and applied approach**,” In *Transportation Research Record: Journal of the Transportation Research Board*, Washington, D.C., Vol. 1831, pp. 1-10.

Vovsha, P., and Petersen, E. (2009), “**Model for person and household mobility attributes**,” *Transportation Research Record: Journal of the Transportation Research Board*, Washington, D.C., Vol. 2132, pp. 95-105.

Walker, J.L., and Ben-Akiva, M. (2011) “**Advances in discrete choice: mixture models**,” In: de Palma, A., Lindsey, R., Quinet, E., and Vickerman, R. (eds) *A Handbook Of Transport Economics*. Edward Elgar Publishing.

Walker, J., and Li, J. (2007), “**Latent lifestyle preferences and household location decisions**,” *Journal of Geographical Systems*, Vol. 9, No. 1, pp. 77-101.

Wang, D., and Li, J. (2009), “**A model of household time allocation taking into consideration of hiring domestic helpers**,” *Transportation Research Part B: Methodological*, Vol. 43, No.2, pp. 204-216.

Weizsäcker, C. C. von (1971), “**Notes on endogenous change of tastes**,” *Journal of Economic Theory*, Vol. 3, No. pp. 345-372.

Welch, L. R. (2003), “**Hidden Markov Models and the Baum-Welch Algorithm**,” *IEEE Information Theory Society Newsletter*, Vol. 53, No. 4., pp. 1, 10-13.

Wen, C.-H., and Lai, S.-C. (2010), “**Latent class models of international air carrier choice**,” *Transportation Research Part E: Logistics and Transportation Review*, Vol. 46, No. 2, pp. 211-221.

Wirthlin Worldwide, and FJCandN (2000). “**Enhancing the visibility and image of transit in the United States and Canada**,” *TCRP Report 63*, Transit Cooperative Research Program, Transportation Research Board, Washington, D.C.

Wollen, P., and Kerr, J., editors (2002), “**Autopia: Cars and Culture**,” London: Reaktion.

Wood, W., Quinn, J.M., and Kashy, D.A. (2002), “**Habits in everyday life: Thought, emotion, and action**,” *Journal of Personality and Social Psychology*, Vol. 83, No.6, pp. 1281–1297.

Wright, L., and K. Fjellström (2005), “**Mass Transit Options**,” *Sustainable transport: A sourcebook for policy-makers in developing cities*.

Yáñez, M. F., Cherchi, E., Ortúzar, J. de D., and Heydecker, B. G. (2009), “**Inertia and shock effects on mode choice panel data: Implications of the Transantiago implementation**,” Presented at the *12<sup>th</sup> International Conference on Travel Behavior Research*, Jaipur, India.

Yáñez, M. F., Mansilla, P., and Ortúzar, J. de D. (2010), “The Santiago Panel: Measuring the effects of implementing Transantiago,” *Transportation*, Vol. 37, No. 1, pp. 125-149.

Zhang, J., Kuwano, M., Lee, B., and Fujiwara, A. (2009), “**Modeling household discrete choice behavior incorporating heterogeneous group decision-making mechanisms**,” *Transportation Research Part B: Methodological*, Vol. 43, No. 2, pp. 230-250.

Zhao, J. (2009), “**Preference accommodating and preference shaping: Incorporating traveler preferences into transportation planning**,” *Ph.D. Dissertation*, Massachusetts Institute of Technology.