The Dissertation Committee for Aruna Sivakumar certifies that this is the approved version of the following dissertation:

TOWARD A COMPREHENSIVE, UNIFIED, FRAMEWORK FOR ANALYZING SPATIAL LOCATION CHOICE

Committee:

Chandra R. Bhat, Supervisor

Randy B. Machemehl

Vijay Mahajan

Michael Oden

C. Michael Walton

TOWARD A COMPREHENSIVE, UNIFIED, FRAMEWORK FOR ANALYZING SPATIAL LOCATION CHOICE

by

Aruna Sivakumar, B.Tech.; M.S.E.

Dissertation

Presented to the Faculty of the Graduate School of

the University of Texas at Austin

in Partial Fulfillment

of the requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2005

To my dearest appudu

ACKNOWLEDGEMENTS

It may be trite to use the phrase ‘friend, philosopher and guide’ but nothing better expresses what my supervisor, Professor Chandra Bhat, has been to me. He is my source of inspiration. He has helped me grow both professionally and personally; supported me in times of need; motivated me to do my best always. Many are the philosophical discussions we have shared and, I hope, will continue to share for the rest of our lives. My respect, admiration, deepest love and sincere thanks to you, Dr. Bhat.

My parents – the source of my faith in life, my pillars of support. They have always encouraged me to explore and expand my interests. They have taught me to be strong, self-sufficient, independent and considerate. I can never express my love for them in words, but I shall always have them in my heart. Everything I achieve, I achieve for them.

I could not have completed this work without the love and unending support of my fiancé, Kaushik, and my dearest friends, especially Cherian and Ratheesh. They cheered me up at my lowest, and cheered me on through the highest points of my life as a PhD student. I must also acknowledge my colleagues at UT Austin, specifically Siva and Jessica - I learnt a great deal from them. This list would be incomplete if I did not include Lisa Weyant. Lisa has been a constant help in all matters, and a good friend.

I must also thank Dr. Kay Axhausen and Dr. Stefan Schönfelder for their valuable inputs on my empirical application and the Mobidrive data. Thanks are due also to Dr. Giray Ökten, for his valuable guidance and feedback on the generation and comparison of the quasi-Monte Carlo sequences.

Last, but definitely not the least, I would like to thank my committee – Dr. Walton, Dr. Machemehl, Dr. Mahajan and Dr. Oden. They were very patient through all the scheduling conflicts. Their feedback during my defense was very stimulating and a good learning experience.

TOWARD A COMPREHENSIVE, UNIFIED, FRAMEWORK

FOR ANALYZING SPATIAL LOCATION CHOICE

Publication No. ____________

Aruna Sivakumar, Ph.D.
The University of Texas at Austin, 2005

Supervisor: Chandra R. Bhat

In today's world of increasing congestion and insufficient scope for infrastructural expansions, urban and transportation planners rely on the accuracy and behavioral realism of travel demand models to make informed policy decisions. The development of accurate and behaviorally realistic travel demand models requires a good understanding of individual travel behavior, and an important step toward this has been the development of the activity-based paradigm, which states that travel is a result of the desire to participate in activities at spatially scattered locations. Activity-based travel demand modeling systems essentially model the activity-travel patterns of individuals, which are characterized by several attributes such as activity purpose, location of activity participation and choice of mode. Of all these attributes, the choice of location of activity

participation is one that has received relatively inadequate attention in the literature. On the other hand, the location of activity participation spatially pegs the daily activity-travel patterns of individuals. Accurate predictions of activity location are, therefore, key to effective travel demand management and air quality control strategies. Moreover, an understanding of the factors that influence the choice of location can contribute to more effective land-use and zoning policies.

The broad objectives of this dissertation research are two-fold. The first objective is to develop a comprehensive econometric model of location choice for non-work activities that incorporates accuracy and behavioral realism in capturing different choice behaviors. This was achieved through the comprehensive introduction of heterogeneity in choice behavior, including observed and unobserved sources of inter- and intra-personal heterogeneity, spatial correlation, variety seeking and loyalty/inertial behavior, and spatial cognition. The estimation of such a flexible model typically requires the use of simulated maximum likelihood estimation (SMLE). The second broad objective of this research is to contribute toward improving the efficiency of the SMLE by comparing the performance of various quasi-Monte Carlo (QMC) sequences and their scrambled versions. Numerical experiments were designed and the Random Linear and Random Digit Scrambled Faure sequences are identified as the most efficient. Finally, all these research efforts contribute to the empirical estimation of non-maintenance shopping location choice models using panel data from the Mobidrive survey.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. INTRODUCTION

1.1 Background and Motivation

The development of accurate and behaviorally realistic travel demand models plays an important role in transportation and land-use planning. Travel demand models enable urban and transportation planners to predict when, how, where, how often and why people travel. Accurate predictions of these travel choices contribute toward better spatial and temporal estimations of travel demand and vehicle miles of travel (VMT), which, in turn, leads to the reliable assessment of travel demand management and/or transportation infrastructure development and emission control measures. However, accurate predictions alone do not suffice as a criterion for good modeling practice. In particular, it is easy to statistically fit a travel demand model to any data and to any level of accuracy by including a number of model parameters. Thus, another important criterion for model development is the incorporation of behavioral realism. In addition to facilitating temporal/spatial transferability and informed policy analysis, the incorporation of behavioral realism in travel demand models helps establish the credibility of these models outside the modeling community.

The development of accurate and behaviorally realistic travel demand models requires a good understanding of the factors and processes that influence the travel behavior of individuals. One of the key contributions toward understanding travel behavior has been the development of the activity-based paradigm that views travel as being derived from the desire to participate in activities. This more intuitive approach has a stronger basis in behavioral realism than the traditional trip-based approach to travel demand modeling (see Bhat et al., 2003, for a detailed comparison of the two approaches). The focus of the activity-based approach is the modeling of the activity-travel patterns of individuals, which may be characterized by six broad attributes: (a) Motivation or, equivalently, the activity purpose (such as work, shopping, recreation etc.), (b) Location of participation of the activity (such as the work place, grocery store or gym), (c) Sequencing of activities and the time of day of activity participation, (d) Mode used to travel to the activity location (for example, auto, transit or a combination of the two), (e) Frequency of activity participation, and (f) Solo or joint activity participation (see, for example, Bowman and Ben-Akiva, 2000, Bhat and Singh, 2000, Bhat and Misra, 2002, and Hamed and Mannering, 1993).

Of all these attributes, the choice of location of activity participation is one that has received relatively inadequate attention in the literature. On the other hand, the location of activity participation spatially pegs the daily activity-travel patterns of individuals. Accurate predictions of activity location are, therefore, key to effective travel demand management and air quality control strategies. Moreover, an understanding of the factors that influence the choice of location can contribute to more effective land-use and zoning policies. For instance, a habit-persistent individual may be more likely to continue shopping at the same grocery store rather than switching in response to a new land-use policy that brings more shopping opportunities closer to home.

The choice of location of activity participation and the factors that influence this choice vary with the activity purpose. Generally, the work location for most people is fixed in the short-term (teleworking individuals may face the choice between working from home and traveling to the office). Non-work activity participation, on the other hand, is typically characterized by a high degree of spatial-temporal flexibility and discretion. The choice of location for non-work activities can thus vary significantly not only across individuals but also across choice occasions of an individual (see Hanson and Huff, 1988, Bhat, 1999). This dissertation research is focused on conceptually analyzing the non-work location choice problem and developing a comprehensive econometric model of non-work location choice that incorporates both accuracy and behavioral realism.

The development of accurate and behaviorally realistic non-work location choice models is useful not only from the transportation and urban planning perspective but also from the perspective of service, retail and real estate businesses. [1] Predictions of where people shop and spend their recreational and leisure time plays an important role in the location and marketing decisions of businesses and firms. Literature, therefore, abounds in shopping location and store choice models (for example, see Burnett, 1977, Roy, 1981, Recker and Schuler, 1981, Fotheringham, 1988, Cadwallader, 1995, Rust and Donthu, 1995, Gonzalez-Benito, 2002, and Beynon et al., 2002), and models of recreational and leisure site location (for example, see Train, 1998, Parsons and Hauber, 1998, and Kemperman et al., 2000, 2002, 2004). Another study area related to store choice modeling is that of location-allocation decisions. The focus of these models is the optimal location of facilities, services, and industries as a function of demand allocation and other factors (for example, see Hansen, 1987, Daskin et al., 2003, Shukla and Waddell, 1991,

---

[1] Location choice models are also referred to as 'destination choice models' and 'attraction-end choice models' in the trip-based travel analysis literature.

and Perl and Ho, 1990). Location choice models have also been applied to model residential choice decisions (for example, see Roy, 1981, Feather, 1994, Ben-Akiva and Bowman, 1998, and Bhat and Guo, 2004) and migration location decisions (for example, see Fotheringham, 1991, Slater, 1992, Pellegrini and Fotheringham, 1999). The evident diversity in location choice modeling applications and the extent of its potential is underscored by the study of Xue and Brown (2003), who have developed a decision model for spatial site selection by criminals. A more comprehensive survey of literature in the field of location choice modeling is presented in chapter 2.

1.2 Research Context and Objectives

As discussed in the previous section, the focus of this dissertation research is the choice of location for non-work activity participation. The broad objectives of this research effort are two-fold. The first objective is to develop a comprehensive econometric model of location choice for non-work activities that incorporates accuracy and behavioral realism in capturing different kinds of choice behaviors. The estimation of such a model typically requires the use of simulated maximum likelihood inference. The second broad objective of this research is to contribute to improving the efficiency of the simulated maximum likelihood estimation (SMLE) by comparing the performance of various quasi-Monte Carlo (QMC) sequences and their scrambled versions. Each of these broad objectives is addressed in the following sections (sections 1.2.1 and 1.2.2) and the specific research objectives are summarized in the final section (section 1.3).

1.2.1 Location Choice Modeling for Non-work Activity Participation

The development of an accurate and behaviorally realistic model of location choice for non-work activity participation necessitates a good understanding of the factors influencing the choice process. Extensive research has therefore been directed toward a better understanding of the cognitive processes, preference behavior and decision rules underlying location choice, and the spatial aspect of travel. Studies in psychology (for example, Pipkin, 1979, 1981, Halperin et al., 1983, Gärling et al., 1984, Anooshian and Seibert, 1996) deal with spatial cognition issues, i.e., the mechanics of the human brain in processing spatial information, and the resulting mental maps of the spatial layout of activities. Studies in geography are directed more toward understanding spatial interaction issues and their effects on choice behavior (see Sheppard, 1979, Birkin and Clarke, 1991, Gould and White, 1974, Fotheringham et al., 2001).

The multi-disciplinary nature of the location choice problem and the diversity of potential applications have encouraged researchers in various fields to develop location choice models that bring together one or more of the various concepts of spatial interaction, cognition, preference behavior (such as habit persistence and variety-seeking) and decision rules. Smith (1978), for instance, develops a framework based on Portfolio theory with a Bayesian approach to account for location choice behavior under uncertainty (due to incomplete information). Although this model considers the effects of individual preferences and choice criteria, it ignores spatial interaction issues and is limited in its exploration of cognition. Hsu and Hsieh (2004) present an individual accessibility model to explain different travel-related decisions including location choice, which is behaviorally descriptive and incorporates spatial-temporal constraints. However, their model is also limited in its exploration of spatial interaction and cognition. Kemperman et al. (2000, 2002) incorporate seasonality and variety-seeking effects in their model of the choice of theme parks, while Train (1998) focuses on inter-personal taste variations in anglers' choices of fishing sites. These studies capture some of the observed and unobserved factors that contribute to inter- and intra-personal heterogeneity in the observed travel patterns, but the models are limited in capturing cognitive influences and do not address spatial interaction. Burnett (1978) applies Markovian theory to explain shopping location choice in terms of cognitive learning. Individuals potentially learn at every opportunity and the effects of learning on choice could potentially be carried infinitely into the future. Therefore, feedback-based mechanisms of learning, in the interest of practicality, must assume a state of learning equilibrium when individuals are supposed to have learnt all there is to learn. Further, they make assumptions regarding the starting point of the learning process that are known as initial conditions. While Burnett's model accounts for the effects of increasing familiarity and new information (in other words, learning) on travel behavior, the model does not address the question of initial conditions and the definition of learning equilibrium. In addition, the model does not consider the differences in individual preferences and spatial interaction issues.

All the above studies incorporate to varying degrees the effects of spatial cognition, preference behavior and attitudes on the choice of location. However, they do not incorporate the effects of interactions between the spatial choice alternatives.

Although the spatial nature of travel choices has been acknowledged in the literature (see Bhat, 2000, and Goulias, 2002), there are few studies that actually incorporate the various spatial interaction effects in modeling location choice (exceptions include Pozsgay and Bhat, 2002, Kanaroglou and Ferguson, 1996, Bolduc et al., 1997, Dellaert et al., 1998, and Bhat and Guo, 2004). The location choice studies that incorporate spatial interaction effects mostly ignore the effects of spatial cognition and preference behavior.

The first broad objective of this research is to develop a behaviorally realistic location choice model for non-work activity participation that comprehensively incorporates the effects of spatial cognition, preference behavior and spatial interaction. The proposed model thus accommodates both inter- and intra-individual variations in location choice behavior due to various factors such as habit persistence, variety-seeking, cognitive learning and spatial-temporal constraints. The model also accommodates different spatial interaction effects such as spatial heterogeneity, spatial autocorrelation, agglomeration and competition.

1.2.2 Comparison of QMC Sequences in SMLE of Discrete Choice Models

The incorporation of behaviorally realistic concepts, such as spatial cognition and spatial interaction, in the proposed econometric model of location choice is achieved through the relaxation of restrictions that impose inappropriate behavioral assumptions regarding the underlying choice process. This relaxation of behavioral restrictions on choice model structures, in many cases, leads to analytically intractable choice probability expressions, which necessitate the use of numerical integration techniques to evaluate the multidimensional integrals in the probability expressions. The most commonly used technique in the literature is the pseudo-Monte Carlo (PMC) simulation method that evaluates a multi-dimensional integral by replacing it with an average of the values of the integrand computed at N discrete and random points (PMC sequence).

Extensive number theory research in the last few decades has led to the development of a more efficient simulation method, the quasi-Monte Carlo (QMC) method that uses the basic principle of the PMC method. However, rather than using random sequences, QMC methods use low-discrepancy, deterministic, quasi-Monte Carlo (or QMC) sequences that are designed to achieve a more even distribution of points in the integration space than the PMC sequences. Research on the generation and application of QMC sequences clearly indicates the superior accuracy of QMC methods over PMC methods in the evaluation of multidimensional integrals (see Morokoff and Caflisch, 1994, 1995). In particular, the advantages of using QMC simulation for such applications in econometrics as simulated maximum likelihood inference, where parameter estimation entails the approximation of several multidimensional integrals at each iteration of the optimization procedure, should be obvious. However, the first introduction of the QMC method for the simulated maximum likelihood inference of econometric choice models occurred only in 1999, when Bhat tested Halton sequences for mixed logit estimation and found their use to be vastly superior to random draws. Since Bhat's initial effort, there have been several successful applications of QMC methods for the simulation estimation of flexible discrete choice models, though most of these applications have been based on the Halton sequence (see, for example, Revelt and Train, 2000; Bhat, 2001; Park et al., 2003; Bhat and Gossen, 2004). Number theory, however, abounds in many other kinds of low-discrepancy sequences that have been proven to have better theoretical and empirical convergence properties than the Halton sequence in the estimation of a single multidimensional integral. For instance, Bratley and Fox (1988) show that the Faure and Sobol sequences are superior to the Halton sequence in terms of accuracy and efficiency. There have also been several numerical studies on the simulation estimation of a single multidimensional integral that present significant improvements in the performance of QMC sequences through the use of scrambling techniques (Kocis and Whiten, 1997; Wang and Hickernell, 2000).

The second broad objective of this research is, therefore, to examine the performances of the different QMC sequences and their scrambled versions in the simulation estimation of flexible discrete choice models such as the location choice model described in the previous section.

1.3 Specific Research Objectives and Tasks

As discussed in the preceding sections, the first broad objective of this dissertation is to perform an in-depth analysis of the choice of location for non-work activity participation. The development of an accurate and behaviorally realistic model of location choice for non-work activity participation necessitates a good understanding of all the observed and unobserved factors influencing location choice, combined with a sound theory relating spatial interaction, cognitive processes, preferences and decision rules to the observed choice. There exists a large body of literature that has made significant contributions toward this goal. The first research task is to conduct an extensive survey of this literature to guide the efforts toward the development of a theory of non-work activity location choice. The next task is to develop a comprehensive framework of location choice decision-making for non-work activity participation that incorporates all the observed and unobserved factors that potentially influence the decision-maker while also considering cognitive and spatial interaction processes. Finally, the conceptual framework is translated into a general econometric model of location choice for non-work activity participation.

The second broad objective of this dissertation is to contribute to improving the efficiency of the simulated maximum likelihood estimation (SMLE) procedure by comparing the performance of various quasi-Monte Carlo (QMC) sequences and their scrambled versions. A suitable experimental design is constructed for the comparison of the various sequences, and numerical experiments conducted to identify the best sequence.

Finally, the two broad objectives of this dissertation research are tied together in the empirical estimation of non-work location choice models using a real-life multi-day dataset. These models will apply the comprehensive location choice model structure developed as a part of this research effort, and use the best QMC sequence identified through experiments for the simulated maximum likelihood estimation.

1.4 Dissertation Outline

The rest of this dissertation is organized as follows.

Chapter 2 presents the results of the survey of spatial choice modeling literature. This extensive review of the literature helps in the identification of the key issues connected with understanding location choice behavior. Based on the observations drawn from the literature survey in Chapter 2, Chapter 3 describes in detail the development of a comprehensive conceptual framework of location choice decisions for non-work activity participation. Chapter 4 presents a general location choice model structure that is developed based on the framework presented in chapter 3.

The next two chapters present the work undertaken in comparing the efficiency of QMC sequences in the simulated maximum likelihood estimation of discrete choice models. Chapter 5 provides a background on the generation of QMC sequences and also describes the specific objectives of our research in examining these sequences. Chapter 6 describes the numerical experiments performed with the QMC sequences and presents the results.

The results of all these research efforts are applied in chapter 7, which presents an empirical analysis of location choice for non-work travel and discusses the policy implications of such an analysis. Chapter 8 concludes this dissertation with a summary of the main findings. The limitations of this research are discussed and potential extensions of the work are identified in this chapter.

CHAPTER 2. SPATIAL CHOICE MODELING

The origins of location choice models lie in the aggregate modeling, undertaken over a century ago, of spatial movements and flows of people and commodities. Since then the need for greater accuracy and predictive capability and the drive for the incorporation of behavioral realism has resulted in disaggregate individual-level location choice models. Location choice models, as we know them today, model an individual's observed choice of location for activity participation as a function of the individual's socio-economic characteristics and the attributes of the alternative locations. The spatial aspect of activity participation, however, is not completely understood and, although most researchers agree on its importance, is often inadequately represented in location choice models (see, for example, Odland, 1981, Miller and O'Kelly, 1983, Cadwallader, 1995, Train, 1998, and Kemperman et al., 2004). This is slowly changing as we see more location choice studies explicitly incorporating spatial interaction effects (for example, see Kanaroglou and Ferguson, 1996, Bolduc et al., 1997, Bhat and Guo, 2004; other activity-based models that are not focused on the location choice problem but incorporate spatial effects include Bhat, 2000, and Goulias, 2002).

The modeling of spatial movements and flows, as mentioned earlier, has been the subject of research for over a century and there is a correspondingly extensive body of literature associated with it. These studies span a wide range of applications, apply a variety of different theories, and implement a number of modeling methodologies. In order to assimilate the contributions of this literature, it is necessary to classify the literature into manageable categories. In the following sections a four-level classification scheme for the literature is first developed (section 2.1), followed by a review of the literature from the perspective of each of the different levels (sections 2.2, 2.3, 2.4 and 2.5).

2.1 Classification of Literature

The general process of developing a location choice model, or any other model for that matter, can be simplified to three broad steps. The first step is the choice of an underlying theory. The location choice problem viewed from a classical microeconomic perspective, for instance, would treat individuals as rational entities with access to full information, who select the location that optimizes their utility subject to constraints. The second step is the choice of a modeling methodology. Continuing with the previous example of classical microeconomic theory, this could be a constrained linear or non-linear optimization model of individual utility. The third step is the application. The optimization model developed in the previous step could be applied to model the shopping location decision of the individual. Following this simplified process of developing a location choice model, the literature may be classified by (a) underlying theory, (b) modeling methodology, and (c) application area. A fourth level in this scheme classifies the literature by the data used in the application. Figure 1 illustrates this classification scheme graphically.

Before proceeding with a review of the literature as per this classification scheme, it is important to take note of a couple of issues. First, in generating a manageable number of categories for each level of classification, it is inevitable that some of the literature will bridge several categories. In such situations, discretion is used to classify this literature as accurately as possible. Second, many research efforts focus solely on the development, rather than the choice, of an underlying theory or modeling methodology. Such studies appear only in the applicable levels of the classification scheme (for example, a research effort that develops a modeling methodology without an empirical application will only be classified by underlying theory and modeling methodology).

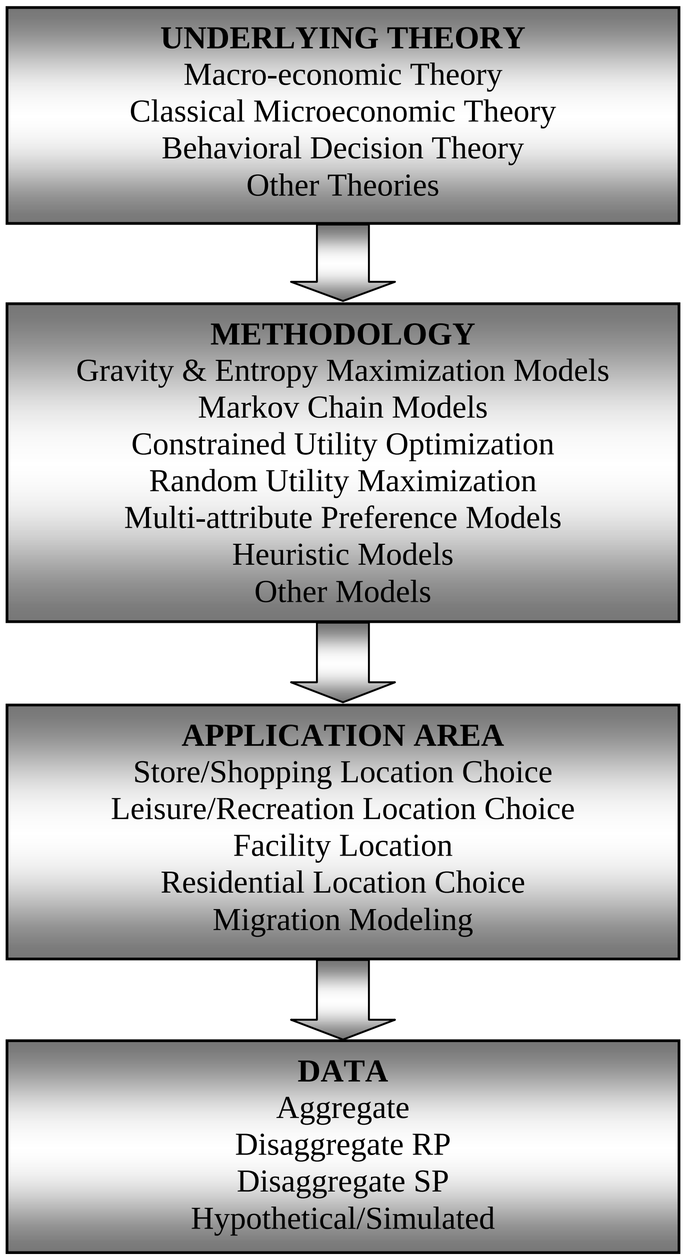

Figure 1. Classification Scheme for Spatial Flow/Choice Models

The following sections review the literature in turn by each of the four levels of the classification scheme.

2.2 Classification Based on Underlying Theory

Table 1 classifies the literature by their underlying theories. There are three widely known and used theories, the macro-economic theory, the classical microeconomic theory and the behavioral decision theory. The fourth category includes several less known theories such as information integration theory, spatial interaction and aggregation theories, and the Dempster-Shafer theory. Each of these categories is discussed in the following sections.

Table 1. Classification by Underlying Theory

Underlying Theory	Papers
Macroeconomic Theory	Carey (1859); Roy (1981); Ravenstein (1885); Hecksher (1919); Ohlin (1933); Stouffer (1940); Stewart (1941); Schneider (1959); Isard (1960); Huff (1963); Wilson (1967, 1971, 1974); Long and Uris (1971); Cesario (1973); Fisk and Brown (1975); Harris and Wilson (1978); Roy and Lesse (1981); Haynes and Fotheringham (1984); Cadwallader (1975);
Classical Micro-Economic Theory	Goodchild (1978); Odland (1981); Berman et al. (1992); Perl and Ho (1990); Drezner et al. (1991); Brandeau and Chiu (1994); Laporte et al. (1994); Cadwallader (1995); Hsu and Hsieh (2004); Gannon (1972); Wesolowsky (1973); Morris and Norback (1980);
Behavioral Decision Theory	Recker and Schuler (1981); Miller and O'Kelly (1983); Dunn and Wrigley (1985); Borgers and Timmermans (1987); Shukla and Waddell (1991); Timmermans et al. (1992); Feather (1994); Rust and Donthu (1995); Bolduc et al. (1997); Eymann and Ronning (1997); Leszczyc and Timmermans (1996); Bell et al. (1998); Fotheringham (1988); Dellaert et al. (1998); Train (1998); Parsons and Hauber (1998); Pellegrini and Fotheringham (1999, 2002); Pozsgay and Bhat (2002); Xue and Brown (2003); Miyamoto et al. (2004); Sermons and Koppelman (1998, 2001); **Learning** Aaker and Jones (1971); Burnett (1977, 1978); Meyer (1979);

	Cognition Golledge and Briggs (1973); Golledge and Zannaras (1973); Jones (1977); Golledge and Spector (1978); Smith (1978); Pipkin (1981); Hanson and Hanson (1981); Halperin et al. (1983); Anooshian and Seibert (1996); Fotheringham and Curtis (1999); Gärling et al. (1984); Burnett (1976, 1978);
Other theories	**Information Integration** Louviere and Woodworth (1983); Louviere (1984); Oppewal et al. (1994); Kemperman et al. (2000, 2004); Levin and Louviere (1979); **Spatial Interaction** Gould and White (1974); Sheppard (1979); Birkin and Clarke (1991); Leonardi and Papageorgiou (1992); **Spatial Aggregation** Kitamura et al. (1979); Kanaroglou and Ferguson (1996); Ferguson and Kanaroglou (1998); Gonzalez-Benito (2002); Orpana and Lampinen (2003); **Dempster-Shafer** Beynon et al. (2002); **Miscellaneous** Aldskogius (1977); MacKay (1973); Thill and Wheeler (2000); Johnson and Payne (1985);

2.2.1 Macroeconomic Theory

The origin of spatial choice modeling lies in the macroeconomic modeling of spatial flows, which was first proposed as a means of modeling commodity flows across the nation (Carey, 1859). Macro-economic theory suggests that spatial flows may be modeled by the large-scale movements and flows of passengers and commodities. Modeling methodologies developed under this theory, such as the gravity model and entropy maximization models[2], are thus concerned with interaction patterns that result from aggregating individual choices across zones. Macroeconomic theory, though useful in the aggregate description of spatial flows, has been subject to substantial criticism for its lack of behavioral realism in explaining individual location choices. Nevertheless, this approach has been around for over a century and continues to enjoy widespread application in transportation planning.

2.2.2 Classical Microeconomic Theory

Classical microeconomic theory was adopted in spatial flow modeling as a means of describing individual choice behavior. It is based on the premise that individuals have access to full information and behave rationally in optimizing their total utility (maximizing profits or minimizing costs) subject to budget constraints. However, the assumptions of full information and perfect rationality render the classical microeconomic theory deficient in describing choice behavior realistically since people typically have access to incomplete information and oftentimes exhibit satisficing behavior. Classical microeconomic theory has therefore seen much more application in facility location modeling, where the profit maximization objective is more critical and appropriate, than individual choice modeling (Wesolowsky, 1973, Morris and Norback, 1980, Laporte et al., 1994, Brandeau and Chiu, 1994, Drezner et al., 1991, Berman et al., 1992, Perl and Ho, 1990). The biggest contribution of this theory to modeling the choice behavior of individuals has been the development of the behavioral decision theory.

---

[2] Gravity and entropy maximization models are often referred to as aggregate spatial interaction models in the literature. The term ‘spatial interaction models’ is somewhat a misnomer as many of these models do not really account for spatial interaction effects. However, this is term has been in use for several decades and is used widely in geographic literature.

2.2.3 Behavioral Decision Theory

Behavioral decision theory has developed over time as it evolved from the classical microeconomic theory. The relaxation of the restrictive assumptions of full information and perfect rationality that made classical microeconomic theory unsuitable for modeling individual choice behavior first resulted in the development of economic consumer theory. Consumer theory, while maintaining its roots in microeconomic theory, allows for satiation, taste variation and growth of needs. Over time this has further evolved to include concepts such as bounded rationality, decision-making under risk and uncertainty, and cognition and evolutionary behavior such as learning. This collection of concepts is now known as Behavioral Decision Theory.

Most of the location choice literature that falls under the category of Behavioral Decision Theory considers one or more behavioral aspects of spatial choice-making. For instance, Recker and Schuler (1981) examine the order in which destination attributes are processed by individuals in their choice of grocery store; Miller and O'Kelly (1983) examine the effects of past choices on shopping location choice through feedback; Train (1998) examines taste variations across anglers in their choice of fishing site; Dellaert et al. (1998) examines shopping location choice from a trip-chaining perspective. Also in this category are studies focused on understanding evolutionary learning behavior through the modeling of a sequence of repeated location choices (Aaker and Jones, 1977, Burnett, 1977, 1978, among others), and studies focused on understanding cognitive processes within the human brain (Golledge and Spector, 1978, Pipkin, 1979, 1981, and Anooshian and Seibert, 1996, and Fotheringham and Curtis, 1999, to name a few). The literature in the field of spatial cognition is very extensive and only those papers that are directly connected with location choice behavior are included in this category (see Golledge and Timmermans, 1990a, for an extensive survey of the literature in spatial cognition).

2.2.4 Other Theories

The “other theories” in the fourth category include information integration theory, spatial interaction and spatial aggregation theories, and Dempster-Shafer theory. And then there are a few miscellaneous studies that are not based on any specific theory, rather they attempt to statistically determine patterns in the observed choice data that can be related to individual characteristics and behavioral processes (Aldskogius, 1977, MacKay, 1973, Thill and Wheeler, 2000, Johnson and Payne, 1985).

Information integration theory (Anderson, 1976) asserts that a response is the result of the integration of information according to simple algebraic rules such as addition, averaging, subtraction and multiplication. The application of this theory to choice behavior is based on the assumption that individuals cognitively integrate their subjective evaluations of the attributes of the alternatives to derive the utility for each choice alternative. The applications of this theory are therefore based on data from experimental designs, also known as stated preference surveys, which enable the collection of subjective evaluation data.

The focus of spatial interaction theory is the premise that the arrangement of decision-makers and alternatives in the study area influences choice behavior. Various spatial interaction effects such as spatial correlation, and spatial heterogeneity have been examined in the literature (see Bhat, 2000, for a description of these spatial interaction effects). The objective of spatial interaction theory is in essence the same as that of spatial aggregation theory i.e. to correctly capture the effects of the spatial structure on choice behavior. The theory of spatial aggregation effects attempts to understand and model the effects of aggregating individual locations and elemental attraction units into zones, which, in fact, translates to the examination of spatial correlation, heterogeneity and other spatial interaction effects. The primary difference between the two theories is that spatial aggregation theory considers the effects of spatial structure by explicitly considering the effects of aggregation on choice models. To that extent, spatial aggregation theory is more a modeling methodology.

Dempster-Shafer theory is based on belief functions and subjective probability quantification (Shafer, 1990). This theory is applied to spatial choice behavior to determine 'favorite locations' for activity participation based on revealed preferences. The definition of 'favorite locations' is fuzzy and based on belief and plausibility measures. This theory proposes to capture the effects of social interactions on choice behavior.

2.3 Classification Based on Modeling Methodology

Table 2 classifies the literature by their modeling methodologies. The modeling methodologies have been broadly classified into gravity and entropy maximization models, constrained optimization models, markov chain models, random utility maximization models, multi-attribute preference models and heuristic models. The other modeling methodologies include log-linear models and verbal hypotheses. As seen in

Figure 2, nearly all the spatial choice studies based on macroeconomic theory use gravity and entropy maximization formulations, and all those based on classical microeconomic theory use constrained optimization formulations. Each of the categories in this level of the classification scheme is discussed in the following sections.

Table 2. Classification by Modeling Methodology

Modeling Methodology	Paper	Underlying Theory
Gravity & Entropy Maximization	Carey (1859); Ravenstein (1885); Hecksher (1919); Ohlin (1933); Stouffer (1940); Stewart (1941); Schneider (1959); Isard (1960); Huff (1963); Wilson (1967, 1971, 1974); Long and Uris (1971); Cesario (1973); Fisk and Brown (1975); Harris and Wilson (1978); Roy (1981); Roy and Lesse (1981); Haynes and Fotheringham (1984); Cadwallader (1975);	**Macroeconomic**
Constrained Optimization	Goodchild (1978); Odland (1981); Berman et al. (1992); Perl and Ho (1990); Drezner et al. (1991); Brandeau and Chiu (1994); Laporte et al. (1994); Hsu and Hsieh (2004); Gannon (1972); Wesolowsky (1973); Morris and Norback (1980);	**Classical Micro-Economic**
Markov Chain	Aaker and Jones (1971); Burnett (1977, 1978);	**Behavioral Decision**
Random Utility Maximization	Recker and Schuler (1981); Miller and O'Kelly (1983); Dunn and Wrigley (1985); Borgers and Timmermans (1987); Dellaert et al. (1998); Shukla and Waddell (1991); Timmermans et al. (1992); Feather (1994); Rust and Donthu (1995); Bolduc et al. (1997); Eymann and Ronning (1997); Leszczyc and Timmermans (1996); Bell et al. (1998); Fotheringham (1988); Train (1998); Parsons and Hauber (1998); Pellegrini and Fotheringham (1999, 2002); Pozsgay and Bhat (2002); Xue and Brown (2003); Kemperman et al. (2004); Miyamoto et al. (2004); Sermons and Koppelman (1998, 2001);	**Behavioral Decision**
	Kitamura et al. (1979); Kanaroglou and Ferguson (1996); Ferguson and Kanaroglou (1998); Gonzalez-Benito (2002); Orpana and Lampinen (2003);	**Spatial Aggregation**

Multi-attribute Preference	Louviere and Woodworth (1983); Louviere (1984); Oppewal et al. (1994); Kemperman et al. (2000); Levin and Louviere (1979);	**Information Integration Theory**
Heuristic Models	Aldskogius (1977); MacKay (1973); Thill and Wheeler (2000); Fischer and Reismann (2002); Johnson and Payne (1985)	**-**
<u>Other Models</u>		
Log-linear	Cadwallader (1995)	**Classical Micro-economic**
Bayesian	Beynon et al. (2002)	**Dempster-Shafer**
Verbal Hypotheses	Horton and Reynolds (1969); Wheeler and Stutz (1971); Schönfelder and Axhausen (2004);	**-**

2.3.1 Gravity and Entropy Maximization

The first approach to address movements and flows across space was the macroeconomic gravity model based on Newton's Theory of Gravity, dating back to the late 1800s. The first gravity model (Carey, 1859), as seen in equation 1, computes the number of trips between origin *i* and destination *j* (T_{ij}) as a simple function of the sizes of the origin and destination (P_i and P_j), and the distance between them (d_{ij}) using a scaling factor *k*.

$$T_{ij} = k\frac{P_i P_j}{d_{ij}} \qquad \textbf{Eq. 1}$$

This formulation then gave way to a more general one that recognizes that the relationships embedded in equation 1 may vary across trip types and with the socio-economic attributes of zones. The new formulation also recognizes that many origin and destination attributes, rather than just the two size variables, could potentially influence the flow patterns. This more general gravity formulation (Haynes and Fotheringham, 1984) computes the number of trips between origin *i* and destination *j* (T_{ij}) as

$$T_{ij} = k\frac{V_{i1}^{\alpha 1}V_{i2}^{\alpha 2}...V_{if}^{\alpha f}V_{j1}^{\lambda 1}V_{j2}^{\lambda 2}...V_{jg}^{\lambda g}}{d_{ij}^{\beta}}, \quad \textbf{Eq. 2}$$

where there are *f* origin attributes, (V_{i1}, V_{i2}, …, V_{if}), and *g* destination attributes, (V_{j1}, V_{j2}, …, V_{jg}), and each attribute contributes to the flows to a different degree, as indicated by the factors $\alpha 1,\ldots,\alpha f$, and $\lambda 1,\ldots,\lambda g$. Although this model has been shown to produce reasonably accurate estimates of spatial flows, it does not possess any theoretical grounding in the travel behavior of individuals.

The next major advance in the macroeconomic modeling of spatial flows was the development of the entropy maximization theory (Wilson, 1967, 1974), which give rise to a family of spatial interaction models including the gravity model, the production-constrained model (equation 3), the attraction-constrained model (equation 4), and the doubly-constrained model (equations 5,6,7).

$$T_{ij} = \frac{O_i P_j^{\lambda} d_{ij}^{\beta}}{\sum_j P_j^{\lambda} d_{ij}^{\beta}}, \quad \textbf{Eq. 3}$$

$$T_{ij} = \frac{D_j P_i^{\alpha} d_{ij}^{\beta}}{\sum_i P_i^{\alpha} d_{ij}^{\beta}}, \text{ and} \quad \textbf{Eq. 4}$$

$$T_{ij} = A_i O_i B_j D_j d_{ij}^{\beta}, \text{ where} \quad \textbf{Eq. 5}$$

$$A_i = \sum_j (B_j D_j d_{ij}^{\beta})^{-1}, \quad \textbf{Eq. 6}$$

$$A_i = \sum_j (B_j D_j d_{ij}^{\beta})^{-1}, \quad \textbf{Eq. 7}$$

O_i is the known total flow from origin *i*, P_j is the population of destination zone *j*, and D_j is the known total inflow into destination *j*. The entropy procedure is based on the enumeration approach of combinatorial analysis, and is derived from statistical mechanics. The basic premise of these models is to enumerate all the possible zone-to-zone flow interchanges and pick the one with the highest uncertainty subject to constraints. However, as with the gravity model, these models have also been criticized for lacking a behavioral perspective in explaining individual travel decisions (despite which the production-constrained spatial interaction model continues to be widely used).

2.3.2 Constrained Optimization

The evident need for spatial flow models that can describe an individual's choice of location resulted in the adoption of classical microeconomic theory and with it the use of constrained optimization models. Constrained optimization models construct the utility of the decision-maker (individual, household, factory etc.) as a demand and supply function, wherein resources are consumed and utility is gained to varying degrees depending on the choice of the alternative and the decision-maker selects the alternative that maximizes utility subject to resource constraints (such as time or money). A general formulation of a constrained optimization model is presented in equation 8, where *U*, the utility of the decision-maker, is a function of the decision variables in vector $\bar{x}$; $\bar{c}$ is a vector of cost functions associated with the decision variables; and *C* is total quantity of resources available to the decision-maker.

$$\max U(\bar{x}), \quad \textit{subject to } \bar{c}\,\bar{x} \le C \qquad \textbf{Eq. 8}$$

Constrained optimization models as per classical microeconomic theory are typically applied to continuous decision variables, though they have been adapted to discrete choices of location. For example, Odland, 1981, adapts the constrained utility optimization model for household location choices, while Gannon (1972), Wesolowsky

(1973), Berman et al. (1992), and others apply them to industrial location choice decisions.

2.3.3 Markov Chain

The use of markov chain models appeared in spatial choice modeling as one of the first attempts at incorporating behavioral realism within the micro-economic approach. Specifically, markov chain models of location choice for shopping (Aaker and Jones, 1971 and Burnett, 1977) were developed to investigate the adaptive or evolutionary learning behavior of decision-makers. In general, if $D = \{D_1,....,D_i,....,D_k)$ is a constant set of mutually exclusive and exhaustive location choice alternatives, then the set of probabilities $(P_{1,n+1},....,P_{i,n+1},....P_{k,n+1})$ that an individual i will choose spatial alternative D_i (i = 1,2,…k) on his or her next (n+1)th choice is given by the assumption of a steady-state Markov process. Under suitable assumptions of Markov processes the probabilities for any chain of location choices may be determined. Log-linear models (Burnett, 1978, Cadwallader, 1995) are simply extensions of simple markov chain processes that allow the probabilities to vary by choice occasion (this type of time-dependent markov processes are referred to as being non-stationary).

2.3.4 Multi-attribute Preference

Multi-attribute preference models, with their roots in information integration theory (Louviere, 1984), are used to model the subjective element of individual choices. These models are typically based on stated-preference (SP) survey data, which are based on individual responses to experimentally designed questions. Individuals indicate their preference for various attributes, both subjective and objective, of the choice alternatives and their utilities are formulated as being an integration of the evaluation of all the attributes. In Louviere and Woodworth (1983) and Louviere (1984), for instance, the utility of the alternatives is modeled as a weighted function of the various attributes. Multi-attribute preference models are also known in the literature as decompositional models or conjoint choice models, though the term 'conjoint choice models' is typically extended to all kinds of formulations that use SP data including multinomial logit formulations based on random utility maximization.

2.3.5 Random Utility Maximization

The derivation of the discrete choice model, based on the theory of random utility maximization, by McFadden in the late 1970s provided a big impetus toward the modeling of spatial flows. The discrete choice paradigm not only brings the problem of understanding spatial flows down to the individual decision-maker but also acknowledges the discrete nature of spatial choices (although space is a continuous entity individuals choose specific locations). It *"assumes the existence of a utility maximizing decision-maker confronted with a set of mutually exclusive, collectively exhaustive alternatives, only one of which can be selected. This decision maker is further assumed to associate some utility with each alternative, and this utility is random for two possible reasons: Either some aspects of the decision problem or decision-maker lead to variability in the evaluation of the alternatives that cannot be explained by any measurable factors, or the analyst, when attempting to explain the observed choices, cannot fully characterize the decision-maker's choice criterion."* (see Lerman, 1983). Thus the discrete choice model effectively addresses the various factors, both observed and unobserved, influencing location choice decisions and hence spatial flows.

The utility that individual i attributes to location j, U_{ij}, is therefore composed of two parts, a measurable component, V_{ij}, and a random error component, ε_{ij}.

$$U_{ij} = V_{ij} + \varepsilon_{ij} \qquad \textbf{Eq. 9}$$

Based on the theory of utility maximization and conditional on assumptions placed on the random component, several formulations such as the multinomial probit or logit models may be derived. The multinomial logit model, for instance, is based on the assumption of extreme value distribution of the error term and is given by

$$p_{ij} = \frac{e^{V_{ij}}}{\sum_{j \in C} e^{V_{ij}}}, \qquad \textbf{Eq. 10}$$

where, p_{ij}, is the probability that individual i selects location j, and C is the choice set of locations available to the individual. The multinomial logit (MNL) formulation in its simplest form (equation 10) suffers from the undesirable property of independence from irrelevant alternatives (IIA) despite which there have been several applications of this model in the location choice literature (Recker and Schuler, 1981, Shukla and Waddell, 1991, Sermons and Koppelman, 1998, 2001, and Bell et al., 1998, to name a few). Some studies work around this shortcoming of the MNL model by using non-linear specifications of the utility function to introduce interaction effects among the spatial alternatives (example, Pozsgay and Bhat, 2002 and Fotheringham, 1988). Extensions of the MNL model, such as the nested logit (NL) and the mixed logit (ML) models relax the restrictive assumptions on the MNL in different ways and provide more flexible formulations for location choice modeling applications (see Miller and O'Kelly, 1983, Hansen, 1987, Eymann and Ronning, 1997, Train, 1998, Kemperman et al., 2004, Miyamoto, 2004).

2.3.6 Heuristics

Another approach to describing and modeling location choice behavior that is seen in the literature is that of heuristic models. Thill and Wheeler (2000) use decision trees to model shopping location choice, while Fischer and Reismann (2002) propose a neural network approach to modeling spatial choice. Others, such as Aldskogius (1977) and MacKay (1973) use informal map analysis, and trend surface analysis to capture patterns in spatial choice behavior.

2.3.7 Other Models

The other models used in the literature are log-linear models (Cadwallader, 1995) that are similar to multiple regression models, models based on the Bayesian theory of beliefs (Beynon et al., 2002) and several verbal hypotheses. Studies such as the ones by Horton and Reynolds (1969) and Wheeler and Stutz (1971) examine activity-travel patterns of individuals and generate hypotheses of activity-travel behavior as a whole. Examples of these include the hypotheses of <u>activity spaces</u> by Horton and Reynolds (1969) and Schönfelder and Axhausen (2004), and <u>social travel fields</u> by Wheeler and Stutz (1971).

2.4 Classification Based on Application Area

The earliest applications of spatial flow models were in the determination of person and freight flows for urban planning and demand/supply decisions. Since then modeling of spatial flows has been an important topic of research in many fields including economics, geography, psychology, marketing, engineering, econometrics, mathematics and transportation, and the many theories proposed by these research efforts contribute in some way to our understanding of the location choice problem. Ultimately, though, all these theories are targeted at developing better location choice models that can be applied reliably and effectively for purposes such as service and facility location, industrial or business location, and urban and transportation planning. The many applications of location choice models can be grouped into 5 broad categories – shopping location and store choice, leisure and recreation location choice, facility location, residential location choice and migration modeling. Table 3 classifies the literature by these categories, each of which is discussed in the following sections.

Table 3. Classification by Application Area

Application Area	Paper	Modeling Methodology	Data	Underlying Theory
Shopping Location and Store Choice	Burnett (1977, 1978)	**Markov Chain**	**RP**	**Learning**
	Miller and O'Kelly (1983)	**RUM**	**RP (t)**	**BDT**
	Cadwallader (1975)	**Gravity**	**Aggregate**	**Macro-Economic**
	Aaker and Jones (1971)	**Markov Chain**	**RP (t)**	**Learning**
	Dunn and Wrigley (1985)	**RUM**	**RP (t)**	**BDT**
	Borgers and Timmermans (1987)	**RUM**	**Synth**	**BDT**
	Dellaert et al. (1998)	**RUM**	**SP**	**BDT**
	Timmermans et al. (1992)	**RUM**	**SP**	**BDT**
	Rust and Donthu (1995)	**RUM**	**RP**	**BDT**
	Leszczyc and Timmermans (1996)			**BDT**
	Bell et al. (1998)	**RUM**	**RP (t)**	**BDT**
	Fotheringham (1988)	**RUM**	**RP**	**BDT**
	Gonzalez-Benito (2002)	**RUM**	**RP**	**BDT**
	Orpana and Lampinen (2003)	**RUM**	**Aggregate**	**BDT**
	Thill and Wheeler (2000)	**Heuristics**	**RP**	**-**
	Cadwallader (1995)	**RUM**	**RP/SP**	**BDT**
	Beynon et al. (2002)	**Bayesian**		**Dempster-Shafer**

Leisure and Recreation Choice	Eymann and Ronning (1997)	**RUM**	**RP**	**BDT**
	Train (1998)	**RUM**	**RP**	**BDT**
	Parsons and Hauber (1998)	**RUM**	**RP**	**BDT**
	Kemperman et al. (2000, 2004)	**RUM**	**SP**	**BDT**
	Pozsgay and Bhat (2002)	**RUM**	**RP**	**BDT**
	Aldskogius (1977)	**Heuristic**	**RP**	**-**
Facility Location	Goodchild (1978)	**Optimization**	**Hypo**	**Micro-Economic**
	Berman et al. (1992)	**Optimization**	**RP**	**Micro-Economic**
	Perl and Ho (1990)	**Optimization**	**RP**	**Micro-Economic**
	Drezner et al. (1991)	**Optimization**	**Hypo**	**Micro-Economic**
	Brandeau and Chiu (1994)	**Optimization**	**RP**	**Micro-Economic**
	Laporte et al. (1994)	**Optimization**	**Hypo**	**Micro-Economic**
	Gannon (1972)	**Optimization**	**Hypo**	**Micro-Economic**
	Wesolowsky (1973)	**Optimization**	**Hypo**	**Micro-Economic**
	Morris and Norback (1980)	**Optimization**	**-**	**Micro-Economic**
	Shukla and Waddell (1991)	**RUM**	**Aggregate**	**BDT**
	Bolduc et al. (1997)	**RUM**	**RP**	**BDT**
	Hansen (1987)	**RUM**	**-**	**BDT**
Residential Location Choice	Roy (1981)			
	Feather (1994)	**RUM**	**RP**	**BDT**
	Ben-Akiva and Bowman (1998)	**RUM**	**RP**	**BDT**
	Sermons and Koppelman (1998, 2001)	**RUM**	**RP**	**BDT**
	Bhat and Guo (2004)	**RUM**	**RP**	**BDT**
	Miyamoto et al. (2004)	**RUM**	**RP**	**BDT**
Migration Modeling	Pellegrini and Fotheringham (1999, 2002)	**RUM**	**RP**	**BDT**
	Kanaroglou and Ferguson (1996)	**RUM**	**-**	**BDT**
	Fotheringham et al. (2001)	**RUM**	**Synth**	**BDT**
	Slater (1992)	**RUM**		**BDT**

2.4.1 Shopping Location and Store Choice

Spatial choice modeling has been used extensively in the development of store choice models, which provide important consumer demand feedback to businesses and serve as inputs to marketing decisions. The type of data usually drives the modeling methodology. Gravity and entropy maximization models are typically used with aggregate data (Cadwallader, 1975), while multi-attribute preference and random utility maximization models are typically used with disaggregate data (Cadwallader, 1995, Gonzalez-Benito, 2002, Dunn and Wrigley, 1985, Borgers and Timmermans, 1987). Most spatial choice models in consumer theory, however, are based on disaggregate data, both from revealed preference (RP) and stated preference (SP) surveys. Other shopping location choice models have also been developed in applied geography, and transportation from the perspective of urban and transportation planning needs. One of the first of this kind was the shopping destination choice model developed by Miller and O'Kelly (1983). They implemented logit models with state dependence to account for the effects of past choices. Other shopping location models in the literature include Thill and Wheeler (2000), Fotheringham (1988), and Timmermans et al. (1992). Aaker and Jones (1971) and Burnett (1977, 1978) examine shopping travel patterns from an evolutionary learning perspective.

2.4.2 Leisure and Recreation Location Choice

Spatial choice models have also found application in the leisure sciences, in determining individuals' choices of recreational and leisure sites such as theme parks. The leisure sciences applications are mostly based on disaggregate data, both RP and SP, and invariably use the random utility maximization theory. For example, Kemperman et al. (2000) incorporate variety-seeking and seasonality effects into a mixed logit model of the SP choice of theme parks. Kemperman et al. (2004), on the other hand, use RP data to model heterogeneity and substitution effects in the propensity of visiting urban parks using a mixed logit model. Train (1998) also uses RP data to estimate random parameter logit models of anglers' choice of fishing sites, thus incorporating taste differences across decision-makers. Pozsgay and Bhat (2002) estimate a nonlinear-in-parameters multinomial logit model for destination choice in urban recreational trips. There have been several other research efforts that attempt to capture the complexity of location choice decisions for non-work trips in general. Horowitz (1980) presents a logit model for destination and mode choice in multi-destination non-work travel, while Kitamura (1984) examines the effects of trip chaining on destination choice decisions.

2.4.3 Facility Location

Somewhat related to store choice models is a class of location-allocation models known as facility location models. These models are developed from the perspective of firms, businesses, industries and public service agencies, in order to help them with their location decisions. These models are, therefore, developed with a view to maximize profits and the allocated consumer demand while ensuring low cost and ease of supply. Location decisions for industries and businesses (example, Wesolowsky, 1977) are treated differently from location decisions for public services (example, Perl and Ho, 1990), the key difference being the elasticity of demand expected for public service facilities. Also, location decisions for discretionary service facilities (example, Berman et al., 1992) are different from the location decisions for essential service facilities since customers tend to visit discretionary service facilities (such as ATMs and gas stations) on their way to other activity locations.

2.4.4 Residential Location Choice

Spatial choice models based on random utility maximization have also been widely used in housing choice analysis. In addition to explaining residential mobility these models also serve to inform the real-estate market of the expected spatial demand. There exists a substantial and rich body of literature in this area. Some examples include McFadden (1978), Onaka and Clark (1983), Timmermans (1992), Waddell (1993, 1996), and Bhat and Guo (2004). Guo (2004) presents an extensive literature review of residential location choice models in the literature.

2.4.5 Migration Modeling

Migration studies are inherently interdisciplinary and are a result of research in various fields including sociology, economics, demography, anthropology, and urban planning. The main aim of migration studies, specifically by geographers, is to understand where and why people migrate, and by this very definition involve the modeling of spatial choices. Fotheringham (1983) made one of the most important contributions to migration studies by developing a new spatial choice model known as the competing destinations model. The competing destinations model is based on the theory that individuals process spatial information hierarchically, and can also derived from the random utility maximization theory.

2.5 Review of Literature by Type of Data

The classification of spatial flow models can also be data-driven (see Table 3, the fourth column indicates the type of data used). Aggregate models predict total zone-to-zone flows, while disaggregate models predict the decision-maker's choice of specific locations, and each of these type of models has different data needs. The disaggregate models can be further classified based on the type of data as revealed preference (RP) and stated preference (SP) models. Some studies use panel revealed preference data (represented as RP (t) in Table 3). In addition, several studies use hypothetical (Hypo) or synthesized (Synth) data to demonstrate their theories (example, Goodchild, 1978, Gannon, 1972, Wesolowsky, 1973 and Fotheringham et al., 2001).

The first models of spatial flow, including the gravity model and the suite of spatial interaction models, were aggregate in nature. They used data on aggregate zone-to-zone flows and ignored the individual entity involved in the movements, be it people or freight (see Cesario, 1973, Roy and Lessee, 1981, and Cadwallader, 1975). Most other modeling methodologies used to analyze location choice in the literature revolve around the entity responsible for the spatial flows, such as the individual making destination choices or the businesses making location choices (see Dunn and Wrigley, 1985, Dellaert et al., 1998, Rust and Donthu, Berman et al., 1992), and are disaggregate in nature. These disaggregate models are, however, flexible enough to handle aggregate data if necessary, as seen in the study by Orpana and Lampinen (2003).

Within the suite of disaggregate models, some studies use revealed preference (RP) data (example, Miller and O'Kelly, 1983, Burnett, 1977 and 1978) while others use stated preference (SP) data (example, Timmermans et al., 1992 and Dellaert et al., 1998). Researchers prefer RP and SP data for different reasons. SP data derived from controlled experiments could potentially reveal the underlying decision processes, unlike RP data where only the final destination selections are observed which could be driven by unobserved constraints rather than true choices. RP data, on the other hand, can be used to simulate real-world decisions and thus develop better predictive models. Studies by

Bhat and Castelar (2002) and Ben-Akiva and Morikawa (1997) present the advantages of using both RP and SP data in model development.

Disaggregate models also vary in the level of aggregation of the choice alternatives modeled. Most studies in marketing and consumer theory model the choice of elemental units of attraction such as grocery stores. This is necessary since the aim of such studies is to examine the effects of marketing policies on the demand at specific store locations. Marketing surveys that drive such modeling efforts are also geared to gather data at this level of disaggregation. Several transportation studies, on the other hand, focus on the zonal aggregates (also called transportation analysis zones or TAZs) of the sources of attraction rather than the elemental units of attraction chosen by the decision-maker. This is usually data driven since most travel surveys only contain zonal information rather than point locations, although this trend is changing now. Also the attributes of zones are more easily available to the travel demand modeler than the attributes of the elemental units of attraction. The focus on zones as alternatives, however, is not too restrictive since the ultimate aim of travel demand models is to predict interzonal flows. Moreover, individuals are likely to perceive elemental units of attraction in clusters (such as shopping districts), and as the size of TAZs are shrinking they may actually match the perceived units of attraction.

2.6 Summary

A review of the extensive spatial choice modeling literature contributes toward the identification of the many issues associated with modeling location choice. These issues are discussed in greater detail in the following chapter.

The literature review also indicates that the behavioral decision theory is one of the best suited to develop accurate and behaviorally realistic location choice models. A sound model of location choice should combine behavioral realism in the representation of the choice process with the incorporation of spatial interaction effects. The next two chapters describe the development of a model of location choice based on behavioral decision theory that accommodates both inter- and intra-individual variations in location choice behavior due to various factors such as habit persistence, variety seeking, cognitive learning and spatial-temporal constraints, as well as different spatial interaction effects such as spatial heterogeneity, spatial autocorrelation, agglomeration and competition.

CHAPTER 3. CONCEPTUAL FRAMEWORK

The objective of this research is to develop a comprehensive conceptual framework and model structure for the choice of location for non-work activity participation. The review of the extensive literature on location choice modeling presented in the previous chapter contributes toward this objective by identifying the key issues connected with understanding location choice behavior. This chapter presents a clear picture of location choice behavior and analyzes the key issues connected with understanding location choice behavior (section 3.1), proposes a comprehensive list of factors that influence the choice of location for non-work activity participation (section 3.2), and develops a conceptual framework that is exhaustive and complete in its consideration of causal factors (section 3.3).

3.1 Understanding Location Choice Behavior

The only observable parts of individual location choice behavior in most cases are the actual choice of location, the associated circumstances (such as mode used, time of day and accompanying individuals), individual socio-demographic characteristics, and the attributes of the alternative locations. In order to clearly understand the motivations behind the observed choice, however, it is important to recognize the unobserved part of the choice behavior i.e. the cognitive processes and the resultant preferences. Understanding the linkages between cognition, preference and behavior is therefore the key to understanding location choice behavior (Pipkin, 1981).

A simple picture of the choice process is as follows. Cognitive processes in the brain are influenced by the social and spatial environments to generate a mental map of preferences, which is then influenced by constraints and social effects to generate the observed choice (see Figure 2 for an illustration). We shall now proceed to examine this picture in detail.

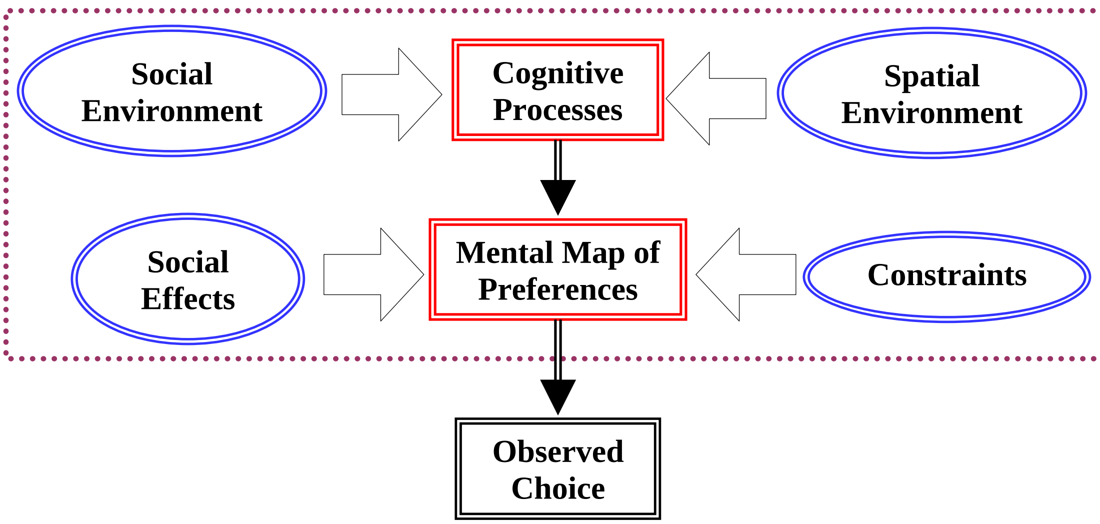

Figure 2. Effects of cognitive processes and preferences on observed choice

Cognitive processes consist of internal psychological processes such as perception (perception of space, perception of the attributes of alternatives), learning (recalling past experiences from memory, recalling the spatial structure of the surroundings) and problem solving. These cognitive processes subsume an individual's subconscious awareness of his or her situation in life, which includes not only factors independent of the choice occasion such as family structure, role in family and availability of resources, but also factors that vary with the choice occasion such as what the rest of the family is doing and whether the household vehicles are available, or whether it will be a joint activity with other family members. Cognitive processes are also influenced by the external environment, both social and spatial, to generate a mental map of preferences. The influence of the social environment manifests itself in two different forms. There are the effects of socially defined norms, preferences and prejudices on one hand, while on the other there is information interchange. The result of social norms can be observed in the way people tend to prefer neighborhoods that are similar to their own socio-economic circumstances. Information interchange, on the other hand, increases an individual's awareness of his surroundings either through conversations with friends and family or through exposure to public media. The influence of the spatial environment is a result of the vastness of the spatial environment, the directionality introduced in the spatial structure due to the transportation infrastructure and land-use variations, and the cognitive processes involved in processing all this spatial information.

The mental map of preferences that results from all the above cognitive processes is a conscious manifestation of the subconscious processes and is defined succinctly by Pipkin (1981) as the representation of environmental learning retrieved from memory. An individual's conscious preferences are further influenced by constraints such as time and income budgets, and social effects such as the preferences of other members of the party in the case of joint activity participation. The observed choice is thus a result of these constraints and social effects on the preferences.

In the process of developing a better understanding of location choice behavior, several key issues were identified that need to be considered in generating a comprehensive conceptual framework for modeling location choice. These issues are discussed in the following sections.

3.1.1 Simultaneity in decision-making

The primary decisions that individuals make to participate in activities separated over space are the location of activity participation, the choice of the mode of travel, the time of day of travel, solo versus joint travel, route choice, and the choice of trip chaining. The order in which these decisions are made is unknown to the observer and there has been no proof that any particular sequence of decisions is more likely than the other (but see Pendyala et al., 2002). In fact, it is very likely that the sequence of decisions varies across individuals and across choice occasions for each individual depending on personal preferences and constraints. It is important to keep this in mind while developing a comprehensive location choice model, since one or more of these decisions may be correlated with the choice of activity location (see Figure 3).

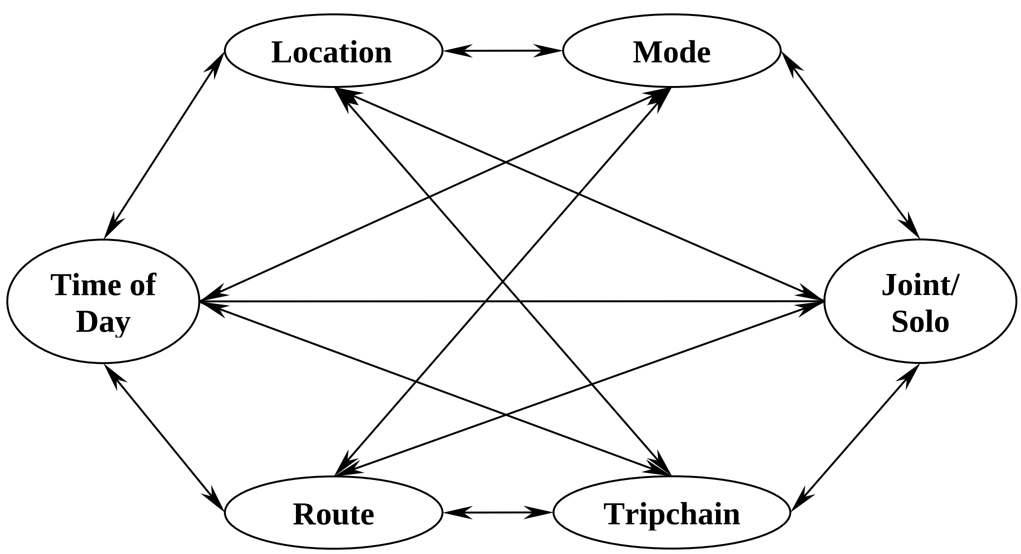

Figure 3. Simultaneity in activity participation choices

3.1.2 The "choice" issue

Another important issue to be aware of in the analysis of choice behavior of any kind is the meaning attributed to the word "choice". The observed choice of alternative is not always a direct result of individual preferences; rather, as discussed earlier, it could be the result of constraints and social effects on individual preferences. For instance, although a particular grocery store may be the most preferred for both individuals A and B, A as a single female with fewer constraints may actual choose to shop at that store while B with an inflexible work schedule and as a father of school-going children may be constrained to shop at a grocery store on the way to the school from work. In this example, B's choice is not so much an ideal "choice" as the best option under the circumstances. Similarly, a person who prefers to go to the gym with a partner may be forced to compromise on the choice of the gym depending on her partner's preferences. A sound model of location choice should attempt to incorporate the effects of such constraints on choice behavior in order to achieve behavioral realism.

3.1.3 Identifying sources of heterogeneity

The flexible and personal nature of non-work activity participation results in highly heterogeneous behavior. A good location choice model should incorporate such varied choice behavior, which is possible only if all the sources of heterogeneity (including constraints and social effects as noted in the preceding section) are identified.

The heterogeneity in choice behavior could be inter-personal or intra-personal, and it could be due to observed or unobserved factors (from the analyst's viewpoint). Interpersonal heterogeneity in non-work location choice could be a result of observed factors such as income or unobserved factors such as habit persistence. In other words, different individuals may choose different shopping locations depending on their income levels and the corresponding spending capacity. Two individuals identical in every respect (from the analyst's viewpoint) may still choose different locations if one of them exhibits habit persistence and therefore selects a location that may not have the highest utility associated with it. Intrapersonal heterogeneity in non-work location choice could also be a result of observed or unobserved factors. Observed factors such as the time of day may constrain an individual to choose a closer shopping location on a weekday, whereas the individual might prefer a different, more distant shopping location in the absence of time constraints on the weekend. Unobserved factors such as variety-seeking or boredom may also result in an individual selecting different shopping locations on consecutive choice occasions.

There are many such possible sources of heterogeneity, and it is important to identify and include all these within the conceptual framework for non-work location choice. Omitted sources of heterogeneity in choice models usually result in underspecified models, and in such situations a good conceptualization can help to correctly interpret such models.

3.1.4 Spatial issues

Travel is a result of the desire for activity participation, constrained by the spatial layout of activity opportunities and other temporal and household-level interpersonal interactions. While the temporal and household-level constraints are relatively simple to conceptualize, the spatial aspect of travel is fairly difficult to characterize. The location choice for non-work trips is, by its very definition, a spatial problem and there are several issues associated with that. First, the unit of choice for non-work location decisions is not well-defined. Take for instance, a shopping trip. An individual could choose to visit a specific store, a shopping mall, or a shopping district. The definition of the choice set for this individual is, therefore, fuzzy. Second, non-work location decisions, like most spatial choice problems, usually involve a very large set of possible alternatives regardless of the spatial definition of the alternative. If the unit of choice is a store, the potential choice set would include all the stores in the city where the decision-maker lives (or even beyond). If the unit of choice is a zonal aggregate of shopping opportunities (i.e., a zone, as in many transportation planning applications), the potential choice set would include all the zones in the planning region, which is still a large number (typically in the hundreds). Third, due to the continuum of space, the various location choice alternatives are likely to share common attractiveness factors (spatial correlation). For instance, we would expect zones that are nearer each other to be more substitutable than zones that are further away. Fourth, in addition to spatial correlation, spatial competition or agglomeration effects are also likely to be in play. In other words, some alternatives benefit from being close to other similar alternatives (agglomeration effect) as in a shopping district, while some alternatives benefit from being away from other similar alternatives (competition effect) by capturing the market in their immediate neighborhood. Fifth, the non-uniformity of spatial structures leads to spatial heterogeneity effects that must also be considered. Sixth, the development of an operational location choice model may involve some level of spatial aggregation of decision-makers (individuals or households into neighborhoods or zones) and elemental units of attraction (stores or parks into zones). In such cases it is important to ensure that the spatial effects of aggregation are accounted for. This includes the correlation between individuals who live in the same zone, and the sampling variance of individuals across zones (see Bhat, 2000).

3.1.5 Spatial information processing

In order to develop a good location choice model, the identification of all factors (observed and unobserved) that can influence location choice decisions and a grasp of the spatial effects must be combined with an understanding of how individuals process spatial information. Understanding how individuals process vast quantities of spatial information is critical in determining their perspective regarding the spatial layout of potential locations, and forms the subject of an entire area of research termed spatial cognition (see Gärling et al., 1984, Anooshian and Seibert, 1996, Fotheringham and Curtis, 1999). Individuals could either consider all possible choice alternatives at once, or they could process the spatial information hierarchically and decide upon a general spatial region prior to picking a specific alternative. The decision rule is accordingly a flat-information processing rule or a hierarchical-information processing rule. This decision rule can potentially vary across individuals and to some extent across choice occasions of an individual. In addition, the boundary defined by individuals who follow the flat-information processing rule can also vary across individuals depending on how large an area the individual can process. Among individuals who follow the hierarchical-information processing rule, the number of levels can also vary.

3.2 Factors Influencing Location Choice Decisions

In this section, a comprehensive list of all the observed and unobserved factors that influence location choice decisions are compiled based on the discussion in the previous section. The following sections enumerate each of the observed and unobserved factors, and examine them in detail.

3.2.1 Time invariant individual preferences

The first factor to account for in location choice decisions is the time invariant preferences that individuals possess for specific locations, attributes of locations or general neighborhoods. These preferences could be due to the following observed and unobserved factors.

1. Individual socio-demographics: Socio-demographic characteristics of individuals (such as race, age, income, gender, marital status, and employment status) can influence their preferences for specific locations, attributes of locations or general neighborhoods. For instance, low-income individuals may prefer certain shopping malls and neighborhoods for their cost-effectiveness, and younger individuals may prefer "hip" shopping districts. It is also conceivable that individuals belonging to a specific socio-demographic group may prefer a specific location for no apparent reason. Although this sounds unreasonable, it is possible that the true reason for this preference is not observable.

2. Unobserved intrinsic preferences: Individuals can also possess unobserved intrinsic preferences for certain attributes of locations, specific locations or general neighborhoods. The preference that individuals attach to the familiarity of visiting the

same location over and over is known as habit persistence or ‘loyalty’. Individuals may exhibit loyalty to specific alternatives or to general neighborhoods. For instance, person A may prefer to visit the exact same grocery store every time, while person B prefers to visit any of the stores in a chain, and person C prefers to visit stores in a particular part of the city. It is also possible that individuals possess intrinsically higher or lower preferences for the attributes of the alternatives. So, while person A prefers larger shopping malls for the variety and browsing opportunities they afford, person B may not prefer large shopping malls as she prefers quick and efficient shopping trips.

3.2.2 Time invariant attractiveness of location attributes

Another important factor that contributes towards location choice decisions is the time invariant attractiveness of the choice alternatives, and the decision-maker’s perception of their attractiveness. For instance, in the case of the choice of a shopping district the attractiveness of a location for an individual could be a function of the following factors:

1. Distance of the location from the individual’s home or work locations,
2. Size of the shopping district and the number of shopping opportunities by product type,
3. Average cost of products sold in the district,
4. The variety of brands of each product available in the district,
5. Quality of products available in the district,
6. Availability of parking within the district and parking costs,
7. Accessibility to the district by public transit modes,

8. Presence of cafes, restaurants and cinema theaters in the district,
9. General atmosphere in the district – hip/children and family friendly/old-fashioned – determined by type of stores, presence of benches and play pens etc.,
10. Racial composition of store owners and employees,
11. Hours of operation of the shops in the district, and
12. Safety.

Some of these factors, such as the cost and quality of products sold or even safety in the neighborhood, may vary with time. In such cases the individual's perception of the reliability of the shopping district is an important component of the attractiveness of the location.

3.2.3 Time variant individual preferences

The time variant preferences of individuals are far more difficult to identify though they may play just as important a role as the other factors. The observed and unobserved factors that contribute to these preferences are listed below:

1. Variety-seeking: Some individuals tire of familiarity and soon reach a point of satiation or boredom either with specific alternatives, with the attributes of alternatives or with entire neighborhoods. Some other individuals may seek variety in a quest to identify a preferred alternative. The degree of variety seeking can vary across individuals and across choice occasions of an individual.

2. Unfulfilled desires: The preferences of an individual for a specific alternative, for the attributes of alternatives or for a neighborhood can also be influenced by unfulfilled desires carried over from past choice occasions. For example, a person who

wished to visit a specific store on a previous occasion but could not do so due to time constraints will show a high preference to visit the store on the next choice occasion.

3. Desire to travel: The preferences of an individual may also be influenced by a desire to travel. In such a case, the person exhibits a higher preference for alternatives that are either located further away and/or are located in the path of scenic neighborhoods, despite the availability of similar alternatives that are located closer.

The degree of habit persistence in individuals can also vary across individuals and across choice occasions of an individual. So some people are more loyal and visit the same store every single time, while others are less loyal and visit a store only every eight times out of ten. The less loyal individuals may sometimes be influenced by unfulfilled desires or a desire for travel or variety-seeking. The varying preferences of an individual could also be a result of the building up of loyalty over time.

3.2.4 Time variant attractiveness of location

The time variant attractiveness of alternatives can be a function of several factors, some of which are easily identified. The following is a list of these factors.

1. Special attraction variables: Special shows at museums and sales at shopping malls temporarily increase the attractiveness of the alternative. The preference of an individual for these ‘special attractions’, however, can vary across persons and across choice occasions of the same person. So some people are more attracted to sale offers in shops than others, while some might give in to the temptation of a sale every once in a while.

2. Advertising: Advertisements through the media, posters and flyers may increase or decrease the attractiveness of alternatives. They could have the effect of increasing an individual’s awareness of his surroundings by introducing him to opportunities whose existence he might not have been aware of. They could also have an effect at a subconscious level by slowly ingraining either a desire to visit or a dislike of a particular choice alternative.

3. Spatial learning: People extend the boundaries of their spatial knowledge when they travel and socialize. The resultant increase in knowledge of new locations and opportunities can also increase the attractiveness of some alternatives. For instance, a grocery shop in a neighborhood hitherto unvisited by person A would be fairly unattractive to the person as an alternative. But if person A happens to meet friends at her local club who like the store very much and share their ideas with her, the attractiveness of the store for person A can be expected to increase.

4. Traffic conditions: The traffic conditions at the time of choice, and the corresponding travel time to a location also influence the attractiveness of the location. So if a person can only visit the grocery store during the evening peak, he can be expected to visit a store situated either near his home or work locations.

3.2.5 Constraints

In addition to the various factors discussed thus far individuals are also influenced in their location choice by the constraints that come into play during each choice occasion. These constraints are listed below.

1. Time-budget: Individuals are restricted in their travel-related decisions by the availability of time. So depending on when a person has decided to travel, he may be more or less sensitive to the distance of the alternative. Time budget constraints vary by time of day and day of week. It has been observed, for instance, that people are typically less constrained for time during the off-peak hours and weekends.

2. Mode choice: The choice of mode may also constrain the choice of activity location. So individuals who do not have access to a personal automobile may be restricted to either shop at locations that are accessible by public transport or shop during times when they can get a ride from a friend.

3. Trip-chaining: The decision to trip chain invariably influences the choice of activity location, since individuals typically select locations so that they can accomplish all their tasks conveniently. For instance, if a person has decided to combine grocery shopping with a social visit to her friend she will be more likely to visit a grocery store either near her residence or near her friend's residence or somewhere between the two locations.

3.2.6 Other decision-makers

The last factor to consider is the presence of other persons with the decision-maker. All along we have discussed the various factors that affect the choices made by the decision-maker. These discussions were based on the assumption that either the decision-maker will be traveling alone, or if there will be multiple individuals traveling together the decision-maker is the primary decision-maker in the group whose preferences dominate that of the others. However, if the preferences of one or more of the other persons traveling with the decision-maker are important the dynamics of the choice process will be significantly different. All the factors discussed in the previous sections still hold, but the preferences of the many individuals will now have to be weighed against each other combined with information regarding the weight carried by each.

3.3 Conceptual Framework

The conceptual framework for non-work location choice presented in this section includes all the observed and unobserved factors discussed in section 3.2. This comprehensive framework thus attempts to incorporate the various factors and processes (contained within the dotted box in Figure 2) that contribute to the observed choice of location.

Figure 4 is an illustration of the proposed conceptual framework. The time invariant factors (in dotted boxes) influence the mental map of the individual on every choice occasion while the time variant factors (in solid boxes), and the spatial information processing rule (in the double-outlined box) are specific to each choice occasion. All these factors are considered simultaneously, to represent the cognitive processes integrated with the effects of the social and spatial environments and subject to constraints and social effects, to generate an observed choice of location. The chosen alternative, in turn influences future choices as individuals' preferences adapt to past experiences. The entire mental map on choice occasion t also influences the mental map on choice occasion t+1, since past preferences (whether satisfied or not) are a part of an individual's memory and therefore cognition.

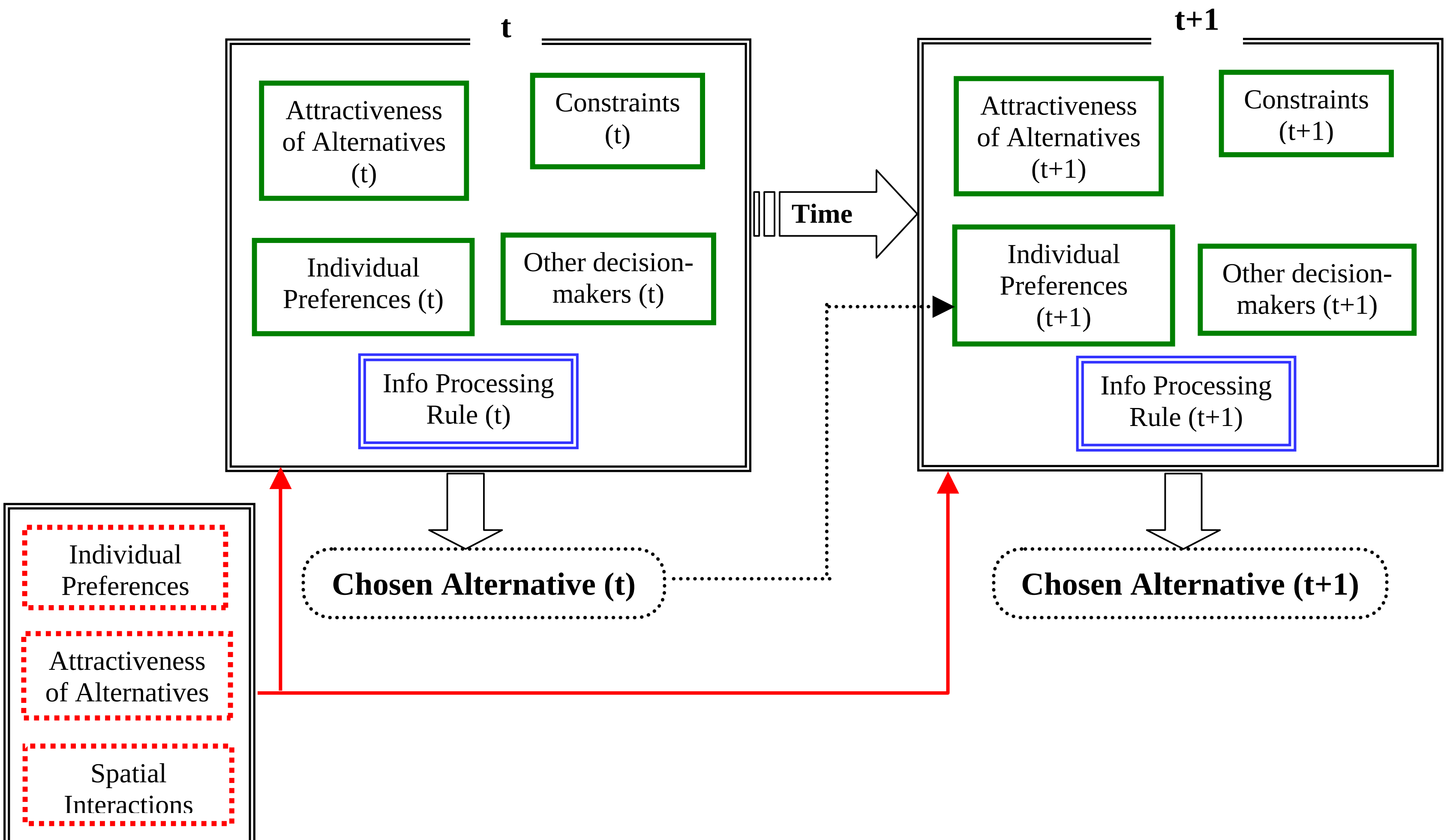

Figure 4. Conceptual framework of the location choice for non-work activity participation

CHAPTER 4. MODEL STRUCTURE

As discussed in chapter 2, several different methodologies have been used in the literature for modeling location choice. The objective of this dissertation, of developing a comprehensive, unified, choice model structure, is achieved using the random utility maximization framework with revealed preference (RP) data. Although a combination of stated preference (SP) and RP data is ideal to examine choice behavior and develop predictive as well as behaviorally realistic models of location choice, most planning efforts typically have access only to RP data. A sound conceptual understanding of location choice behavior, combined with a powerful model structure based on RP data that can accommodate the varied choice behaviors and cognitive processes, can compensate for the disadvantage of the lack of SP data. The conceptual framework presented in the previous chapter contributes toward this objective. In this chapter, the conceptual framework is supplemented with a powerful model structure that can be used to estimate accurate and behaviorally realistic location choice models.

The following sections discuss the reasons for the choice of a random utility maximization framework for the model structure (section 4.1); present the proposed model in its entirety and discuss, with examples, the ability of the proposed model to subsume a variety of choice behaviors (section 4.2); indicate the several well-known forms of models nested within the proposed model (section 4.3); and finally describe the estimation procedure for the proposed model (section 4.4).

4.1 Random Utility Maximization (RUM)

The random utility maximization (RUM) framework is based on the simple premise that the utility (U_{ijt}) an individual i attributes to a choice alternative j on choice occasion t consists of a systematic, deterministic component (V_{ijt}) and a random error component (ε_{ijt}). In other words,

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt}, \qquad \textbf{Eq. 11}$$

and individuals are assumed to choose the alternative that maximizes their utility. From a philosophical perspective it is intuitive to perceive individuals as utility maximizing entities provided we make allowances for mathematically non-rational behavior such as variety-seeking, habit persistence, social influences and satisficing. The important point here is that such “non-rational” behavior as variety-seeking and habit persistence actually enhances the utility perceived by individuals for specific alternatives. So provided the specification of the utility function can include varied kinds of behavior, the utility maximization principle is an intuitive approach to modeling choice behavior. Moreover, the presence of the random error component adds to the behavioral realism of the utility function since it ensures that two observationally identical individuals can differ in their choice behavior.

This dissertation research uses a mixed logit model structure based on the RUM framework to develop the comprehensive location choice model (refer Bhat, 2002, for a description of the mixed logit model). The mixed logit model structure is a powerful and flexible tool that utilizes the potential of the RUM framework. For instance, it is capable of handling a very large number of choice alternatives (Bhat and Guo, 2004), incorporating spatial interaction effects (Bhat, 2000), including various sources of observed and unobserved heterogeneity (Train, 1998) and incorporating evolutionary learning behavior through feedback (Miller and O'Kelly, 1983). The only issue with using a mixed logit model structure is the associated computational burden of estimation. Chapters 5 and 6 discuss this issue in greater detail and present the results of the research efforts to improve the efficiency of the estimation process.

4.2 Proposed Location Choice Model

The proposed location choice model expresses the utility that an individual i (i = 1,…,I) associates with an alternative j (j = 1,…,J_i) on choice occasion t (t = 1,…,T_i), as

$$\begin{aligned} U_{ijt} &= Z_j(\alpha_i + \eta_{it} + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij}(\beta_i + \gamma_{it} + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i) + (\xi_i\ )L_{jt} \\ &+ PREATT_{ijt}(\chi_i^0 + \chi_{it}^1 + \chi^2 X_i) + PRECHO_{ijt}(\zeta_i^0 + \zeta_{it}^1 + \zeta^2 X_i) \\ &+ \lambda_0[\widetilde{U}_{ij(t-1)} + \lambda_1 \widetilde{U}_{ij(t-2)} + \lambda_1^2 \widetilde{U}_{ij(t-3)} + \ldots \lambda_1^t \widetilde{U}_{ij1}] + \varepsilon_{ijt} + \rho \sum_{j' \in J'} \varepsilon_{ij't} \qquad \text{Eq. 12} \end{aligned}$$

where, Z_j is a vector of observed time invariant attributes of zone j,

X_i is a vector of observed socio-demographic attributes of individual i,

C_{it} is a vector of characteristics of choice occasion t for individual i (including constraints faced by the individual),

D_{ij} is a matrix of distance or time and cost measures between i and j,

L_{jt} is a vector of special attraction variables associated with alternative j on choice occasion t,

$PREATT_{ijt}$ is a function of the similarities between the attributes of previously chosen alternatives (on choice occasions *t-1, t-2,…,1*) and alternative j, *PRECHO*$_{ijt}$ is a function of the number of times alternative *j* has been chosen on choice occasions *t*-1, *t*-2,…,1,

$\tilde{U}_{ij(t-1)}$, $\tilde{U}_{ij(t-2)}$,… are the utilities that individual *i* associated with alternative *j* on choice occasions prior to occasion *t* excluding the effects of constraints, and

$\{\alpha_i, \eta_{it}, \delta_1, \delta_2, \delta_3, \beta_i, \gamma_{it}, \omega_1, \omega_2, \omega_3, \xi_i, \chi_i^0, \chi_{it}^1, \chi^2, \zeta_i^0, \zeta_{it}^1, \zeta^2, \lambda_0, \lambda_1, \rho\}$ are the parameters of the model that are explained in the following paragraphs.

The term $\alpha_i + \delta_1 X_i$ represents the vector of time invariant preferences (or dispreferences) of individual *i* for the attributes Z_j of the choice alternative (the importance of the time varying terms η_{it} and γ_{it} will de discussed toward the end of this section). The vector of parameters δ_1 represents the extent of the preferences that can be captured by observed socio-demographic characteristics of the individual, while α_i represents the unobserved preferences of the individual that makes her choice behavior different from that of an observationally identical individual. The vector of parameters α_i, therefore, accounts for inter-personal response heterogeneity that is not accounted for by other factors such as variety-seeking and the desire for travel. The term $\beta_i + \omega_1 X_i$, similarly, represents the vector of time invariant preferences (or dispreferences) of individual *i* for the time and costs, D_{ij}, associated with the choice alternative.

The parameter ξ_i represents the time invariant preferences of individual *i* for the special attractions associated with alternative *j* on choice occasion *t*. For instance, if a shopping mall has a big sale, the individual might want to visit that mall on that particular occasion. Constraints might, however, bring the utility of the mall down despite the ‘special attraction’. The vectors of parameters $(\delta_2, \delta_3, \omega_2, \omega_3)$ represent the effects of constraints on individual *i*. This could include time budget, trip chaining and mode availability constraints. So, given an individual’s time schedule and availability, the choice alternatives that require longer distances of travel might incur a disutility large enough to overcome their utility due to all other kinds of preferences.

The terms $(\chi_i^0 + \chi_{it}^1 + \chi^2 X_i)$ and $(\zeta_i^0 + \zeta_{it}^1 + \zeta^2 X_i)$, and the parameters (λ_0, λ_1) represent the time variant preferences of individual *i* that are a result of learning, variety seeking and unfulfilled desires, respectively. The term $(\chi_i^0 + \chi_{it}^1 + \chi^2 X_i)$ represents the preference of individual *i* for alternative *j* that is due to the degree of similarity in attributes between *j* and other alternatives chosen by the individual on previous choice occasions. For instance, if alternative *j* is assigned a higher utility due to its proximity to other recently chosen locations (all else being identical) the hypothesis is that it is the effect of spatial learning. A higher preference exhibited for alternative *j* due to its similarity in some other attribute (such as size of the store, in the case of store choice) with recently chosen locations could, on the other hand, be the result of habit persistence in preference for that particular attribute. The term $(\zeta_i^0 + \zeta_{it}^1 + \zeta^2 X_i)$ represents the preference of individual *i* for alternative *j* due to effects of previous choice occasions when *j* was chosen. This captures variety seeking in choice of alternative. An individual who exhibits habit persistence is likely to have a higher preference for locations he has visited in the past, while one who exhibits variety seeking is likely to have a lower preference for locations she has visited in the past. The term

$\lambda_0[\tilde{U}_{ij(t-1)} + \lambda_1\tilde{U}_{ij(t-2)} + \lambda_1^2\tilde{U}_{ij(t-3)} + \ldots\lambda_1^t\tilde{U}_{ij1}]$ represents the carryover effects and unfulfilled desires from past choice occasions on the utility individual *i* associates with alternative *j*. The terms $\tilde{U}_{ij(t-1)}$, $\tilde{U}_{ij(t-2)}, \ldots,$ $\tilde{U}_{ij1}$ are the utilities that individual *i* associated with alternative *j* on choice occasions prior to occasion *t*, excluding the effects of constraints.

The effects of any other factors (that have not already been accounted for) that cause intra-personal heterogeneity in observed choices are captured in the utility function by η_{it} and γ_{it}, the time variant preferences of the individual for the attributes of the alternative and for the travel time and costs associated with the alternative.

The term ε_{ijt} is the random error component of the utility individual *i* attributes to alternative *j* on choice occasion *t*. The inclusion of the term $\rho\sum_{j^{'}\in J^{'}}\varepsilon_{ij^{'}t}$ captures the spatial correlation of alternative *j* with other choice alternatives that are adjacent to *j* (represented by the set *J'*), with the parameter ρ capturing the degree of spatial correlation.

The proposed location choice model is thus a mixed logit model that accommodates spatial interaction effects, and response heterogeneity due to various observed and unobserved factors including state dependent effects such as variety seeking, habit persistence, carryover effects and spatial learning.

4.3 Models Nested Within the Proposed Model Structure

As discussed in the previous section, the proposed location choice model in equation 12 can capture varied choice behaviors including variety seeking and habit persistence, and incorporate various factors such as that of unfulfilled desires, spatial learning, spatial interaction, temporal and modal constraints and special attraction variables. Different assumptions imposed on this model will, therefore, result in simpler (restricted) models that represent specific circumstances or behavior types. Some of these simpler models are presented in the following sections.

4.3.1 Multinomial Logit Model

The simple multinomial logit model (MNL) of location choice is a very special case of the general model presented in equation 12. The MNL is a result of the following assumptions on our general location choice model.

$$\alpha_i = \alpha, \beta_i = \beta , \ \ \xi_i = \xi \qquad \textbf{Eq. 13}$$

$$\eta_{it} = \gamma_{it} = 0 \qquad \textbf{Eq. 14}$$

$$\chi_i^0 = \chi_{it}^1 = \chi^2 = \zeta_i^0 = \zeta_{it}^1 = \zeta^2 = \lambda_0 = \lambda_1 = 0 \qquad \textbf{Eq. 15}$$

$$\rho = 0 \qquad \textbf{Eq. 16}$$

These are basically the assumptions of no inter-personal heterogeneity (equation 13), no intra-personal heterogeneity (equation 14), no state dependence or carryover effects (equation 15) and no spatial correlation (equation 16). The resulting model is the familiar MNL utility expression.

$$U_{ijt} = Z_j (\alpha + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij} (\beta + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i) + \xi L_{jt} + \varepsilon_{ijt} \qquad \textbf{Eq. 17}$$

The MNL model is widely applied in the literature in various fields. Some of the applications to modeling location choice include Recker and Schuler, 1981, Shukla and Waddell, 1991, and Sermons and Koppelman, 1998, 2001.

4.3.2 First-order State Dependence Model

The assumption that only the immediately previous choice of individuals has an effect on their choice behavior is known as first-order state dependence. This assumption coupled with the MNL assumptions (equations 13-16) results in the classic first-order state dependence model with a variable to indicate the previous choice. This variable is typically just a dummy variable indicating whether the specific alternative was the chosen alternative on the previous occasion ($SAME_{ijt}$). The assumptions in this model are therefore

$$\alpha_i = \alpha, \quad \beta_i = \beta, \quad \xi_i = \xi, \quad \zeta_i^0 = \zeta^0, \qquad \textbf{Eq. 18}$$

$$\eta_{it} = \gamma_{it} = 0, \qquad \textbf{Eq. 19}$$

$$\chi_i^0 = \chi_{it}^1 = \chi^2 = \zeta_{it}^1 = \zeta^2 = \lambda_0 = \lambda_1 = 0 \text{ , PRECHOijt = SAMEijt , and} \qquad \textbf{Eq. 20}$$

$$\rho = 0, \qquad \textbf{Eq. 21}$$

which result in the following first-order model.

$$U_{ijt} = Z_j(\alpha + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij}(\beta + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i)$$

$$+ \xi L_{jt} + \zeta^0 SAME_{ijt} + \varepsilon_{ijt} \qquad \textbf{Eq. 22}$$

The pure first-order state dependence model in equation 22 can be further extended to incorporate spatial interaction effects and unobserved heterogeneity, if desired (Miller and O'Kelly, 1983, for instance, estimate a first order state dependence model with random parameter heterogeneity). Conversely, under the assumption of *memory-less choice behavior* equation 22 collapses to equation 17 (the MNL model), which is the case when the past choices of individuals have no effect on their choice behavior.

4.3.3 Spatially Correlated Logit Model

If the assumption of zero spatial correlation in the MNL model of equation 17 is relaxed, we get the following spatially correlated logit (SCL) model.

$$U_{ijt} = Z_j(\alpha + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij}(\beta + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i) + \xi L_{jt} + \varepsilon_{ijt} + \rho \sum_{j' \in J'} \varepsilon_{ij't}$$

Eq. 23

The use of a Generalized Extreme Value (GEV) structure results in substantial computational efficiency gains in the estimation of the SCL model. Bhat and Guo (2004) formulated and applied an SCL model that uses a GEV-based structure to accommodate correlation in the utility of spatial units.

4.3.4 Mixed Logit Model

The MNL model of equation 17, with the relaxation of the zero unobserved inter-individual heterogeneity assumption, yields the following random parameters Mixed Logit (MxL) model.

$$U_{ijt} = Z_j(\alpha_i + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij}(\beta_i + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i) + \xi_i L_{jt} + \varepsilon_{ijt} \quad \textbf{Eq. 24}$$

There have been several applications of this model in the literature, including specific applications to modeling location choice such as Train (1998) and Kemperman et al. (2004).

4.3.5 Mixed Spatially Correlated Logit Model

The relaxation of both the assumptions of zero spatial correlation and zero unobserved inter-individual heterogeneity in the MNL model of equation 17, yields the Mixed Spatially Correlated Logit Model (MSCL).

$$U_{ijt} = Z_j(\alpha_i + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij}(\beta_i + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i) + \xi_i L_{jt} + \varepsilon_{ijt} + \rho \sum_{j' \in J'} \varepsilon_{ij't}$$

Eq. 25

Alongside the SCL model described in section 4.3.3, Bhat and Guo (2004) also formulated and applied an MSCL model of residential location choice that uses a GEV-based structure to accommodate correlation in the utility of spatial units and superimposes a mixing distribution over the GEV structure to capture unobserved response heterogeneity.

4.3.6 Bi-level Mixed Logit Model

The MNL model of equation 17, with the relaxation of the zero unobserved inter- and intra-individual heterogeneity assumptions, yields the following Bi-level Mixed Logit (BiMxL) model.

$$U_{ijt} = Z_j(\alpha_i + \eta_{it} + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij}(\beta_i + \gamma_{it} + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i) + \xi_i L_{jt} + \varepsilon_{ijt}$$

Eq. 26

Bhat and Castelar (2002) formulated and applied a unified mixed-logit framework for the joint analysis of revealed and stated preference data in their paper. Although their model uses a bi-level integration technique it differs in structure from the bi-level model presented here.

4.3.7 Bi-level Mixed Spatially Correlated Logit Model

The relaxation of the assumptions of zero spatial correlation and zero unobserved inter- and intra-individual heterogeneity in the MNL model of equation 17, yields the following Bi-level Mixed Spatially Correlated Logit (BiMSCL) model.

$$U_{ijt} = Z_j(\alpha_i + \eta_{it} + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij}(\beta_i + \gamma_{it} + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i) +$$

$$\xi_i L_{jt} + \varepsilon_{ijt} + \rho \sum_{j^{'} \in J^{'}} \varepsilon_{ij^{'}t} \quad \textbf{Eq. 27}$$

The estimation of this model can be achieved using a GEV-based structure to accommodate correlation in the utility of spatial units, superimposed with a two-level mixing distribution, one to capture unobserved intra-individual response heterogeneity and the other for unobserved inter-individual response heterogeneity.

4.4 Model Estimation

The vector of parameters to be estimated in a location choice model based on the proposed model structure is, as seen in the previous sections, some subset of $\{\alpha_i, \eta_{it}, \delta_1, \delta_2, \delta_3, \beta_i, \gamma_{it}, \omega_1, \omega_2, \omega_3, \xi_i, \chi_i^0, \chi_{it}^1, \chi^2, \zeta_i^0, \zeta_{it}^1, \zeta^2, \lambda_0, \lambda_1, \rho\}$. Of these parameters, $\{\alpha_i, \beta_i, \xi_i, \chi_i^0, \zeta_i^0\}$ vary across individuals and capture unobserved inter-individual response heterogeneity, while $\{\eta_{it}, \gamma_{it}, \chi_{it}^1, \zeta_{it}^1\}$ vary across choice occasions of an individual and capture unobserved intra-individual response heterogeneity. For convenience, let $\Psi = \{\alpha_i, \beta_i, \xi_i, \chi_i^0, \zeta_i^0\}$, $\Omega = \{\eta_{it}, \gamma_{it}, \chi_{it}^1, \zeta_{it}^1\}$ and μ represent the rest of the fixed response parameters $\{\delta_1, \delta_2, \delta_3, \omega_1, \omega_2, \omega_3, \chi^2, \zeta^2, \lambda_0, \lambda_1\}$. ρ is the dissimilarity parameter that captures the degree of spatial correlation. Let the distribution of unobserved inter- and intra-individual heterogeneities be multivariate normal, so that the elements of Ψ and Ω are realizations of the random multivariate normally distributed variables that comprise $\tilde{\Psi}$ and $\tilde{\Omega}$ respectively. Let θ be a vector of true parameters characterizing the mean and variance-covariance matrix of $\tilde{\Psi}$, and σ be a vector of true parameters characterizing the mean and variance-covariance matrix of $\tilde{\Omega}$.

In its most general form, the utility associated by individual i with zone j on choice occasion t is given by $U_{ijt} = V_{ijt} + \varepsilon_{ijt}$, where

$$V_{ijt} = Z_j(\alpha_i + \eta_{it} + \delta_1 X_i + \delta_2 C_{it} + \delta_3 C_{it} X_i) + D_{ij}(\beta_i + \gamma_{it} + \omega_1 X_i + \omega_2 C_{it} + \omega_3 C_{it} X_i) + (\xi_i\ )L_{jt}$$
$$+ PREATT_{ijt}(\chi_i^0 + \chi_{it}^1 + \chi^2 X_i) + PRECHO_{ijt}(\zeta_i^0 + \zeta_{it}^1 + \zeta^2 X_i)$$
$$+ \lambda_0[\tilde{U}_{ij(t-1)} + \lambda_1 \tilde{U}_{ij(t-2)} + \lambda_1^2 \tilde{U}_{ij(t-3)} + \ldots \lambda_1^t \tilde{U}_{ij1}] \qquad \textbf{Eq. 28}$$

As per the notations, the parameters $\{\alpha_i, \beta_i, \xi_i\ , \chi_i^0, \zeta_i^0\}$ and $\{\eta_{it}, \gamma_{it}, \chi_{it}^1, \zeta_{it}^1\}$ in the above expression are drawn from the random variables that comprise $\tilde{\Psi}$ and $\tilde{\Omega}$. V_{ijt} may therefore be represented as $V_{ijt}(\tilde{\Psi}, \tilde{\Omega}, \mu)$. Now the probability function depends on the spatial correlation assumption.

Under the assumption of no spatial correlation, the probability that individual i will choose alternative j at the t^{th} choice occasion, conditional on $\tilde{\Psi}$, $\tilde{\Omega}$ and μ, is the usual multinomial logit form (see McFadden, 1978):

$$P_{ijt} \mid (\tilde{\Psi}, \tilde{\Omega}, \mu) = \frac{e^{V_{ijt}(\tilde{\Psi}, \tilde{\Omega}, \mu)}}{\sum_{k=1}^{J} e^{V_{ikt}(\tilde{\Psi}, \tilde{\Omega}, \mu)}} \qquad \textbf{Eq. 29}$$

The assumption of spatial correlation, on the other hand, combined with a GEV-based structure leads to the following expression for the probability that individual i will choose alternative j at the t^{th} choice occasion, conditional on $\tilde{\Psi}$, $\tilde{\Omega}$, μ and ρ (see Bhat and Guo, 2004). For ease in presentation, let us absorb ρ, the scalar that represents spatial correlation, into the parameter vector μ.

$$P_{ijt} \mid (\tilde{\Psi}, \tilde{\Omega}, \mu) = \frac{\sum_{m \neq j} (\alpha_{j,jm} e^{V_{ijt}(\tilde{\Psi},\tilde{\Omega},\mu)})^{1/\rho} [(\alpha_{j,jm} e^{V_{ijt}(\tilde{\Psi},\tilde{\Omega},\mu)})^{1/\rho} + (\alpha_{m,jm} e^{V_{imt}(\tilde{\Psi},\tilde{\Omega},\mu)})^{1/\rho}]^{\rho-1}}{\sum_{k=1}^{J-1} \sum_{l=j+1}^{J} [(\alpha_{k,kl} e^{V_{ikt}(\tilde{\Psi},\tilde{\Omega},\mu)})^{1/\rho} + (\alpha_{l,kl} e^{V_{ilt}(\tilde{\Psi},\tilde{\Omega},\mu)})^{1/\rho}]^{\rho}}, \quad \textbf{Eq. 30}$$

where $\alpha_{i,ij}$ is an allocation parameter.

The unconditional probability can be obtained thereafter as:

$$P_{ijt} = \int_{\tilde{\beta}=-\infty}^{\infty} \int_{\tilde{\Omega}=-\infty}^{\infty} (P_{ijt} \mid \tilde{\Psi}, \tilde{\Omega}, \mu) dF(\tilde{\Omega} / \sigma) dF(\tilde{\Psi} / \theta) \quad \textbf{Eq. 31}$$

where F is the multivariate cumulative normal distribution. The dimensionality of the above integration is dependent on the number of elements in the Ψ and Ω vectors.

Therefore, the parameters to be estimated under the assumption of zero spatial correlation are the σ, θ and μ vectors corresponding to equations 29 and 31. Whereas, the assumption of spatial correlation yields the model represented by equations 30 and 31, and the parameters to be estimated include the scalar ρ representing spatial correlation, which has been absorbed into the vector of fixed response parameters μ, and the vectors σ and θ characterizing the multivariate normal distributions of the parameters in Ω and Ψ. To develop the likelihood function for parameter estimation, we need the probability of each sample individual *i*'s sequence of observed choices on choice occasions $1, \ldots T_i$. Conditional on $\tilde{\Psi}$, the likelihood function for individual *i*'s observed sequence of choices is:

$$L_i\left(\tilde{\Psi}, \sigma, \mu\right) = \prod_{t=1}^{T_i} \left[\int_{\tilde{\Omega}=-\infty}^{+\infty} \left\{ \prod_{j=1}^{J} P_{ijt}\left(\tilde{\Psi}, \tilde{\Omega}, \mu\right) Y_{ijt} \right\} f(\tilde{\Omega} \mid \sigma) d\tilde{\Omega} \right], \quad \textbf{Eq. 32}$$

where, Y_{ijt} takes the value 1 if individual i chose alternative j on choice occasion t, and 0 otherwise.

The unconditional likelihood function of the choice sequence is:

$$L_i(\theta,\sigma,\mu) = \int L_i(\tilde{\Psi},\sigma,\mu) f(\tilde{\Psi} \mid \theta) d\tilde{\Psi}$$

$$= \int_{\tilde{\Psi}=-\infty}^{+\infty} \left\{ \prod_{t=1}^{T_i} \left[\int_{\tilde{\Omega}=-\infty}^{+\infty} \left\{ \prod_{j=1}^{J} P_{ijt}\left(\tilde{\Psi},\tilde{\Omega},\mu\right) Y_{ijt} \right\} f(\tilde{\Omega} \mid \sigma) d\tilde{\Omega} \right] \right\} f(\tilde{\Psi} \mid \theta) d\tilde{\Psi} \qquad \textbf{Eq. 33}$$

The log-likelihood function is $L(\theta,\sigma,\mu) = \sum_i \ln L_i(\theta,\sigma,\mu)$.

The likelihood function in equation 33 is quite different from those in previous applications of the mixed logit model, such as Bhat (1998), Hensher (2001), and Brownstone and Train (1999). In particular, there are two levels of integration rather than one. This arises because, from an estimation standpoint, the random coefficients formulation that accommodates taste variations within individuals across choice occasions operates at the choice level, while the random coefficients formulation that accommodates taste variation across individuals operates at the individual level.

Quasi-Monte Carlo (QMC) simulation techniques are applied to approximate the integrals in the likelihood function and maximize the logarithm of the resulting simulated likelihood function across all individuals with respect to θ, σ and μ. The procedure to simulate each individual's likelihood function $L_i(\theta,\sigma,\mu)$, is as follows: (a) For a given value of the parameter vector θ, draw a particular realization of $\tilde{\Psi}$ from its distribution, (b) For a given value of the σ vector, draw several sets of realizations of $\tilde{\Omega}$ from its distribution, each set corresponding to a choice occasion of the individual, (c) compute the probability of the chosen alternative for each choice occasion (*i.e.*, the likelihood function of that choice occasion) at that choice occasion's set of $\tilde{\Omega}$ realizations, and for the current $\tilde{\Psi}$ realization, (d) Average the likelihood functions across the various realizations of $\tilde{\Omega}$ for each choice occasion, (e) Compute the individual likelihood function as the product of the averaged likelihood functions across all choice occasions of the individual, (f) Repeat steps a through e several times with fresh realizations of $\tilde{\Psi}$ and new sets of draws of $\tilde{\Omega}$, and (g) Compute the average across all individual likelihood function evaluations. Mathematically, the individual likelihood function is approximated as:

$$SL_i(\theta,\sigma,\mu)=\frac{1}{N}\sum_{n=1}^{N}\left[\prod_{t=1}^{T_i}\left\{\frac{1}{M}\sum_{g_n=1}^{M}\left(\prod_{j=1}^{J}P_{ijt}(\tilde{\Psi}^n\,|\,\theta,\tilde{\Omega}^{g_n}\,|\,\sigma,\mu)^{Y_{ijt}}\right)\right\}\right], \qquad \textbf{Eq. 34}$$

where $SL_i(\theta,\sigma,\mu)$ is the simulated likelihood function for the i^{th} individual's sequence of choices given the parameter vectors θ, σ and μ, $\tilde{\Psi}^n\,|\,\theta$ is the n^{th} draw (n=1,2,…,N) from $f(\tilde{\Psi}\,|\,\theta), \tilde{\Omega}^{g_n}\,|\,\sigma$ is the g_n^{th} draw (g_n=1,2,…,M) from $f(\tilde{\Omega}\,|\,\sigma)$ at the n^{th} draw of $\tilde{\Psi}$, and other variables are as defined earlier. $SL_i(\theta,\sigma,\mu)$ is an unbiased estimator of the actual likelihood function $L_i(\theta,\sigma,\mu)$. Its variance decreases as N and M increase. It also has the appealing properties of being smooth (*i.e.*, twice differentiable) and being strictly positive for any realization of the draws.

The simulated log-likelihood function is constructed as:

$$SL(\theta, \sigma, \mu) = \sum_i \ln\left[SL_i(\theta, \sigma, \mu)\right] \qquad \textbf{Eq. 35}$$

The parameter vectors θ, σ and μ are estimated as the values that maximize the above simulated function. Under rather weak regularity conditions, the simulated maximum likelihood estimator is consistent, asymptotically efficient, and asymptotically normal (see Hajivassiliou and Ruud, 1994; Lee, 1992; McFadden and Train, 1998).

Depending on the number of parameters in θ and σ, and the number of draws *N* and *M*, however, the simulated maximum likelihood estimation of this bi-level model can be very time consuming. Most applications of mixed logit models in the literature use QMC sequences, such as the Halton sequence, to draw realizations for $\tilde{\Psi}$ and $\tilde{\Omega}$ from their normal population distributions. Although Halton sequences are a vast improvement over pseudo-Monte Carlo (PMC) methods in the efficiency of the simulated estimation process, there are several other QMC sequences that are potentially superior to the Halton sequences. The following two chapters present the results of research undertaken to identify a more efficient QMC sequence for the purpose.

CHAPTER 5. QUASI-MONTE CARLO SEQUENCES

5.1 Background

The incorporation of behavioral realism in econometric models helps establish the credibility of the models outside the modeling community, and can also lead to superior predictive and policy analysis capabilities. As demonstrated in the previous chapter, behavioral realism is incorporated in econometric models of choice through the relaxation of restrictions that impose inappropriate behavioral assumptions regarding the underlying choice process. For example, the extensively used multinomial logit (MNL) model has a simple form that is achieved by the imposition of the restrictive assumption of independent and identically distributed error structures (IID), which leads to the not-so-intuitive property of independence from irrelevant alternatives (IIA).

The relaxation of behavioral restrictions on the model structures, in many cases, leads to analytically intractable choice probability expressions, which necessitate the use of numerical integration techniques to evaluate the multidimensional integrals in the probability expressions. The numerical evaluation of such integrals has been the focus of extensive research dating back to the late 1800s. One of the first approaches was the extension of the one-dimensional numerical quadrature rules (such as the trapezoid rule and the Simpson's rule) to multidimensional polynomial-based cubature methods. However, the theory of polynomial-based cubature methods is complex in multiple dimensions, and so these methods are generally not considered for multidimensional integration. The only exception is when the multidimensional integral can be transformed into the product of 's' single integrals for which well-known quadrature formulas exist, so that an appropriate product formula may be constructed (see Press et al., 1992). However, such product formulas are unable to compute integrals with sufficient precision and speed in more than 2 dimensions (see Hajivassiliou and Ruud, 1994). This problem was alleviated with the development of a new method proposed in the 1940s; the Monte Carlo simulation approach; the basic concept for which seems to have existed as early as 1899.

The Monte Carlo (MC) approach to evaluating multidimensional integrals involves computing the integrand at a sequence of N *random* points and computing the average of the integrand values[3]. The MC simulation approach has an expected integration error of the order of $N^{-0.5}$, which is independent of the number of dimensions 's' and thus provides a great improvement over the quadrature-based methods. However, an integration error of the order of $N^{-0.5}$ implies that to obtain one additional decimal digit of accuracy, it is necessary to increase the number of draws from the random (MC or PMC) sequence by a hundredfold. This realization led to the development of several variance reduction techniques for the MC methods, which potentially lead to more accurate integral evaluation with fewer draws. One such technique is stratified random sampling, such as Latin Hypercube Sampling (or LHS, see McKay et al., 1979). Despite the improvements achieved by the variance reduction techniques, the convergence rate of MC methods is generally slow for simulated likelihood estimation of choice models.

[3] In actual implementation, however, the generation of truly random sequences using random physical effects such as radioactive emissions was found to be slow and inconvenient. This led to the use of deterministic pseudorandom sequences called pseudo-Monte Carlo (PMC) sequences, which appear random when subjected to simple statistical tests. The resulting simulation procedure is known as pseudo-Monte Carlo (PMC) simulation.

Extensive number theory research in the last few decades has led to the development of a more efficient simulation method, the quasi-Monte Carlo (QMC) method. This method uses the basic principle of the MC method in that it evaluates a multidimensional integral by replacing it with an average of the values of the integrand computed at discrete points. However, rather than using random sequences, QMC methods use cleverly-crafted, low discrepancy, deterministic quasi-Monte Carlo (or QMC) sequences. These QMC sequences are designed to achieve a more even distribution of points in the integration space than the MC and PMC sequences.

Over the years, several different quasi-random sequences have been proposed for QMC simulation. Among these are the reverse radix-based sequences (such as the Halton sequence) and the *(t,s)*-sequences (such as the Sobol and Faure sequences). The even distribution of points provided by these low discrepancy sequences leads to efficient convergence for the QMC method, generally at rates that are higher than the MC method. In particular, the theoretical upper bound for the integration error in the QMC method is of the order of N^{-1} for one-dimensional integration[4], where N is the number of draws of the quasi-random sequence used for the evaluation of the integral. Despite these obvious advantages, the QMC method has two major limitations. First, the deterministic nature of the quasi-random sequences makes it difficult to estimate the error in the QMC simulation procedure (while there are theoretical results to estimate integration error via upper bounds with the QMC sequence, these are much too difficult to compute and are very conservative upper bounds). Second, a common problem with many low-

[4] In general the upper bound for the integration error in the QMC method is of the order of $(\log N)^s / N$

discrepancy sequences is that they exhibit poor properties in higher dimensions. The Halton sequence, for example, suffers from significant correlations between the radical inverse functions for different dimensions, particularly in the larger dimensions. A growing field of research in QMC methods has resulted in the development, and continuous evolution, of efficient *randomization* strategies (to estimate the error in integral evaluation) and *scrambling* techniques (to break correlations in higher dimensions). The randomization procedure involves introducing some randomness into the quasi-random sequence, while preserving the equidistribution property of the underlying sequence (Owen, 1997, 1998; Tuffin, 1996). The resulting sequences, called hybrid or randomized QMC (RQMC) sequences, provide better accuracy than PMC sequences while also providing the ability to estimate the integration error. Scrambling techniques, on the other hand, were devised by number theorists to scramble the numbers of different dimensions to break the correlations in the higher dimensions of QMC sequences (see, for example, Braaten and Weller, 1979)[5].

Research on the generation and application of randomized and scrambled QMC sequences clearly indicates the superior accuracy of QMC methods over PMC methods in the evaluation of multidimensional integrals (see Sarkar and Prasad, 1986; Morokoff and Caflisch, 1994, 1995; Kocis and Whiten, 1997; Wang and Hickernell, 2000). In particular, the advantages of using QMC simulation for such applications in econometrics as simulated maximum likelihood inference, where parameter estimation entails the

[5] The reader will note, however, that the terms ‘randomization’ and ‘scrambling’ are not mutually exclusive. For instance, Owen’s scrambling technique breaks correlations while introducing randomness at the same time. Braaten-Weller scrambling, on the other hand, is a method that does not use any randomness.

approximation of several multidimensional integrals at each iteration of the optimization procedure, should be obvious. However, the first introduction of the QMC method for the simulated maximum likelihood inference of econometric choice models occurred only in 1999 when Bhat tested Halton sequences for mixed logit estimation and found their use to be vastly superior to random draws. Since Bhat's initial effort, there have been several successful applications of QMC methods for the simulation estimation of flexible discrete choice models, though most of these applications have been based on the Halton sequence (see, for example, Train, 1999; Revelt and Train, 2000; Bhat, 2001; Park et al., 2003; Bhat and Gossen, 2004; Bhat and Srinivasan, 2004). Number theory, however, abounds in many other kinds of low-discrepancy sequences that have been proven to have better theoretical and empirical convergence properties than the Halton sequence in the estimation of a single multidimensional integral. For instance, Bratley and Fox (1988) conduct a numerical comparison of the performance of Halton, Faure and Sobol sequences in the estimation of a single multidimensional integral. Their results clearly indicate that the Faure and Sobol sequences are superior to the Halton sequence in terms of accuracy and efficiency. There have also been several numerical studies on the simulation estimation of a single multidimensional integral that present significant improvements in the performance of QMC sequences through the use of scrambling techniques (see Kocis and Whiten, 1997 and Wang and Hickernell, 2000). It is, therefore, of interest to examine the performances of the different QMC sequences and their scrambled versions in the simulation estimation of flexible discrete choice models.

Section 5.2 discusses the specific objectives of this study. Section 5.3 presents the background for the generation of alternative sequences. The evaluation framework used in this study and the computational results form the contents of chapter 6.

5.2 Objectives

As described section 1.2.2, the second broad objective of this dissertation research is to compare the performance of different kinds of low discrepancy sequences, and their scrambled and randomized versions, in the simulated maximum likelihood estimation of the mixed logit class of discrete choice models. Specifically, the extensively used Halton sequence and a special case of *(t.m.s)*-nets known as the Faure sequence are selected. The choice of the Faure sequence is motivated by two reasons. First, the generation of the Faure sequence is a fairly straightforward and non-time consuming procedure. Second, it has been proved that the Faure sequence performs better than the Halton sequence in the evaluation of a single multidimensional integral (Kocis and Whiten, 1997).

The performance of the Halton and Faure sequences is compared against the performance of a stratified random sampling PMC sequence (the Latin Hypercube Sampling or LHS sequence) by constructing numerical experiments within a simulated maximum likelihood inference framework. Further, the numerical experiments also include a comparison of scrambled versions of the QMC sequences against their standard versions to examine potential improvements in performance through scrambling. The performances of the various non-scrambled and scrambled sequences are evaluated based on their ability to efficiently and accurately recover the true model parameters.

The total number of draws required for the estimation of a mixed multinomial logit (MMNL) model on a dataset of Q observations is N×Q, where N is the number of draws used to simulate the probabilities for each observation. The N×Q draws of a QMC sequence can be generated either as one long sequence of N×Q draws, or as a set of N draws which is scrambled Q times to obtain Q different sets of N draws. The first approach is referred to in this dissertation as the generation of draws *without scrambling across observations*, and the second approach as the generation of draws *with scrambling across observations*. In the numerical experiments, both these approaches are compared in the generation of each of the standard and scrambled QMC sequences. Another important point to note is that the standard and scrambled versions of the QMC and the LHS sequences are all generated as uniformly-distributed sequences of points. The estimation of an MMNL model, however, requires the simulation of a normal mixing distribution and therefore needs as input a low discrepancy normally-distributed sequence of points. Two different transformation procedures to convert the uniformly-distributed sequences to normally-distributed sequences – the Box-Muller and the Inverse Normal transform procedures are tested and compared.

To summarize, the specific objectives of this research are three-fold. The first objective is to experimentally compare the overall performance of the Halton and Faure sequences (and their scrambled versions) against each other and against the LHS sequence[6]. The second objective is to compare the efficiency of the QMC sequences with

[6] Sandor and Train (2004) perform a comparison of four different kinds of *(t,m,s)*-nets (created from Niederreiter nets), the standard Halton, and random start Halton sequences against simple random draws. Their study considers the estimation of a 5-dimensional mixed logit model using 64 QMC draws per

and without scrambling across observations. The third objective is to compare the Box-Muller and the Inverse Normal transform procedures for translating uniformly distributed sequences to normally distributed sequences.

5.3 Background for generation of alternative sequences

This section describes the various procedures to generate PMC and QMC sequences. Specifically, the following sections discuss the generation of PMC sequences using the Latin Hypercube Sampling (LHS) procedure (Section 5.3.1), and the generation of the QMC sequences proposed by Halton and Faure (Section 5.3.2); the scrambling techniques (Section 5.3.3) and randomization techniques (Section 5.3.4) applied in this study; the generation of sequences *with* and *without* scrambling across observations (Section 5.3.5); and basic descriptions of the Box-Muller and Inverse Normal transforms (Section 5.3.6).

5.3.1 PMC Sequences

The basic idea of the PMC simulation technique is to evaluate a multidimensional integral by computing the average value of the integrand over a sequence of N pseudorandom points (also referred to as a PMC sequence). PMC sequences can be easily generated using standard random number generators available in most software packages. A typical PMC simulation uses a simple random sampling (SRS) procedure to generate a uniformly distributed PMC sequence over the integration space. An alternate approach

---

observation, and compares the bias, standard deviation and RMSE associated with the estimated parameters. In this study we have conducted numerical experiments both in 5 and 10 dimensions in order that the comparisons may capture the effects of dimensionality. For the 5-dimensional mixed logit estimation problem we also examined the impact of varying number of draws (25, 125 and 625). Finally, we examine the performance of the Faure sequence and LHS method, along with the Halton sequence, and consider different scrambling variants of these sequences.

known as Latin Hypercube sampling (LHS), that yields asymptotically lower variance than SRS, is described in the following section.

5.3.1.1 Latin Hypercube Sampling

The LHS method was first proposed as a variance reduction technique (McKay et al., 1979) within the context of PMC simulation-based simulation. The basis of LHS is a full stratification of the integration space, with a random selection inside each stratum. This method of stratified random sampling in multiple dimensions can be easily applied to generate a well-distributed sequence. The LHS technique involves drawing a sample of size N from multiple dimensions such that for each individual dimension the sample is maximally stratified. A sample is said to be maximally stratified when the number of strata equals the sample size N, and when the probability of falling in each of the strata equals N^{-1}.

To draw a uniform LHS sequence of size N in K dimensions, the i^{th} sample element for dimension j is given by

$$u_{ij} = ((p_{ij} - \xi_{ij}) / N), \qquad \textbf{Eq. 36}$$

where, for each $j = 1,\ldots,K$, p_{ij} ($i = 1,\ldots,N$) is a random permutation of the numbers $1,\ldots,N$; ξ_{ij} is a uniform distributed random number between 0 and 1; and the K permutations and the NK uniform variates ξ_{ij} are mutually independent. The LHS sequence is then given by

$$\psi_{lhs}^{(N)} = ((p - \xi) / N), \qquad \textbf{Eq. 37}$$

where, $\psi_{lhs}^{(N)}$ is an NxK matrix consisting of N draws of a K-dimensional LHS sequence, p is an NxK matrix consisting of K different random permutations of the numbers 1,…,N, and ξ_{ij} is an NxK matrix of uniformly distributed random numbers between 0 and 1.

In essence, the LHS sequence is obtained by slightly shifting the elements of an SRS sequence, while preserving the ranks (and rank correlations) of these elements, to achieve maximal stratification. Figure 5 presents a maximally stratified and uniformly distributed 2-dimensional LHS sequence with N = 6. As can be seen in the figure, each stratum in either dimension contains exactly one point. This is achieved by dividing the integration space into the required number of strata and randomly selecting a point in each stratum. Figure 6 plots the first 100 points of a 2-dimensional LHS sequence.

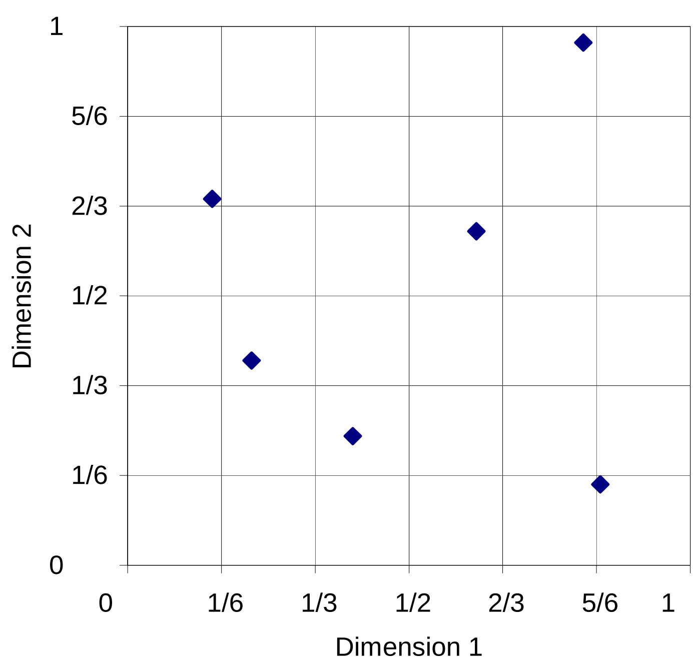

Figure 5. Uniformly-distributed LHS sequence in 2 dimensions (N = 6)

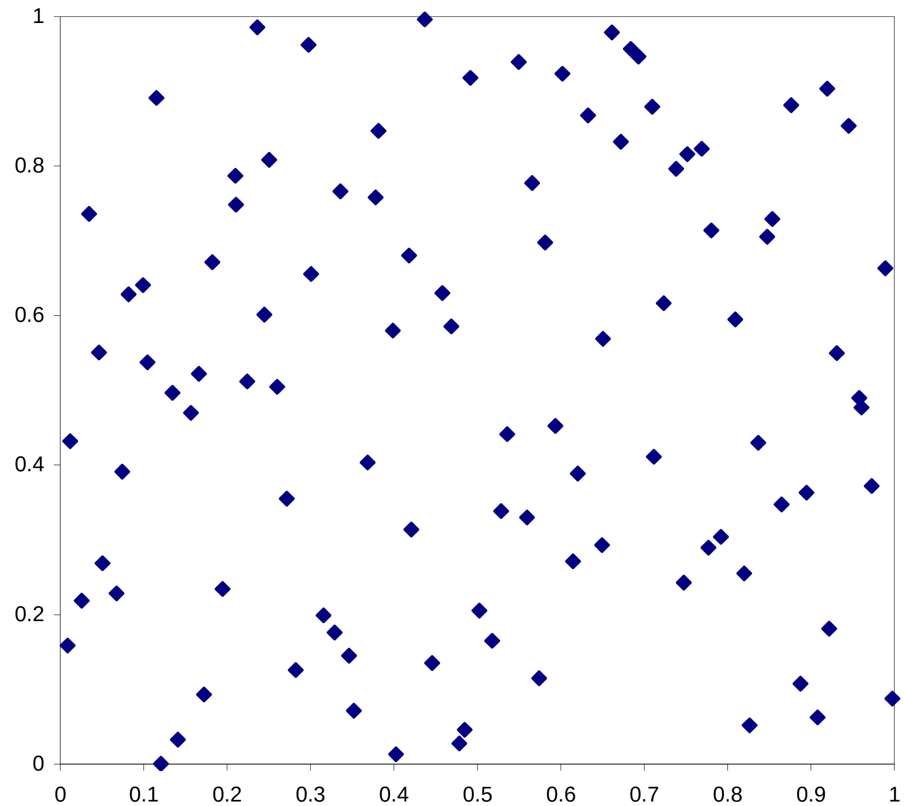

Figure 6. First 100 points of a 2-dimensional LHS sequence

5.3.2 QMC Sequences

The underlying idea of the QMC simulation technique is to evaluate a multidimensional integral by computing the average value of the integrand over a deterministic set of low-discrepancy points that are generated to be evenly distributed over the integration space. Many of the low-discrepancy sequences in use today are linked to the van der Corput sequence, which was originally introduced for dimension s = 1 and base b = 2 (van der Corput 1935a, 1935b). Sequences based on the van der Corput sequence are also referred to as the reverse radix-based sequences. To find the nth term,

x_n, of a van der Corput sequence, let us first write the unique digit expansion of n in base b as:

$$n = \sum_{j=0}^{\infty} a_j(n) b^j \text{ , where } 0 \le a_j(n) \le b-1 \text{ and } b^J \le n \le b^{J+1} . \qquad \textbf{Eq. 38}$$

This is a unique expansion of *n* that has only finitely many non-zero coefficients $a_j(n)$. The next step is to evaluate the radical inverse function in base *b*, which is defined as

$$\phi_b(n) = \sum_{j=0}^{\infty} a_j(n) b^{-j-1} . \qquad \textbf{Eq. 39}$$

The van der Corput sequence in base *b* is then given by $x_n = \phi_b(n)$, for all $n \ge 0$. This idea that the coefficients of the digit expansion of an increasing integer *n* in base *b* can be used to define a one-dimensional low-discrepancy sequence inspired Halton (1960) to create an *s*-dimensional low-discrepancy Halton sequence by using *s* van der Corput sequences with relatively prime bases for the different dimensions.

An alternative approach to the generation of low-discrepancy sequences is to start with points placed into certain equally sized volumes of the unit cube. These fixed length sequences are referred to as (t,m,s)-nets, and related sequences of indefinite lengths are called (t,s)-sequences. Sobol (1967) suggested a multidimensional (t,s)-sequence using base 2, which was further developed by Faure (1982) who suggested alternate multidimensional (t,s)-sequences with base $b \ge s$.

The following sub-sections describe the procedures used in this paper to generate the standard Halton and Faure sequences.

5.3.2.1 Halton Sequences

The standard Halton sequence in s dimensions is obtained by pairing s one-dimensional van der Corput sequences based on s pairwise relatively prime integers, $b_1, b_2, ..., b_S$ (usually the first s primes) as discussed earlier. The Halton sequence is based on prime numbers, since the sequence based on a non-prime number will partition the unit space in the same way as each of the primes that contribute to the non-prime number. Thus, the nth multidimensional point of the sequence is as follows:

$$\phi(n) = (\phi_{b_1}(n), \phi_{b_2}(n), ..., \phi_{b_s}(n)) . \qquad \textbf{Eq. 40}$$

The standard Halton sequence of length N is finally obtained as

$$\psi_h^{(N)} = [\phi(1)', \phi(2)', ..., \phi(n)']' . \qquad \textbf{Eq. 41}$$

The Halton sequence is generated number-theoretically as described above rather than randomly and so successive points of the sequence "know" how to fill in the gaps left by earlier points, leading to a uniform distribution within the domain of integration. This is illustrated in Figure 7 where the first 100 points of a 2-dimensional Halton sequence are plotted. The points 51 through 100 (denoted by ∗) clearly fill in the gaps left by the previous 50 points (denoted by ◆). The resulting set of 100 points in Figure 7 are more evenly distributed than the randomly generated LHS sequence in Figure 6, which is observed to exhibit some clumping of points.

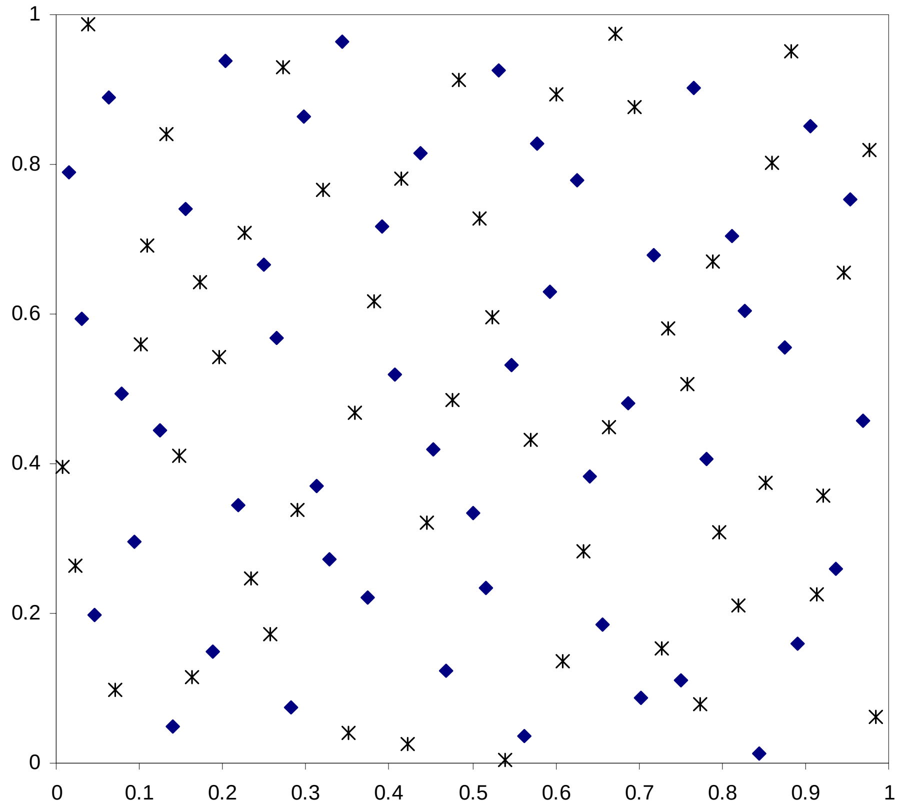

Figure 7. First 100 points of a 2-dimensional Halton sequence

5.3.2.2 Faure Sequences

The standard Faure sequence is a *(t,s)*-sequence designed to span the domain of the *s*-dimensional cube uniformly and efficiently. In one dimension, the generation of the Faure sequence is identical to that of the Halton sequence. In *s* dimensions, while the Halton sequence simply pairs *s* one-dimensional sequences generated by the first *s* primes, the higher dimensions of the Faure sequence are generated recursively from the elements of the lower dimensions. So if *b* is the smallest prime number such that $b \geq s$ and $b \geq 2$, then the first dimension of the *s*-dimensional Faure sequence corresponding to *n* can be obtained by taking the radical inverse of *n* to the base *b*:

$$\phi_b^1(n) = \sum_{j=0}^{J} a_j^1(n) b^{-j-1} \qquad \textbf{Eq. 42}$$

The remaining dimensions are found recursively. Assuming we know the coefficients $a_j(n)$ corresponding to the first (*k-1*) dimensions, the coefficients for the kth dimension are generated as follows:

$$a_j^k(n) = \sum_{i \geq j}^{J} {}^iC_j a_i^{k-1}(n) \bmod b, \qquad \textbf{Eq. 43}$$

where ${}^iC_j = i!/j!(i-j)!$ is the combinatorial function. Thus the next level of coefficients required for the k^{th} element in the *s*-dimensional sequence is obtained by multiplying the coefficients of the (*k-1*)th element by an upper triangular matrix *C* with the following elements.

$$C = \begin{bmatrix} {}^0C_0 & {}^1C_0 & {}^2C_0 & {}^3C_0 & \cdots \\ 0 & {}^1C_1 & {}^2C_1 & {}^3C_1 & \\ 0 & 0 & {}^2C_2 & {}^3C_2 & \\ 0 & 0 & 0 & {}^3C_3 & \\ \vdots & & & & \end{bmatrix}$$

These new coefficients $a_j^k(n)$ are then reflected about the decimal point to obtain the k^{th} element as follows:

$$\phi_b^k(n) = \sum_{j=0}^{J} a_j^k(n) b^{-j-1}, \ \ 2 \leq k \leq s \qquad \textbf{Eq. 44}$$

This recursive procedure generates the *s* points corresponding to the integer *n* in the Faure sequence based on $b \ (\geq s)$. Thus the n^{th} multidimensional point in the sequence is

$$\phi(n) = (\phi_b^1(n), \phi_b^2(n), ..., \phi_b^S(n))$$

The standard Faure sequence of length N is then obtained in the same manner as the standard Halton sequence:

$$\psi_f^{(N)} = [\phi(1)', \phi(2)', ..., \phi(n)']' \quad \textbf{Eq. 45}$$

Faure sequences are essentially *(t,m,s)*-nets in any prime *b* with $b \geq s \; and \; t = 0$. A Faure sequence of b^m points is generated to be evenly distributed over the integration space, such that if we plot the sequence in the integration space together with the elementary intervals of area b^{-m}, exactly one point will fall in each elementary interval. Take, for example, a Faure sequence of 8 (2^3) points in 2 dimensions that is a (0,3,2)-net in base 2, as plotted in Figure 8. Figure 9 consists of four different plots of this Faure sequence within the integration space. Each plot presents a different construction of elementary intervals of area 2^{-3} within the same integration space. As seen from these plots, exactly one point of the Faure sequence falls within every elementary interval of area 2^{-3} (or 1/8) regardless of how these intervals are constructed, thus achieving an even distribution of points over the domain of integration.

Earlier studies have shown that for higher dimensions, the properties of the Faure sequence are poor for small values of *n* in equation 45 (refer, for example, Fox, 1986). To overcome this issue the first 100000 multidimensional points are dropped for all the standard and scrambled Faure sequences generated.

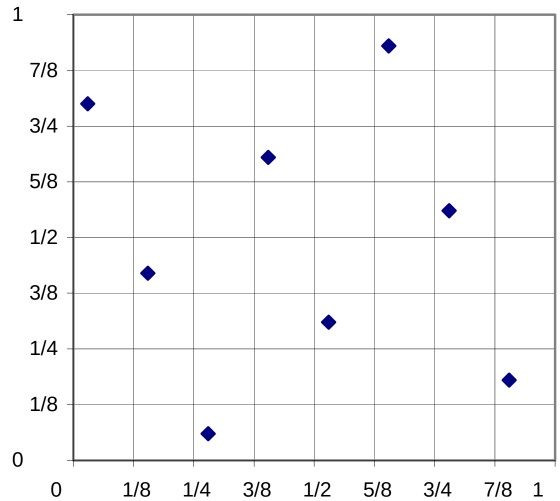

Figure 8. (0,3,2)-net in base 2

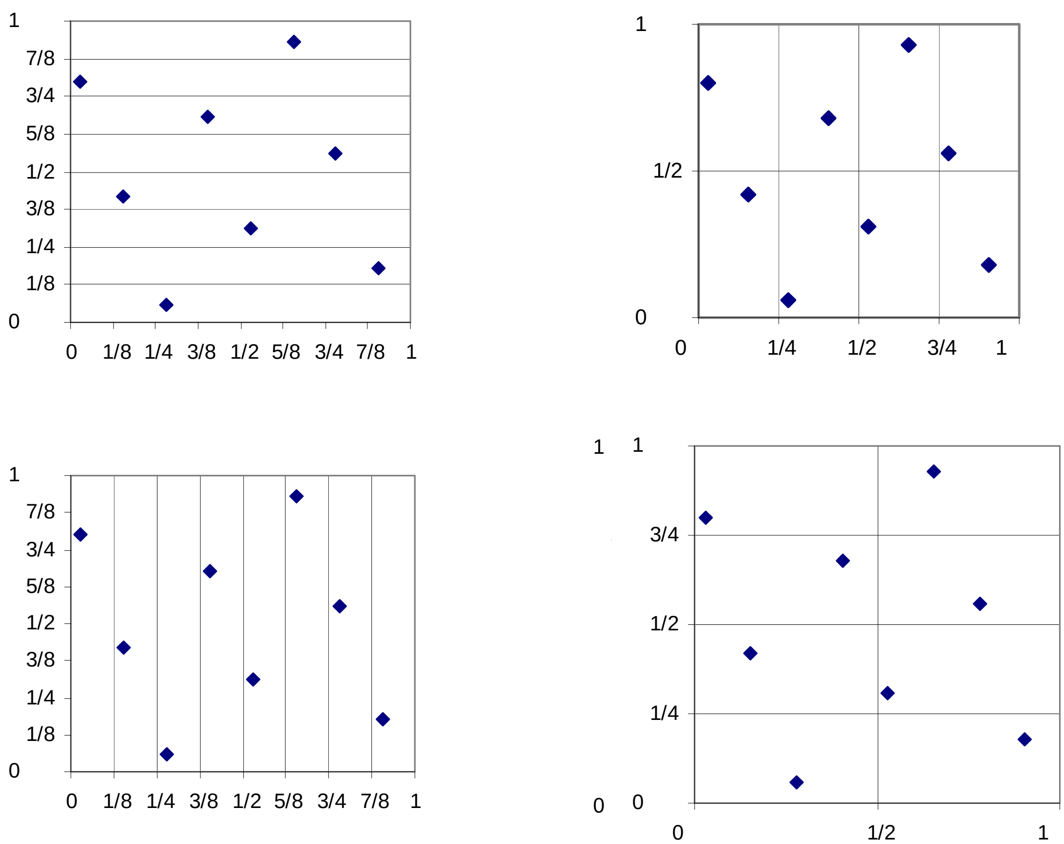

Figure 9. (0,3,2)-net in base 2 with elementary intervals of area 1/8 (Modified from Ökten and Eastman, 1988)

5.3.3 Scrambling Techniques used with QMC Sequences

Research has shown that the finite parts (for moderate sizes) in higher dimensions of many QMC sequences have poor properties, which can be alleviated using suitable scrambling techniques. The standard Halton sequence, for instance, suffers from significant correlations between the radical inverse functions at higher dimensions. For example, the fourteenth dimension (corresponding to the prime number 43) and the fifteenth dimension (corresponding to the prime number 47) consist of 43 and 47 increasing numbers, respectively. This generates a correlation between the fourteenth and fifteenth coordinates of the Halton sequence as illustrated in Figure 10. The standard Faure sequence, on the other hand, forms distinct patterns in higher dimensions that also cover the unit integration space in diagonal strips, thus showing significantly higher discrepancies in the higher dimensions. Figure 11 illustrates this in a plot of the fifteenth and sixteenth coordinates of the Faure sequence.

Several methods have been suggested to improve the uniformity of the QMC sequences in higher dimensions. Since most of these methods involve some form of permutation (or scrambling) of the coefficients in each of the radical inverse functions in an effort to redistribute the points of the sequence more uniformly, they are referred to as *scrambling* techniques. This study implements the Braaten-Weller scrambling for Halton sequences, and the Random Digit and Random Linear scrambling for Faure sequences. Each of these methods is described in greater detail in the following sections.

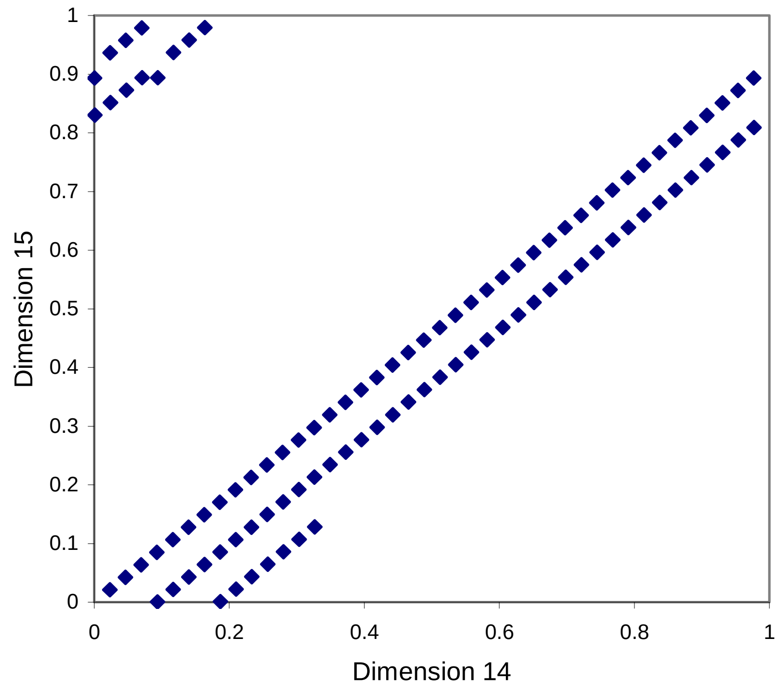

Figure 10. Standard Halton sequence: first 100 points (Source: Bhat, 2003)

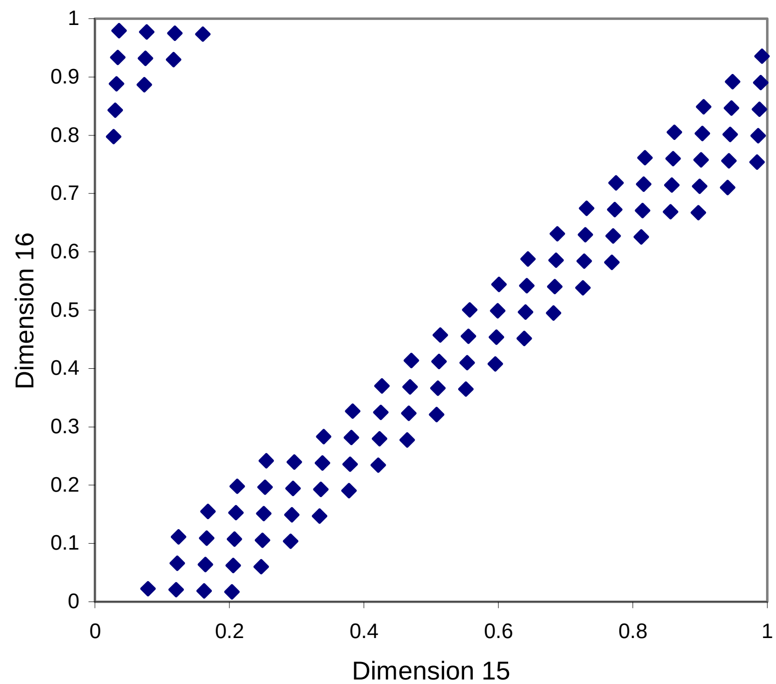

Figure 11. Standard Faure sequence; first 100 points

5.3.3.1 Braaten-Weller scrambling

Braaten and Weller (1979) describe a permutation of the coefficients $a_j(n)$ in equation 39 that minimizes the discrepancy of the resulting scrambled Halton sequence. Their method suggests different permutations for different prime numbers, thus effectively breaking the correlation across dimensions. Braaten and Weller have also proved that their scrambled sequence retains the theoretically appealing N^{-1} order of integration error of the standard Halton sequence.

Figure 12 presents the Braaten-Weller scrambled Halton sequence in the fourteenth and fifteenth dimensions. The same sequence before scrambling is presented in Figure 10. The effectiveness of the Braaten-Weller scrambling method in breaking the correlations is evident from these plots.

To illustrate the Braaten-Weller scrambling procedure, take the 5th number in base 3 of the Halton sequence, which in the digitized form is 0.21. The suggested permutation for the coefficients (0, 2, 1) for the prime 3 is (0, 1, 2), which when expanded in base 3 translates to $1x3^{-1} + 2x3^{-2} = 5/9$. The first 8 numbers in the standard Halton sequence corresponding to base 3 are 1/3, 2/3, 1/9, 4/9, 7/9, 2/9, 5/9, 8/9. The Braaten-Weller scrambling procedure yields the following scrambled sequence: 2/3, 1/3, 2/9, 8/9, 5/9, 1/9, 7/9, 4/9.

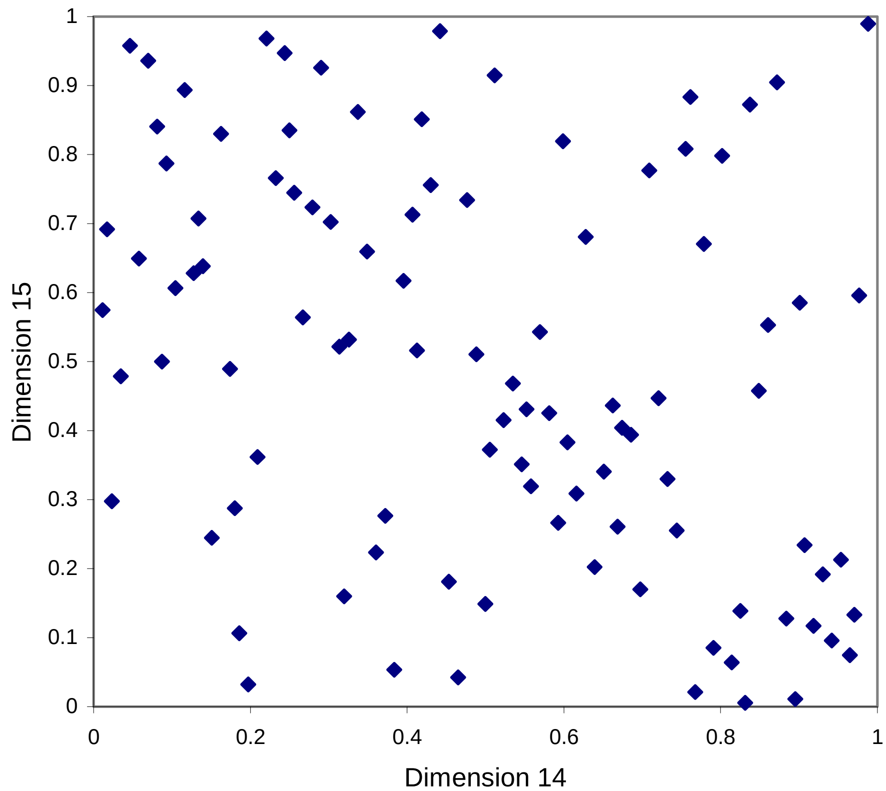

Figure 12. Braaten-Weller Scrambled Halton Sequence: first 100 points

5.3.3.2 Random Digit Scrambling

The Random Digit scrambling approach for Faure sequences is conceptually similar to the Braaten-Weller method, and suggests random permutations of the coefficients $a_j^k(n)$ to scramble the standard Faure sequence. Matoušek (1998) describes this scrambling technique and its theoretical properties in detail.

Figure 13 presents the Random Digit scrambled Faure sequence in the fifteenth and sixteenth dimensions. The same sequence before scrambling is presented in Figure 11. A comparison of the plots in these two figures indicates that the Random Digit scrambling technique is very effective in breaking the patterns in higher dimensions and generating a more even distribution of points.

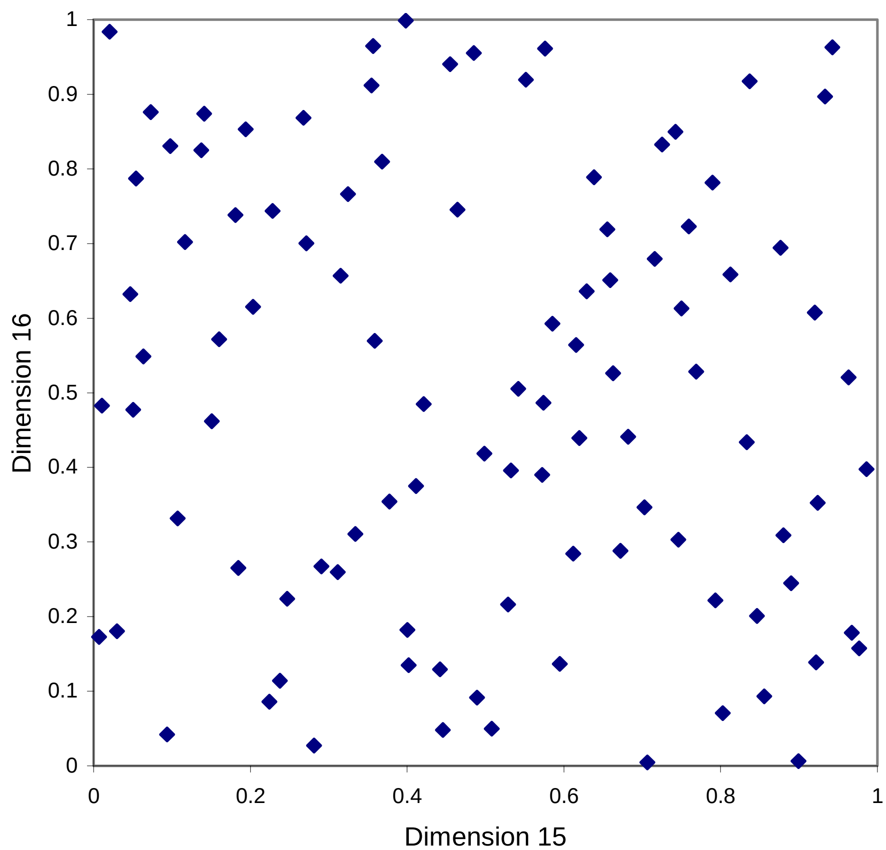

Figure 13. Random Digit Scrambled Faure Sequence: first 100 points

The Random Digit scrambling technique essentially uses independent random permutations for each coefficient in each dimension of the sequence. For example, consider the first two elements in a 5-dimensional Faure sequence consisting of the following coefficients,

{ {(2, 1, 0), (2, 3, 1), (2, 4, 2), (4, 2, 3), (1, 0, 4)},

{(1, 0, 0), (3, 2, 1), (0, 2, 4), (0, 4, 4), (4, 4, 0)} }.

In each of the 5 dimensions, the vector's base 5 expansion has 3 digits, which implies that we need 15 independent random permutations $\pi = (\pi_1, \ldots\ldots, \pi_{15})$. π_1, for example, could be the following permutation

$\pi_1(0) = 4$; $\pi_1(1) = 2$; $\pi_1(2) = 0$; $\pi_1(3) = 1$; $\pi_1(4) = 3$.

So when all 15 permutations are applied to the above two elements, we obtain the scrambled Faure sequence as follows

$$\{\{(\pi_1(2), \pi_2(1), \pi_3(0)), (\pi_4(2), \pi_5(3), \pi_6(1)), (\pi_7(2), \pi_8(4), \pi_9(2)),$$
$$(\pi_{10}(4), \pi_{11}(2), \pi_{12}(3)), (\pi_{13}(1), \pi_{14}(0), \pi_{15}(4))\},$$
$$\{(\pi_1(1), \pi_2(0), \pi_3(0)), (\pi_4(3), \pi_5(2), \pi_6(1)), (\pi_7(0), \pi_8(2), \pi_9(4)),$$
$$(\pi_{10}(0), \pi_{11}(4), \pi_{12}(4)), (\pi_{13}(4), \pi_{14}(4), \pi_{15}(0))\}\}$$

For each application, a different digit-scrambled version of the Faure sequence should be used. This is achieved by generating new random permutations π for each run.

5.3.3.3 Random Linear Scrambling

The Random Linear Scrambling technique for Faure sequences proposed by Matoušek (1998) is a variant of a procedure used by Tezuka (1995) in generating what he called "generalized Faure sequences". This scrambling approach is based on the concept of cleverly introducing randomness in the recursive procedure of generating the coefficients for each successive dimension.

Figure 14 presents the Random Linear scrambled Faure sequence in the fifteenth and sixteenth dimensions. The same sequence before scrambling is presented in Figure 11. These plots indicate that the Random Linear scrambling method results in a much more even distribution of points in the fifteenth and sixteenth coordinates than the Random Digit scrambling method (Figure 13)[7].

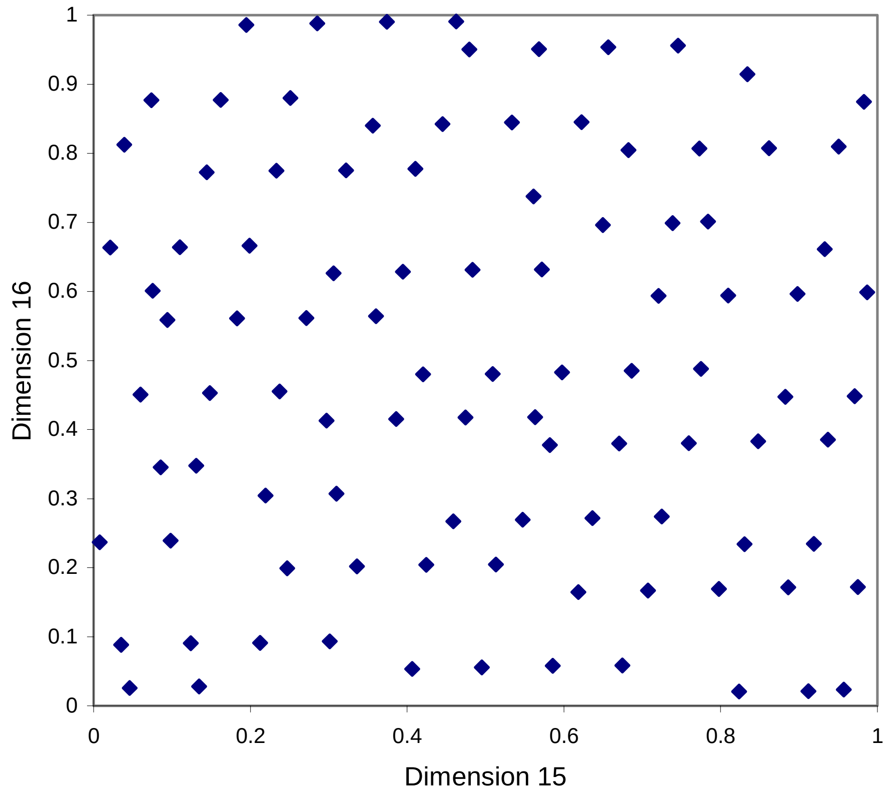

Figure 14. Random Linear Scrambled Faure Sequence: first 100 points

The Random Linear scrambling approach of Matoušek is easily implemented by modifying the upper triangular combinatorial matrix C used in generating Faure sequences (see Section 5.3.2.2). A linear combination $AC+B$ is used in the place of the

[7] The behavior of the Random Linear scrambling technique seemed to not always be predictable in terms of uniformity of coverage. In particular, the results of the Random Linear scrambling method for the nineteenth and twentieth dimensions of the Faure sequence were observed to be rather poor as the redistribution of points occurs in a fixed pattern. In the seventeenth and eighteenth dimensions, although the redistribution of points occurs in a fixed pattern, the pattern results in a more even coverage than in the case of the nineteenth and twentieth dimensions.

matrix C, where *A* is a randomly generated matrix and B is a random vector, both consisting of uniform random variates *U[0, b-1]*.

5.3.4 Randomization of QMC Sequences

QMC sequences, such as the standard Halton sequence described in Section 5.3.2, are fundamentally deterministic and do not permit the practical estimation of integration error. Since a comparison of the performance of these sequences necessitates the computation of simulation variances and errors, it is necessary to randomize these QMC sequences. Randomization of QMC sequences is a technique that introduces randomness into a deterministic QMC sequence while preserving the equidistribution property of the sequence (see Shaw, 1988; Tuffin, 1996). This study uses Tuffin's randomization procedure, which is based on the following concept. Let $\psi^{(N)}$ be a QMC sequence of length N over the *s*-dimensional unit cube and consider any *s*-dimensional uniformly distributed vector *u* in the *s*-dimensional unit cube. Construct a new sequence $\chi^{(N)}$ whose elements χ_{ns} are obtained as $\psi_{ns} + u_s$ if $\psi_{ns} + u_s \leq 1$, and $\psi_{ns} + u_s - 1$ if $\psi_{ns} + u_s > 1$. Tuffin proves that the sequence $\chi^{(N)}$ so formed is also a QMC sequence of length N over the *s*-dimensional unit cube. Intuitively, the vector *u* simply shifts the points of each coordinate of the original QMC sequence $\psi^{(N)}$ by a certain value. Since all the points within each coordinate are shifted by the same amount, the new sequence will preserve the equidistribution property of the original sequence (see Bhat, 2003, for a detailed explanation of the randomization procedure).

In the numerical experiments in this dissertation research, Tuffin's randomization is used to perform 20 estimation runs for each test scenario. The results of these 20 estimation runs are used to compute the relevant statistical measures.

5.3.5 Generation of Draws With and Without Scrambling Across Observations

The previous sections describe the process of generation of the LHS sequence, and the standard and scrambled versions of the Halton and Faure sequences. This section examines the generation of these sequences specifically in the context of the estimation of an MMNL model.

The simulated maximum likelihood estimation of an MMNL with a K-dimensional mixing distribution involves generating a K-dimensional PMC or QMC sequence for a specified number of draws 'N' for each individual in the dataset. Therefore estimating an MMNL model on a dataset with Q individuals will require an N×Q K-dimensional PMC or QMC sequence, where each set of N K-dimensional points computes the contribution of an individual to the log-likelihood function. A PMC or QMC sequence of length N×Q can be generated either as one continuous sequence of length N×Q or as Q independent sets of length N each. In the case of PMC sequences, both these approaches amount to the same since a PMC sequence is identical to a random sequence with each point of the sequence being independent of all the previous points. In the case of QMC sequences, Q independent sets of length N can be generated by first constructing a sequence of length N and then scrambling it Q times, which is known as generation *with scrambling across observations*. The other alternative of generating a continuous QMC sequence of length N×Q is known as generation *without scrambling across observations*.

QMC sequences generated with and without scrambling across observations exhibit different properties. A QMC sequence generated with scrambling across observations leads to a higher degree of randomness in simulating the contribution of the individuals to the log-likelihood function. This method of generation can also be expected to result in time savings especially when dealing with high-dimensions and a large number of individuals. A continuously generated QMC sequence of length N×Q, on the other hand, leads to an averaging out of simulation errors across individuals. This occurs due to the basic property of QMC sequences in that each set of N points fills in the gaps left by the set of N points used for the previous individuals thereby causing the simulated probabilities to be negatively correlated across observations (Train, 1999; Bhat, 2003).

In this study the performance of the various scrambled and standard QMC sequences generated both with and without scrambling across observations are examined.

5.3.6 Box-Muller and Inverse Normal Transforms

The standard and scrambled versions of the Halton and Faure sequences (with or without scrambling across observations), and the LHS sequence are generated to be uniformly distributed over the multidimensional unit cube. Simulation applications, however, may require these sequences to take on other distributional forms. For example, the estimation of the MMNL model described in Section 6.1.1 requires normally distributed multivariate sequences that span the multidimensional domain of integration.

The transformation of the uniformly distributed LHS and QMC sequences to normally distributed sequences can be achieved using either the inverse standard normal distribution or one of the many approximation procedures discussed in the literature, such as the Box-Muller Transform (1958), Moro's method and Ramberg and Schmeiser approximation (1972). The performances of the inverse normal and the Box-Muller transforms are compared in this study.

If Y is a K-dimensional matrix of length N*Q containing the uniformly distributed LHS or QMC sequence, the inverse normal transformation yields $X = \Phi^{-1}(Y)$, where X is a normally distributed sequence of points in K-dimensions. The Box-Muller method approximates this transformation as follows. The uniformly distributed sequence of points in Y are transformed to the normally distributed sequence X using the equations

$$X_{ij} = \cos(2\pi Y_{i(j+1)})\sqrt{-2\log Y_{ij}} \text{ and } X_{i(j+1)} = \sin(2\pi Y_{i(j+1)})\sqrt{-2\log Y_{ij}}, \qquad (11)$$

for all i = 1, 2, … N*Q, and j = 1, 3, 5, … K-1, assuming that K is even. If K is odd, then simply generate an extra column of the sequence and perform the Box-Muller transform with the K+1 even columns. The (K+1)th column of the transformed matrix X can then be dropped.

CHAPTER 6. COMPARISON OF ALTERNATE QMC SEQUENCES

6.1 Evaluation framework

The performance of the sequences presented in the previous chapter is evaluated within the context of the simulated maximum likelihood estimation of the MMNL model. This section first discusses the simulated maximum likelihood estimation of the MMNL model (Section 6.1.1), then presents the experimental design used in generating a simulated dataset of 2000 observations (Section 6.1.2), and finally describes the scenarios tested in this study and the evaluation criteria used in comparing the performance of these sequences (Section 6.1.3). The computational results are presented in Section 6.2. All the numerical experiments in this study are implemented using the GAUSS matrix programming language.

6.1.1 Simulated Maximum Likelihood Estimation of the MMNL Model

In the numerical experiments in this research, a random-coefficients interpretation of the MMNL model structure is used. However, the results from these experiments can be generalized to any model structure with a mixed logit form. The random-coefficients structure essentially allows heterogeneity in the sensitivity of individuals to exogenous attributes. The utility that an individual q associates with alternative i is written as:

$$U_{qi} = \beta_q^{'} x_{qi} + \varepsilon_{qi} \qquad \textbf{Eq. 46}$$

where, x_{qi} is a vector of exogenous attributes, β_q is a vector of coefficients that varies across individuals with density $f(\beta)$, and ε_{qi} is assumed to be an independently and identically distributed (across alternatives) type I extreme value error term. With this specification, the unconditional choice probability of alternative i for individual q is given by the following mixed logit formula:

$$P_{qi}(\theta) = \int_{-\infty}^{\infty} L_{qi}(\beta) f(\beta \mid \theta)\, d(\beta), \quad L_{qi}(\beta) = \frac{e^{\beta' x_{qi}}}{\sum_j e^{\beta' x_{qj}}}, \qquad \textbf{Eq. 47}$$

where, P_{qi} is the probability that individual q chooses alternative i, β represents parameters which are random realizations from a density function *f(.)* also known as the mixing distribution, and θ is a vector of underlying moment parameters characterizing *f(.)*. While several density functions may be used for *f(.)*, the most commonly used is the normal distribution with θ representing the mean and variance.

The objective of simulated maximum likelihood inference is to estimate the parameters 'θ ' of the mixing distribution by numerical evaluation of the choice probabilities for all the individuals using simulation. Using '*N*' draws from the mixing distribution *f(.)*, each labeled β^n, $n = 1,\ldots,N$, the simulated probability for an individual can be calculated as

$$SP_{qi}(\theta) = (1/N) \sum_{n=1,\ldots,N} L_{qi}(\beta^n)\,. \qquad \textbf{Eq. 48}$$

$SP_{qi}(\theta)$ has been proved to be an unbiased estimate of $P_{qi}(\theta)$ whose variance decreases as the number of draws 'N' increases. The simulated log-likelihood function is then computed as

$$SLL(\theta) = \sum_{q=1,\ldots,Q} \ln(SP_{qi}(\theta))' \qquad \textbf{Eq. 49}$$

where i is the chosen alternative for individual q. The parameters 'θ' that maximize the simulated log-likelihood function are then calculated. Properties of this estimator have been studied, among others, by Lee (1992) and Hajivassiliou and Ruud (1994).

6.1.2 Experimental Design

The data for the numerical experiments conducted in this study were generated by simulation. Two sample data sets were generated containing 2000 observations (or individuals q in equation 46) and four alternatives per observation. The first data set was generated with 5 independent variables to test the performance of the sequences in 5 dimensions. The values for each of the 5 independent variables for the first two alternatives were drawn from a univariate normal distribution with mean 1 and standard deviation of 1, while the corresponding values for each independent variable for the third and fourth alternatives were drawn from a univariate normal distribution with mean 0.5 and standard deviation of 1. The coefficient to be applied to each independent variable for each observation was also drawn from a univariate normal distribution with mean 1 and standard deviation of 1 (i.e., $\beta_{qi} \sim N(1,1)$, q = 1, 2, … 2000 and i = 1, …, 4). The values of the error term, ε_{qi} in equation 46, were generated from a type I extreme value (or Gumbel) distribution, and the utility of each alternative was computed. The alternative with the highest utility for each observation was then identified as the chosen alternative. The second data set was generated similarly but with 10 independent variables to test the performance of the sequences in 10 dimensions.

6.1.3 Test Scenarios

This study uses the simulated datasets described above to numerically evaluate the performance of the LHS sequence, and the standard and scrambled versions of the Halton and Faure sequences within the MMNL framework. First random-coefficients mixed logit models are estimated in 5 and 10 dimensions, using a simulated estimation procedure with 20,000 random draws (N = 20,000 in equation 48). The resulting estimates are declared to be the "true" parameter values. The various sequences are then evaluated by computing their abilities to recover the "true" model parameters. This technique has been used in several simulation-related studies in the past (see Bhat, 2001; Hajivassiliou et al., 1996).

The test sequences include the standard Halton, Braaten-Weller scrambled Halton, standard Faure, Random Digit Scrambled Faure, Random Linear Scrambled Faure, and LHS sequences. For each of these six sequences, cases with 25, 125 and 625 draws (N in equation 48) are tested for 5 dimensions and cases with 100 draws for 10 dimensions.

6.2 Computational results

The estimation of the 'true' parameter values served as the benchmark to compare the performances of the different sequences. The performance evaluation of the various sequences was based on their ability to recover the true model parameters accurately. Specifically, the evaluation of the proximity of estimated and true values was based on two performance measures: (a) root mean square error (RMSE), and (b) mean absolute percentage error (MAPE). Further, for each performance measure two properties were computed: (a) bias, or the difference between the mean of the relevant values across the 20 runs and the true values, and (b) total error, or the difference between the estimated and true values across all runs[8].

One general note before the presentation and discussion of the results. The Box-Muller transform method to translate uniformly distributed sequences to normally distributed sequences resulted in higher bias and total error than the inverse normal transform method almost universally for all the scenarios tested (this is consistent with the finding of Tan and Boyle, 2000). Therefore only the results of the inverse transform procedure are presented here.

The computational results have been grouped into four tables. Table 4 presents the results corresponding to the two evaluation criteria (RMSE and MAPE) for the test scenarios with 25 draws for 5 dimensions; Table 5 presents the results for the test scenarios with 125 draws for 5 dimensions; Table 6 presents the results for the test scenarios with 625 draws for 5 dimensions; and Table 7 presents the results for the test scenarios with 100 draws for 10 dimensions. In each table, the first column specifies the type of sequence used, which will be one of the following (a) the standard Halton, (b) the Braaten-Weller scrambled Halton, (c) the standard Faure, (d) the Random Digit scrambled Faure, (e) the Random Linear scrambled Faure, and (f) the LHS sequence. The second column indicates whether the sequence is generated with or without scrambling

[8] The simulation variance was also computed, i.e.; the variance in relevant values across the 20 runs and the true values. However, the results of those computations are not discussed here in order to simplify presentation and also because the total error captures simulation variance.

across observations ("Scrambling" or "No Scrambling"). The remaining columns list the RMSE and MAPE performance measures for the estimators in each case.

In the following sections, the results are first examined and interpreted separately for each of the 25 draws, 125 draws, 625 draws and 100 draws (10 dimensions) cases; and then finally the overall trends in the results are examined.

6.2.1 5 Dimensions and 25 draws

Table 4 indicates that the standard and scrambled Halton sequences generated with scrambling across observations yield lower RMSE and MAPE bias and total error than the corresponding sequences generated without scrambling across observations. A similar result holds for the standard Faure sequence. However, for the scrambled Faure sequences, the sequences generated without scrambling across observations yield about equal or lower RMSE and MAPE total error than the sequences that are generated with scrambling across observations.

Table 4. Evaluation of ability to recover model parameters (5 dimensions, 25 draws)

Sequence Type	Scrambling across observations	RMSE		MAPE	
		Bias	Total error	Bias	Total error
Standard Halton	No Scrambling	0.2987	0.3275	30.6976	30.6976
	Scrambling	0.2817	0.2997	29.7409	29.7409
Braaten-Weller Scram. Halton	No Scrambling	0.3157	0.3515	32.5745	32.5745
	Scrambling	0.2948	0.3259	30.4528	30.4544
Standard Faure	No Scrambling	0.2586	0.2869	27.2551	27.2551
	Scrambling	0.2374	0.2887	24.0570	24.0937
Random Digit Scram. Faure	No Scrambling	0.2955	0.3332	28.8420	28.8420
	Scrambling	0.2947	0.3541	29.8144	29.8144
Random Linear Scram Faure	No Scrambling	0.2677	0.2978	27.9082	27.9082
	Scrambling	0.2848	0.3209	29.4035	29.4035
LHS	N/A	0.2650	0.3059	27.7668	27.7668

Overall, the following inferences can be made regarding the performance of the sequences in 5 dimensions and with 25 draws:

(a) The standard Halton sequence yields lower RMSE and MAPE bias and total errors than the Braaten-Weller scrambled Halton sequence.

(b) The standard Faure sequence also yields lower RMSE and MAPE bias and total errors than the corresponding scrambled versions.

(c) The standard Faure sequence performs better than the corresponding standard Halton sequence on all counts. The LHS sequence performs at about the same level as all other sequences except the standard Faure.

(d) The standard Faure sequence with scrambling across observations provides the best results in the overall.

6.2.2 5 Dimensions and 125 draws

Table 5 indicates that, for the standard Halton sequence, the case without scrambling across observations provides lower bias for both the RMSE and MAPE cases, but slightly higher total error. For the scrambled Halton, the case without scrambling across observations dominates (this latter result is the reverse of what was found in the 25 draws case). For the Faure sequences, no scrambling across observations provides better results than scrambling across observations for the standard and Random Digit scrambled Faure versions. However, the reverse is the case for the Random Linear Faure sequence.

Table 5. Evaluation of ability to recover model parameters (5 dimensions, 125 draws)

Sequence Type	Scrambling across observations	RMSE		MAPE	
		Bias	Total error	Bias	Total error
Standard Halton	No Scrambling	0.0538	0.0672	5.6565	6.0881
	Scrambling	0.0560	0.0627	5.9892	6.0709
Braaten-Weller Scram. Halton	No Scrambling	0.0383	0.0560	4.0664	5.1062
	Scrambling	0.0445	0.0646	4.7313	5.9334
Standard Faure	No Scrambling	0.0393	0.0553	4.1668	4.5773
	Scrambling	0.0455	0.0630	4.8227	5.3210
Random Digit Scram. Faure	No Scrambling	0.0298	0.0489	3.1551	4.2517
	Scrambling	0.0432	0.0563	4.5803	5.0752
Random Linear Scram Faure	No Scrambling	0.0364	0.0534	3.9041	4.4663
	Scrambling	0.0310	0.0450	3.2947	4.1762
LHS	N/A	0.0715	0.0789	7.5294	7.6367

Overall, the following inferences can be made regarding the performance of the sequences in 5 dimensions and with 125 draws:

(a) The Braaten-Weller scrambled Halton sequence, in general, does better than the standard Halton, a reversal from the case with 25 draws.

(b) The Braaten-Weller scrambled Halton sequence with no scrambling across observations is the "winner" across all standard and scrambled Halton sequences.

(c) The scrambled versions of the Faure sequence perform better than the standard Halton, the scrambled Halton, and the standard Faure sequences.

(d) The Random Linear scrambled Faure sequence with scrambling across observations performs the best in terms of total error. In terms of bias, the Random Digit scrambled Faure sequence with no scrambling across observations performs the best, although the Random Linear scrambled sequence with scrambling across observations comes a close second.

(e) The LHS yields the highest bias and total error across all the sequences

6.2.3 5 Dimensions and 625 draws

As shown in Table 6, we observe that the standard and scrambled Halton sequences yield lower bias and total error when they are generated with scrambling across observations rather than without scrambling across observations. The same result also extends to the standard Faure and Random Linear scrambled Faure sequences, but the case without scrambling across observations does better than with scrambling across observations for the Random Digit scrambled Faure.

The following inferences can be made regarding the overall performance of the sequences in 5 dimensions and with 625 draws:

(a) The Braaten-Weller scrambled Halton does better than the standard Halton in terms of bias. But in terms of total error, the Braaten-Weller scrambled Halton is better than the standard Halton only for the case when there is no scrambling across observations.

Table 6. Evaluation of ability to recover model parameters (5 dimensions, 625 draws)

Sequence Type	Scrambling across observations	RMSE		MAPE	
		Bias	Total error	Bias	Total error
Standard Halton	No Scrambling	0.0088	0.0189	0.8701	1.6096
	Scrambling	0.0065	0.0161	0.6021	1.3830
Braaten-Weller Scram. Halton	No Scrambling	0.0069	0.0177	0.7053	1.5221
	Scrambling	0.0060	0.0170	0.6013	1.4086
Standard Faure	No Scrambling	0.0070	0.0131	0.7148	1.1309
	Scrambling	0.0047	0.0129	0.3596	1.0538
Random Digit Scram. Faure	No Scrambling	0.0025	0.0138	0.2354	1.1987
	Scrambling	0.0059	0.0174	0.5914	1.4629
Random Linear Scram Faure	No Scrambling	0.0049	0.0161	0.4702	1.4698
	Scrambling	0.0035	0.0152	0.3423	1.2542
LHS	N/A	0.0152	0.0311	1.5890	2.7455

(b) Curiously, the standard Halton with scrambling across observations does the best among the many Halton sequences in terms of total error. However, the Braaten-Weller scrambled Halton with scrambling across observations does almost as well.

(c) The total error values in Table 6 indicate that the standard Faure performs better than the scrambled versions. However, the bias associated with the standard Faure is generally higher than the best alternatives among the scrambled Faure sequences. Among the scrambled Faure sequences, the Random Digit scrambled Faure with no scrambling across observations has the lowest bias and total error values. The Random Linear scrambled Faure sequence with scrambling across observations is the next best alternative among the scrambled Faure sequences.

(d) All the Faure sequences clearly perform better than the Halton sequences in terms of yielding lower bias and total error.

(e) The LHS shows the worst performance across all test scenarios, with the highest bias and total error.

(f) The standard and scrambled Faure sequences exhibit the best performance. While it is not possible to clearly pick a "winner" among the many Faure sequences, it should be noted that the Random Digit scrambled Faure with no scrambling across observations has the lowest bias among all the sequences. The standard Faure sequences yield the lowest total error across all the alternatives, but also yield amongst the highest bias values.

6.2.4 10 Dimensions and 100 draws

The results in Table 7 indicate that the standard Halton sequence exhibits a better performance when it is generated with scrambling across observations, whereas the scrambled Halton sequence performs better when it is generated without scrambling across observations. The standard and scrambled Faure sequences generally exhibit better performances when they are generated without scrambling across observations.

The following conclusions can be drawn regarding the overall performance of the sequences from Table 7:

(a) The standard Halton sequence with scrambling across observations performs better than the standard Halton without scrambling across observations; however, the reverse is the case for the Braaten-Weller scrambled Halton sequence. Overall, the Braaten-Weller scrambled Halton with no scrambling across observations appears to do best.

Table 7. Evaluation of ability to recover model parameters (5 dimensions, 625 draws)

Sequence Type	Scrambling across observations	RMSE		MAPE	
		Bias	Total error	Bias	Total error
Standard Halton	No Scrambling	0.2224	0.2692	26.6145	26.8211
	Scrambling	0.1953	0.2489	23.5067	23.9490
Braaten-Weller Scram. Halton	No Scrambling	0.1681	0.2500	19.8661	21.4625
	Scrambling	0.3297	0.3666	30.2559	30.5939
Standard Faure	No Scrambling	0.1969	0.3114	22.1754	26.5580
	Scrambling	0.2337	0.3068	27.7484	29.8256
Random Digit Scram. Faure	No Scrambling	0.1844	0.2577	21.8181	22.4525
	Scrambling	0.1998	0.2585	24.5396	24.7051
Random Linear Scram Faure	No Scrambling	0.1740	0.2266	20.9043	21.2949
	Scrambling	0.1802	0.2679	20.7861	22.5148
LHS	N/A	0.2213	0.3013	25.6583	26.5579

(b) Among the standard and scrambled Faure sequences, the Random Linear scrambled Faure sequence performs better than the Random Digit scrambled Faure sequence, which in turn performs better than the standard Faure sequence.

(c) Interestingly, in 10 dimensions, the LHS sequence performs comparably with the standard Halton sequence.

(d) There is no clear winner in this case. In terms of total error, the Random Linear scrambled Faure sequence with no scrambling across observations clearly performs the best. In terms of bias, on the other hand, the Braaten-Weller scrambled Halton with no scrambling across observations performs the best. The Random Linear scrambled Faure with no scrambling across observations is, however, close on its heels.

6.2.5 General trends

The different test scenarios of the QMC sequences in 5 dimensions clearly indicate that a larger number of draws results in lower bias, and total error. However, the margin of improvement decreases as the number of draws increases. The following are other key observations from our analysis.

1. At very low draws, the standard versions of the Halton and Faure sequences perform better than the scrambled versions. However, the bias and total error of the estimates is very high and the research results strongly recommend against the use of 25 or less draws in simulation estimation.
2. The scrambled versions of both the Halton and Faure sequences perform better than the standard versions of the sequences at 125 draws (for 5 dimensions) and

100 draws (for 10 dimensions). At 625 draws for 5 dimensions, the standard versions of both the Halton and Faure sequences perform marginally better than their scrambled versions in terms of total error but yield much higher bias. Overall, using about 100-125 draws with scrambled versions of QMC sequences seems appropriate (though one would always gain by using a higher number of draws at the expense of more computational cost).

3. The Faure sequence generally performs better than the Halton sequence across both 5 and 10 dimensions. The only exception is the case of 100 draws for 10 dimensions, which indicates that in terms of bias values the Braaten-Weller scrambled Halton sequence with no scrambling across observations performs slightly better than the Random Linear scrambled Faure with no scrambling across observations. However, this difference is marginal and the Random Linear scrambled Faure clearly yields the lowest total error.
4. Among the Faure sequences, the Random Linear and Random Digit scrambled Faure sequences perform better than the standard Faure (except the case with 25 draws for 5 dimensions, which is anyway not recommended because of high bias and total error values; see point 1 above). However, between the two scrambled Faure versions there is no clear winner.
5. The Random Linear scrambled Faure with scrambling across observations performs better than without scrambling across observations for 5 dimensions (for 125 and 625 draws). For 10 dimensions, the Random Linear scrambled Faure with

scrambling across observations performs slightly less well than without scrambling across observations. However, this difference is rather marginal.

6. The Random Digit scrambled Faure with no scrambling across observations performs better than with scrambling across observations in all the cases.
7. Overall, this analysis concludes that the Random Linear and Random Digit scrambled Faure sequences are amongst the most effective QMC sequences for simulated maximum likelihood estimation of the MMNL model. While both the scrambled versions of the Faure sequence perform well in 5 dimensions, the Random Digit scrambled Faure with no scrambling across observations performs marginally better. In 10 dimensions, on the other hand, the Random Linear scrambled Faure with no scrambling across observations yields the best performance both in terms of bias and total error.
8. This study also strongly recommends the use of the inverse transform to convert uniform QMC sequences to normally distributed sequences.

CHAPTER 7. EMPIRICAL APPLICATION

The comprehensive non-work location choice model proposed in chapter 4 relates spatial interaction, cognitive processes, preferences and decision rules to the observed choice of location through the incorporation of, among other factors, observed and unobserved sources of inter- and intra-individual heterogeneity, feedback and spatial correlation. Such an accurate and behaviorally realistic model structure places significant computational burden on the estimation process. Chapters 5 and 6 address this issue and present the results of research undertaken to identify the most efficient Quasi-Monte Carlo (QMC) sequence for simulated maximum likelihood estimation (SMLE). This chapter presents an empirical application of the proposed location choice model for non-maintenance shopping activities that utilizes the most efficient QMC sequence identified for SMLE.

For an empirical application to fully utilize the potential of the comprehensive model structure developed, a rich data source is needed. The chapter begins with section 7.1 that presents the data sources used in this application. Section 7.2 describes the process of sample formation from the various data sources, while section 7.3 presents a description of the sample. Section 7.4 describes the model formulations that were estimated, section 7.5 presents the categories of variables available for model estimations, and section 7.6 discusses the results of the model estimations. The chapter concludes with section 7.7, which summarizes the results and discusses policy implications of the estimated location choice models.

7.1 Data Sources

The richness of the location choice model structure developed as part of this dissertation research necessitates the use of a rich data source that can utilize the potential of the proposed model structure. Table 8 lists the criteria to be satisfied by the data source in order to exploit the proposed model structure.

Table 8. Criteria to be satisfied by a data source in order to capture various aspects of the proposed model structure

Criteria	Necessary to…
Multi-day Data	To capture (observed & unobserved) intra-personal heterogeneity & state dependence
Detailed Non-work Activity Information	Capture effects of occasion-specific constraints and attributes; Also the proposed model structure is designed only for non-work activities
Detailed Socioeconomic Data	To capture effects of observed sources of inter- and intra- personal heterogeneity
Zonal Layout, Level-of-Service and Land-Use Data	To capture spatial effects
Substantial Data Size	Ensure correct/stable estimation results, and capture different types of heterogeneity
Metropolitan/City Study Region	To capture 'typical' non-discretionary activity patterns
Geocoded Locations	To match individual activity-patterns with the spatial layout

In order to estimate a location choice model that captures the effects of past choices (state dependence), inter- and intra-individual heterogeneity, a multi-day data source is an absolute requirement. While multi-day data surveys are not common in the United States (the 2-day Bay Area Travel Survey being an exception), there are a handful of multi-day datasets from Europe that consist of 5 or more weeks of data. These are: (a) the Uppsala Travel Survey from Uppsala, Sweden, (b) the Mobidrive Survey from the cities of Halle and Karlsruhe in Germany, (c) the SVI Leisure Project from Zürich, Switzerland, (d) the SVI Stabilität Diary from the canton of Thurgau in Switzerland, (e) the ISA Rättfart GPS Study from Borlänge, Sweden, and (f) the AKTA GPS Study from Copenhagen, Denmark (see Schönfelder and Axhausen, 2004, for further details on each of these datasets). The ISA Rättfart and AKTA GPS studies are not of much use in the current context since they do not specifically identify non-work activities. Of the remaining multi-day data sources, the Mobidrive data is the by far the largest in terms of the number of survey respondents and the number of reported trips. The Mobidrive Survey is also one of the most recent studies, contains data on multiple non-work activity types, and includes geocoded location information.

The Mobidrive data is thus the best choice for this empirical application, and is the primary data source used. The Mobidrive data is the result of a 6-week travel survey conducted in the Fall of 1999 in the cities of Karlsruhe (West Germany) and Halle (East Germany), as part of a larger Mobidrive project sponsored by the Federal Republic of Germany Ministry of Research and Education. The main objective of this travel survey data collection was to facilitate a better understanding of the rhythms, routines, and habits of individuals over an extended time period of several weeks. The data collection effort was initiated by contacting a sample of households randomly selected from a phonebook database in each of the two cities. A subsample of this larger sample of households was selected for administration of the travel survey, based on eligibility considerations and willingness to participate (only households who did not plan to take a vacation of more than a week during the survey period and who did not have children under the age of 6 years were deemed eligible).

The final sample from the survey included information on 361 individuals from 162 households. Of these, 44 individuals from 23 households in Karlsruhe participated in a pre-test survey, and 317 individuals from 139 households in Karlsruhe and Halle participated in the main survey. The structure and administration procedures were identical in the two surveys. Both the pre-test and main surveys were conducted in two waves to capture seasonal variations in activity-travel patterns and to avoid the Christmas and summer holidays. The pre-test travel survey was administered between May 31st and July 25th, and the main survey was administered between September 13th and November 14th. Slightly less than 10% of the total sample (approximately 15% of eligible households) participated. Basic information on non-participating households was collected, and research by Axhausen et. al., 2000, has revealed no significant self-selection or fatigue effects.

The survey itself comprised three parts. First, a face-to-face interview was administered where the interviewer assisted the household in filling out three forms (gathering information on the sociodemographic characteristics of households and their members, and car fleet size and composition). Second, a travel diary was mailed to each household and individuals in the household were asked to maintain a record of all their trips and out-of-home activities over a 6-week period. Finally, an attitude questionnaire was administered to participants at least 16 years old once the final week's diary had been turned in (see Schlich et al., 2000, and Axhausen et al., 2000, for more information on the survey data).

While the Mobidrive data provides the non-work activity and travel information, other secondary sources of data that describe the location alternatives are also required for the estimation of the location choice models. The study area (Karlsruhe core city) consists of 69 transportation analysis zones (TAZs), which form the choice set. The models estimated in this study predict the individual choice of travel to these zones and not to specific shopping opportunities (or elemental alternatives) within the zones. This approach was adopted for several reasons. One, the use of elemental alternatives would create a substantial number of alternatives in the individual's choice set. This would pose infeasible data processing requirements, and make the modeling process and definition of alternatives difficult. Two, for transportation planning, the desired end-result is the prediction of trip-interchanges between zonal pairs, not between elemental attraction alternatives. From this standpoint, location choice models with zonal alternatives are easy to apply for forecasting. Three, the attributes of zones are more easily available to the travel demand modeler than the attributes of the elemental units of attraction. Moreover, individuals are likely to perceive elemental units of attraction in clusters (such as shopping districts), and as the size of TAZs are shrinking they may actually match the perceived units of attraction. Figure 15 presents the zonal configuration of the study area, with the shaded zones representing the Central Business District (CBD).

The secondary data sources used in this empirical application include Geographic Information Systems (GIS) files of the transportation network and zonal land-use for the core-city of Karlsruhe. In addition, the empirical analysis uses data on personal business, shopping, recreational and physical activity opportunities in each of the zones in Karlsruhe, collected from the yellow pages.[9]

The following section describes how the primary and secondary sources of data were assembled to prepare a dataset for the empirical application.

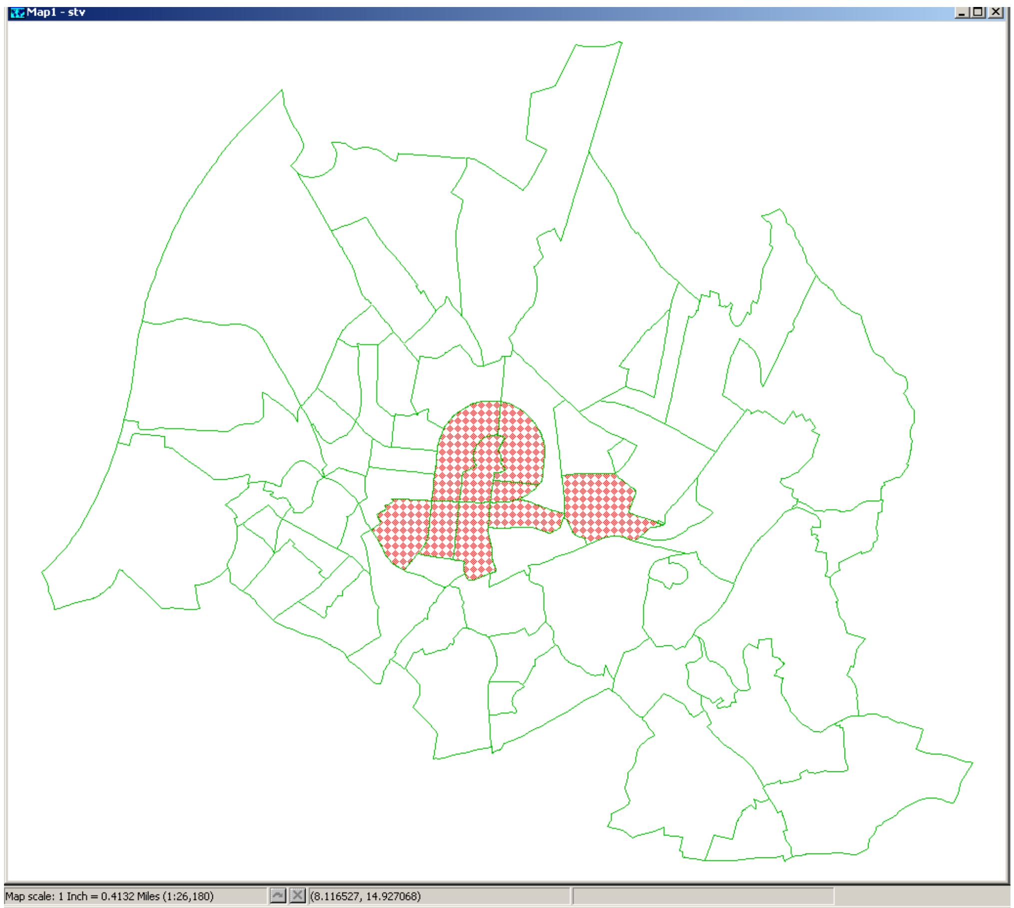

Figure 15. Study Area – 69 core city zones in Karlsruhe

[9] All the land use data was obtained from Dr. Kay Axhausen and Dr. Stefan Schönfelder, at the Institute for Transport Planning and Systems, Swiss Federal Institute of Technology, Zürich.

7.2 Sample Formation

The trip file from the Mobidrive data contains a detailed record of all the trips and out-of-home activities for every individual over a 6-week period. The activity types are classified into 21 categories, including non-work activities such as meeting friends, cultural excursions, active sports, daily shopping and shopping for non-daily demand. Of these non-work categories, shopping for non-daily demand (or non-maintenance shopping) was selected for the empirical application of the proposed location choice model. The choice of non-work activity type was motivated by two main reasons. First, daily shopping, non-maintenance shopping and meeting friends are the most frequently occurring non-work activities in the Mobidrive data (see Table 9) and therefore provide the richest data sources for model estimations. Second, among these three frequently occurring non-work activities, non-maintenance shopping is likely to involve the most variety and discretion in the choice of location (see section 7.3).

The first step in sample formation, therefore, was to extract the non-maintenance shopping activity occasions from the Mobidrive trip file. The second step was to assemble the zonal land-use and level-of-service data from the GIS files. This involved the generation of inter-zonal distances through network skims and the extraction of zonal land-use variables such as area of the zone, area under mixed development, area of roadways, and presence of daycares and churches in the zone. This data provides the zonal attributes and impedances that form the explanatory variables in the estimated location choice models. Since the GIS files, and therefore the land-use and level-of-service data, were available only for Karlsruhe, it was decided to focus the study on survey respondents from Karlsruhe. The third step, then, was to extract the non-maintenance shopping activities undertaken specifically by Karlsruhe survey respondents from the dataset generated in the first step. This corresponded to 1052 choice occasions undertaken by 190 individuals belonging to 93 households in Karlsruhe. The fourth step was to clean this dataset. This involved two primary tasks. First, since land-use and level-of-service data was available only for the 69 core city zones[10], only individuals who live in the core city of Karlsruhe were retained. Also, non-maintenance shopping activities undertaken at external zone locations were excluded for the same reason. Second, the individuals who participated in only one non-maintenance shopping activity during the survey period were excluded. This was done in order to ensure that state dependence and intra-personal heterogeneity effects can be inferred from the data. The cleaned dataset comprises 903 non-maintenance shopping activity occasions undertaken by 158 individuals belonging to 81 households. The final step in sample formation was to append the zonal land-use and impedance data to the cleaned dataset.

7.3 Exploratory Analysis

The final sample thus consists of the non-maintenance shopping episodes undertaken by 158 individuals during a 6-week period, with zonal descriptors for each of the alternative locations. Of these 158 individuals, 55% are female, 53% are employed and 75% have a driving license; 15% belong to single-person households, 28% to couple

[10] The core city zones correspond to the 69 zones in the GIS zonal file for Karlsruhe. The other destination zones in the Mobidrive data are external zones visited by the survey respondents.

families and 31% to nuclear families. The average household income of the sample is 4800 DM.

The 158 individuals participated in 2-29 non-maintenance shopping episodes during the survey period and visited between 1 and 10 unique zones, indicating reasonable intra-individual heterogeneity in location choice. A measure termed as the variety-seeking ratio[11] was computed for the sample of 158 individuals, in order to quantify the degree to which people exhibit heterogeneity in choosing a non-maintenance shopping location. This measure was defined as the number of unique zones visited during the survey period divided by the number of non-maintenance shopping trips undertaken by the individual during the survey period. The variety-seeking ratio thus ranges between 0 and 1 (excluding 0 and including 1), and a higher value implies a greater heterogeneity in location choice. The average variety-seeking ratio for non-maintenance shopping in the assembled sample of 158 individuals is 0.71.

An examination of the average variety-seeking ratios for the various non-work activity types in the Mobidrive data further strengthens the argument for the choice of non-maintenance shopping for this empirical application. Among the non-work activity types that have a sizeable number of reported trips, say greater than 200, non-maintenance shopping has the highest average variety-seeking ratio of 0.71 (see Table 9). Visits to discos, pubs, restaurants, cinemas etc. is very close behind, and has an average

---

[11] Although the measure is termed ‘variety-seeking ratio’, it only captures the degree of heterogeneity in the choice of location. This heterogeneity could be the result of a variety of causes including the effects of constraints and variety-seeking choice behavior.

variety-seeking ratio of 0.70. However, the total number of reported trips (across both Halle and Karlsruhe) is much higher for non-maintenance shopping.

Table 9. Variety-Seeking Ratio by Non-work Activity Type for the Mobidrive Study Area (Halle + Karlsruhe)

Non-work Activity Type	Number of Persons	Total Number of Trips	Variety-Seeking Ratio
Non-Maintenance Shopping	292	1806	0.71
Disco/Pub/Restaurant/Cinema	204	1296	0.70
Meeting Friends	288	2586	0.60
Active Sports	163	1316	0.45
Shopping: Daily Demand	321	4631	0.41
Group/Club Meeting	111	749	0.41
Excursion: Nature	49	157	0.89
Excursion: Culture	52	152	0.84
Meeting Relatives & Family	15	44	0.78

A regression of the variety-seeking ratio for non-maintenance shopping against person and household sociodemographic characteristics captures the correlations between the sociodemographics and observed heterogeneity.

Table 10. Variety-Seeking Ratio Regression Model

Variables	Parameter	t-stat
Constant	0.755	8.158
Female (dummy)	-0.050	-1.303
Number of working hours	-0.001	-0.993
Self-Employed (dummy)	-0.111	-1.329
Number of motor vehicles	0.068	3.035
National (dummy)	-0.085	-1.027
Distance Home to Bus stop (km)	7.36E-05	1.498
Number of children less than 16 years	-0.030	-1.394
Roommate Household (dummy)	-0.330	-2.038
R-Squared	0.122	

The regression results in Table 10 indicate that females exhibit lower heterogeneity in the choice of shopping location than males, while individuals of foreign origin exhibit higher heterogeneity than nationals (as inferred from the coefficient on the national dummy). The coefficients on the employment characteristics are intuitive. They indicate that individuals who work longer hours have lower variety-seeking ratios, as do individuals who are self-employed and consequently handle more responsibilities. On a similar note, people with children less than 16 years are typically more constrained by their responsibilities and therefore exhibit lower variety-seeking ratios. Individuals in roommate households also exhibit less heterogeneity in the choice of non-maintenance shopping location. This is also expected, since individuals who choose to live in roommate households are usually students or people who prefer to live frugally for a variety of reasons, and are hence likelier to exhibit low variety-seeking ratios. Finally, the coefficients on the variables that represent mode availability suggest that individuals who own motor vehicles and live far from bus-stops exhibit more heterogeneity in their location choice. Moreover, the variety-seeking ratio increases with the number of motor vehicles owned and the distance to the bus-stop. This is also intuitive, since individuals who own more vehicles are likelier to live farther away from public transport and such individuals also tend to visit more locations that are rendered accessible to them on account of their higher mobility.

Although the regression model presents intuitive correlations between sociodemographic characteristics and the degree of heterogeneity in location choice, the heterogeneity in the choice of location may be more than just the result of variety-seeking behavior and in order to gain an understanding of choice behavior it is important to understand the sources of heterogeneity. Heterogeneity in choice could be the result of several choice-occasion specific constraints that cannot be captured by the cross-sectional analysis in this regression model. For instance, the choice of non-maintenance shopping location on a specific choice occasion could be influenced by the decision to chain the trip with other activities, such as dinner at a particular restaurant. Heterogeneity could also be the result of spatial correlation effects or variety-seeking behavior. In order to capture the different sources of heterogeneity in the choice of location, it is necessary to estimate location choice models in a panel context. The following sections describe the estimation of location choice models, based on the proposed model structure, that attempt to capture the various observed and unobserved sources (from the analyst's viewpoint) of heterogeneity.

7.4 Model Formulation

The non-maintenance shopping location choice models estimated as a part of this empirical application are based on the comprehensive model structure proposed in chapter 4. A panel data set of non-maintenance shopping episodes undertaken by 158 individuals in the city of Karlsruhe was assembled for the purpose.

Several models were estimated starting with the simple MNL model, extending to the Mixed Logit (MxL) model, and leading up to the Mixed Spatially Correlated Logit (MSCL) model and the Bi-level MxL and MSCL models. The MNL model that does not incorporate state dependence, unobserved heterogeneity or spatial correlation, is estimated as the benchmark against which all other models are evaluated. This MNL model, which is the most restrictive case of the proposed model structure, is widely used in location choice modeling. The other model types estimated are also derived from the general model structure of equation 12 by placing suitable restrictions. The model estimation procedures for these model types are described in chapter 4 (Section 4.4). All estimations and computations were carried out using the GAUSS programming language. Gradients of the log simulated likelihood function with respect to the parameters were also coded.

7.5 Variable Specifications

The assembled panel dataset consists of a variety of variables that can be used to describe the utility associated with each of the alternative core city zones in the various model formulations. The variables include several zonal *size* and non-size attributes, as well as interactions of the socio-demographic characteristics of individuals and attributes of the choice occasions with these zonal attributes. These variables can be categorized into five groups, each of which is discussed in sections 7.5.1-7.5.5. Section 7.5.6 describes the creation of the state dependence (feedback) variable.

7.5.1 Zonal Size Attributes

There are several zonal *size* attributes in the assembled dataset that capture the attractiveness of a zone. These include: (i) zonal area, (ii) area of the zone covered by parking, (iii) area of the zone covered by roadways, (iv) area of the zone covered by industries, (v) area of the zone covered by mixed development, (vi) area of the zone covered by commercial/trade related enterprises, (vii) population of the zone, (viii) number of shopping opportunities, (ix) number of personal business opportunities, (x) number of recreational opportunities, and (xi) number of physical activity opportunities. Since zonal *size* attributes indirectly represent the elemental alternatives within a zone, it is important to introduce them in a form that ensures that a large zone will have a higher probability of being chosen than a small zone. For this purpose a non-linear composite size measure is introduced, which is defined as follows (a similar approach is used in Pozsgay and Bhat, 2002).

$$CompositeSize_j = \ln(x_{1j} + \delta_1 x_{2j} + \delta_2 x_{3j} + \delta_3 x_{4j} + ...), \qquad \textbf{Eq. 50}$$

where, (x_{1j}, x_{2j}, x_{3j}, x_{4j}, …) are zonal *size* attributes, and $(\delta_1, \delta_2, \delta_3, ...)$ are parameters to be estimated. Nonlinear-in-parameters MNL (NLMNL) models with the following utility expression were estimated and the best specification was identified for the composite size term (see Table 11).

$$U_{ijt} = \alpha\ distance_{ijt} + \beta\ Composite\ Size_j + \varepsilon_{ijt} \qquad \textbf{Eq. 51}$$

Table 11. Best specification model for the composite size term

Variables	**NLMNL**	
	Param.	t-stat.
Distance from home	-1.771	-28.518
Composite Size	0.638	17.125
No. of shopping opportunities	2.558	1.810
No. of recreational opportunities	89.779	2.514
Area covered by mixed developments	17.088	2.341
Number of observations	62307	
Log-likelihood at convergence	-3052.663	

The best specification for the composite size includes the number of shopping and recreational opportunities and the area of the zone covered by mixed developments. In other words, a zone with a greater number of shopping and recreational opportunities and more mixed developments is assigned a larger composite size. The other zonal *size* measures did not turn out to be significant in this specification, largely due to the correlations between these *size* measures. In all the models estimated in this study, the composite size computed with the parameters and variables identified in Table 11 is used as the only zonal *size* measure.

7.5.2 Zonal Non-size Attributes

The zonal non-size attributes include population density, central district (dummy), presence of daycare (dummy), and presence of church (dummy). These variables, along with the composite size, are introduced in the location choice model as measures of zonal attractiveness. The church dummy was excluded from the model specifications because nearly every zone in the study area has a church and this correlation could show up in the models as a spurious effect. All the other non-size attributes were significant in all the models estimated.

7.5.3 Zonal Impedance Measures

The distance of each zone from the home zone of an individual is treated as the impedance associated with that zone. Zones which are farther away from an individual's home zone will be less preferred, and therefore distance from home is a good impedance (or cost) variable.

Several studies have shown that people tend to visit locations that are either around their home or around their school/work place for non-work activities (see Schonfelder and Axhausen, 2002). In keeping with this, the distance of zones from the work/school zone of an individual is introduced as another impedance measure.

7.5.4 Socioeconomic and Demographic Variables

The zonal composite size and non-size attributes, and impedance measures were interacted with individual sociodemographic characteristics in order to capture the observed sources of heterogeneity across individuals in their response to the zonal attributes and impedance. For instance, the parameter estimated on the distance interacted with gender captures the difference in sensitivities to the impedance measure across males and females.

The various sociodemographic characteristics available in the data include age, gender, marital status, employment status, schooling status, number of working hours, education level, license (dummy), club member (dummy), national (dummy), number of season tickets held, household size, number of children less than 16 years in the household, number of dogs, number of automobiles, number of cycles, household type, household income, and distance from home to bus-stop, light rail transit (LRT) and heavy rail.

After some testing, the household type variables were excluded from the model specifications as they are correlated with other household attributes such as household income and marital status. All the other sociodemographic variables were tested in the model estimations and the significant ones retained.

7.5.5 Attributes of Choice Occasions

The zonal *size* and non-size attributes, and impedance measures were also interacted with attributes of the choice occasions in order to capture time-dependent effects and constraints on the response to zonal attributes and impedances. For example, the parameters estimated on distance interacted with time of day variables capture the time-varying effects of time-budget constraints on the sensitivity to impedance.

The attributes of the choice occasions available in the data include time-of-day, day of week, mode chosen, trip chaining (dummy), number of household members accompanying, number of others accompanying, and activity duration. The mode chosen for the choice occasion was found to be highly correlated with the choice of zone. However, after careful consideration it was decided to exclude this variable from the models estimated, since mode and destination are often simultaneous decisions and warrant joint mode-destination choice models. Moreover, due to the strong correlation between mode and the chosen zone other explanatory variables drop out when mode is introduced rendering it difficult to estimate a good model for forecasting purposes.

7.5.6 Feedback Effects

The proposed model structure (equation 12) introduces feedback through the terms $PREATT_{ijt}$, which is a function of the similarities between zone *j* and the attributes of previously chosen alternatives (on choice occasions *t-1, t-2,…,1*), and $PRECHO_{ijt}$, which is a function of the number of times zone *j* has been chosen on choice occasions *t*-1, *t*-2,…,1. In this application, a simple form of the $PRECHO_{ijt}$ function, $SAME_{ijt}$ (first-order Markov process, first order state dependence, or lagged choice indicator), is used, which is defined as follows (a similar approach is used by Miller and O'Kelly, 1983).

$$SAME_{ijt} = \begin{cases} 1, & \text{if zone } j \text{ was chosen on the choice occasion } t-1 \\ 0, & \text{otherwise} \end{cases}$$

The introduction of feedback in a model must be accompanied by a specification of the initial conditions. This study makes two assumptions regarding the initial conditions. First, it assumes that the survey respondents have reached a state of equilibrium in their activity-travel patterns, so that the survey period will be representative of their choice behavior. Second, the first non-maintenance shopping episode of each survey respondent is assumed to be exogenous to the estimation.

7.6 Empirical Results

As discussed earlier, a basic MNL model of location choice was estimated as the benchmark against which all other models were compared. This model (MNL-1) does not incorporate state dependence, unobserved heterogeneity or spatial correlation. An MNL model with state dependence (MNL-2) was also estimated to assess the impacts of introducing feedback. Table 12 presents the best specification MNL models of both types. The MNL-1 model was then extended to incorporate unobserved inter-individual response heterogeneity, which yields a mixed logit model (MxL-1). MxL-2 is a mixed logit model that incorporates state dependence in addition to unobserved inter-individual response heterogeneity. The results of the MxL model estimations are presented in Table 13. The MNL-2 and MxL-2 models were further extended to incorporate spatial correlation effects, which produces the SCL and MSCL models respectively. However, as the SCL and MSCL model estimation results in Table 14 indicate, there are no significant spatial correlation effects in the study area. Finally, Bi-level Mixed Logit (BiMxL) models that incorporate unobserved intra-individual response heterogeneity, in addition to the unobserved inter-individual response heterogeneity in the MxL models, were also estimated. The results of the BiMxL model estimations are, however, not presented here since unobserved intra-individual heterogeneity in the assembled sample was found to be insignificant.

The following sections compare the above models in terms of their goodness-of-fit, discuss the responses of individuals to the zonal attractiveness measures and the sensitivities to the zonal impedance measures, and examine the effects of feedback and unobserved inter-individual heterogeneity. The potential causes for the absence of unobserved intra-individual heterogeneity and spatial correlation are also discussed here.

Table 12. Best Specification Multinomial Logit Models of Location Choice

Variables	MNL-1		MNL-2	
	Param.	t-stat.	Param.	t-stat.
Attributes of Alternatives				
Composite Size of zone	0.254	4.465	0.196	2.996
Central District (dummy)	0.953	2.758	0.833	2.067
Presence of Daycare (dummy)	-0.286	-3.001	-0.195	-1.85
Population density of zone	-0.008	-9.274	-0.006	-6.371
Impedance Measures assoc. with Alternatives				
Distance of zone from home	-3.397	-4.959	-3.078	-3.446
Distance of zone from work/school	-1.035	-8.95	-0.861	-6.626
Interactions between Zonal Attributes and Impedance				
Composite Size of zone x Distance of zone from home	0.283	4.994	0.253	3.831
Interactions of Zonal Attributes with Sociodem.				
Composite Size of Zone x				
Low Income (dummy)	0.273	4.209	0.211	2.869
No. Season Tickets	0.125	2.291	0.108	1.71
Central District (dummy) x				
Age	-0.009	-1.837	-0.005	-0.954
Female (dummy)	0.262	1.729	0.045	0.263
Presence of Daycare (dummy) x				
Married (dummy) x No. of children	0.688	2.661	0.592	2.236
Interactions of Zonal Attributes with Choice Occasion-Specific Constraints				
Composite Size of Zone x				
Trip Chained with Other Shopping (dummy)	-0.149	-2.363	-0.112	-1.608
Central District (dummy) x				
Trip Chained with Other Activities (dummy)	0.323	1.973	0.349	1.846
No. of other Accompanying Adults	-0.331	-2.846	-0.400	-2.862
Activity Duration	0.003	3.137	0.005	3.765
Time-of-day Morning (dummy)	-0.409	-1.882	-0.428	-1.723
Interactions of Zonal Impedance with Sociodem.				
Distance of zone from home x				
Female (dummy)	-0.491	-2.94	-0.267	-1.363
Retired (dummy)	-0.796	-3.05	-0.877	-2.73
No. of Cars owned	0.373	2.801	0.211	1.336
National (dummy)	-0.859	-1.665	-0.624	-0.884
No. Season Tickets owned	-0.425	-2.046	-0.288	-1.203
Interactions of Zonal Impedance with Choice Occasion-Specific Constraints				
Distance of zone from home x				
Weekend (dummy)	-0.499	-2.315	-0.462	-1.957

No. of other Accompanying Adults	0.468	3.714	0.362	2.546
Trip Chained with Other Activities (dummy)	0.446	2.657	0.398	2.058
Activity Duration	0.003	3.062	0.002	1.401
State Dependence Variables				
First Order Feedback of Chosen Zone			1.675	17.529
Number of observations	62307		51405	
Log-likelihood at convergence	-2862.371		-2250.881	
Log-likelihood at equal shares	-3823.398		-3154.409	

Table 13. Best Specification Mixed Logit Models of Location Choice

Variables	MxL-1		MxL-2	
	Param.	**t-stat.**	**Param.**	**t-stat.**
Attributes of Alternatives				
Composite Size of zone	0.346	4.756	0.283	3.439
Central District (dummy)	1.133	2.469	0.820	1.395
Presence of Daycare (dummy)	-0.525	-3.337	-0.332	-1.897
Population density of zone	-0.017	-8.841	-0.012	-6.532
Impedance Measures assoc. with Alternatives				
Distance of zone from home	-3.363	-3.71	-3.002	-2.797
Distance of zone from work/school	-1.087	-5.837	-0.975	-4.628
Interactions between Zonal Attributes and Impedance				
Composite Size of zone x Distance of zone from home	0.239	3.855	0.195	2.484
Interactions of Zonal Attributes with Sociodem.				
Composite Size of Zone x				
Low Income (dummy)	0.264	2.472	0.212	1.604
No. Season Tickets	0.161	1.845	0.151	1.514
Central District (dummy) x				
Age	-0.014	-2.087	-0.008	-1.028
Female (dummy)	0.204	0.892	0.113	0.413
Presence of Daycare (dummy) x				
Married (dummy) x No. of children	0.699	0.999	0.665	0.878
Interactions of Zonal Attributes with Choice Occasion-Specific Constraints				
Composite Size of Zone x				
Trip Chained with Other Shopping (dummy)	-0.188	-2.427	-0.145	-1.744
Central District (dummy) x				
Trip Chained with Other Activities (dummy)	0.273	1.318	0.316	1.411
No. of other Accompanying Adults	-0.331	-2.165	-0.403	-2.148
Activity Duration	0.003	2.557	0.006	3.786
Time-of-day Morning (dummy)	-0.187	-0.696	-0.280	-0.969
Interactions of Zonal Impedance with Sociodem				
Distance of zone from home x				
Female (dummy)	-0.248	-1.011	-0.021	-0.078
Retired (dummy)	-0.780	-2.517	-0.815	-2.297
No. of Cars owned	0.254	1.287	0.077	0.322
National (dummy)	-0.831	-1.314	-0.602	-0.800
No. Season Tickets owned	-0.391	-1.479	-0.160	-0.534
Interactions of Zonal Impedance with Choice Occasion-Specific Constraints				
Distance of zone from home x				
Weekend (dummy)	-0.544	-1.933	-0.511	-1.710

No. of other Accompanying Adults	0.460	2.541	0.409	2.146
Trip Chained with Other Activities (dummy)	0.411	1.998	0.398	1.595
Activity Duration	0.004	1.903	0.003	1.171
State Dependence Variables				
First Order Feedback of Chosen Zone			1.361	10.531
Std. Deviation in Response to				
Composite Size of zone	0.297	6.153	0.174	2.257
Central District (dummy)	0.693	4.092	0.648	3.191
Presence of Daycare (dummy)	0.789	4.049	0.647	2.452
Population density of zone	0.013	5.641	0.010	3.112
Distance of zone from home	0.603	3.228	0.506	1.846
Distance of zone from work/school	0.688	2.760	0.377	1.040
First Order Feedback of Chosen Zone			0.549	3.335
Number of observations	62307		51405	
Log-likelihood at convergence	-2786.861		-2223.456	

Table 14. Best Specification Spatially Correlated Logit Models of Location Choice

Variables	SCL		MSCL	
	Param.	t-stat.	Param.	t-stat.
Attributes of Alternatives				
Composite Size of zone	0.201	2.128	0.321	2.551
Central District (dummy)	0.837	1.482	0.838	1.036
Presence of Daycare (dummy)	-0.282	-1.663	-0.339	-1.252
Population density of zone	-0.012	-6.614	-0.019	-5.829
Impedance Measures assoc. with Alternatives				
Distance of zone from home	-3.632	-3.586	-3.622	-2.669
Distance of zone from work/school	-0.871	-4.881	-1.005	-3.95
Interactions between Zonal Attributes and Impedance				
Composite Size of zone x Distance of zone from home	0.390	4.554	0.292	2.546
Interactions of Zonal Attributes with Sociodem.				
Composite Size of Zone x				
Low Income (dummy)	0.275	2.082	0.302	1.619
No. Season Tickets	0.147	1.581	0.261	1.843
Central District (dummy) x				
Age	-0.006	-0.855	-0.009	-0.827
Female (dummy)	0.118	0.609	0.314	0.904
Presence of Daycare (dummy) x				
Married (dummy) x No. of children	0.932	1.135	0.915	0.995
Interactions of Zonal Attributes with Choice Occasion-Specific Constraints				
Composite Size of Zone x				
Trip Chained with Other Shopping (dummy)	-0.157	-1.733	-0.220	-2.01
Central District (dummy) x				
Trip Chained with Other Activities (dummy)	0.448	1.758	0.381	1.242
No. of other Accompanying Adults	-0.457	-2.353	-0.469	-1.903
Activity Duration	0.008	4.644	0.009	4.931
Time-of-day Morning (dummy)	-0.385	-1.383	-0.214	-0.569
Interactions of Zonal Impedance with Sociodem				
Distance of zone from home x				
Female (dummy)	-0.395	-1.772	-0.069	-0.218
Retired (dummy)	-1.057	-3.337	-0.955	-2.151
No. of Cars owned	0.240	1.269	0.070	0.242
National (dummy)	-1.103	-1.478	-0.927	-1.029
No. Season Tickets owned	-0.210	-0.853	-0.052	-0.147
Interactions of Zonal Impedance with Choice Occasion-Specific Constraints				
Distance of zone from home x				
Weekend (dummy)	-0.665	-2.329	-0.689	-1.966

No. of other Accompanying Adults	0.421	2.204	0.494	2.272
Trip Chained with Other Activities (dummy)	0.529	2.277	0.558	1.921
Activity Duration	0.003	1.357	0.003	1.225
State Dependence Variables				
First Order Feedback of Chosen Zone	2.409	12.563	1.970	8.032
Std. Deviation in Response to				
Composite Size of zone			0.221	1.463
Central District (dummy)			0.901	2.413
Presence of Daycare (dummy)			0.580	1.162
Population density of zone			0.012	2.342
Distance of zone from home			0.612	1.973
Distance of zone from work/school			0.451	0.972
First Order Feedback of Chosen Zone			0.898	2.783
Spatial Correlation Effects as Captured by				
Dissimilarity Parameter	3.742	4.184	3.936	4.193
Number of observations	51405		51405	
Log-likelihood at convergence	-2229.594		-2201.645	

7.6.1 Comparison of Goodness-of-Fit

The log-likelihood value at convergence for the MNL-1 model with 26 parameters is -2862.4, while the corresponding value for the MNL feedback model (MNL-2) with 27 parameters is -2250.9 (Table 12). The log-likelihood for the naïve model that assigns equal shares to all the zones is -3823.4. Clearly, the simplest MNL model, MNL-1, is better than the naïve model (the likelihood ratio test statistic is of the order of 1922.0, which is much larger than the chi-squared statistic with 26 degrees of freedom at any reasonable level of significance). Moreover, a likelihood ratio test between the MNL models with and without feedback (MNL-1 and MNL-2) indicates significant effects of past choices on current choice behavior (the likelihood ratio test statistic is of the order of 1223.0, which is greater than the chi-squared statistic with 1 degree of freedom at any reasonable level of significance).

The likelihood ratio test between the Mixed Logit model without feedback (MxL-1) and the corresponding MNL model (MNL-1) indicates statistically significant unobserved response heterogeneity across individuals (the likelihood ratio test statistic is of the order of 151.0, which is greater than the chi-squared statistic with 6 degrees of freedom at any reasonable level of significance). A comparison between the models MxL-2 and MxL-1 further indicates the presence of significant feedback effects in addition to the unobserved inter-individual heterogeneity (the likelihood ratio test statistic is of the order of 1127.0, which is greater than the chi-squared statistic with 2 degrees of freedom at any reasonable level of significance).

Although a statistical comparison of the log-likelihoods of the Mixed Spatially Correlated Logit model (MSCL) and the MxL-2 model seems to indicate significant spatial correlation effects over and above the unobserved inter-individual heterogeneity and feedback effects, both the spatial correlation models SCL and MSCL are actually rejected since the dissimilarity parameters estimated are invalid. This will be further discussed in the section on spatial correlation effects (Section 7.6.5).

7.6.2 Effects of Zonal Attractiveness and Impedance Measures

As shown in Tables 12 and 13, the responses to the zonal attributes (and sensitivities to the zonal impedances) captured by the estimated parameters remain more or less the same across the different model types. The responses and sensitivities to the various zonal attributes and impedances are discussed below based on the MNL-1 estimates in Table 12. Any deviations in these results across the model types are also discussed.

7.6.2.1 Composite Size Measure

The zonal composite size measure (defined in section 7.5.1) has a positive coefficient indicating that larger composite size zones are preferred more than zones of smaller composite size. This is to be expected since larger composite size zones contain more elemental units of attraction such as shopping malls and recreation centers.

The parameter on the interaction term of composite size with low-income indicates that individuals belonging to low income households show a higher preference for larger composite size zones than other individuals. It could be that low-income individuals prefer to comparison shop for the best value for money and larger composite size zones provide more opportunities for the purpose. Also, individuals who own season tickets for public transport show a higher preference for larger composite size zones than others. Since large zones with lots of mixed development are typically better connected by public transport than zones with fewer opportunities, this is also an intuitive result. All interactions of composite size with other sociodemographic characteristics turned out to be statistically insignificant.

The interaction of the composite size measure with choice occasion-specific constraints yielded only one significant term. The parameter on this term indicates that larger composite size zones are less preferred when a non-maintenance shopping activity is chained with other shopping activities. This is again intuitive as individuals usually pick convenient locations that can be combined with other destinations when they chain trips.

The effects of the composite size measure on the utility of a zone and the observed sources of heterogeneity in these effects, vary very little across the different models. Overall, according to the MNL models, composite size has a positive effect on the utility of a zone for all the individuals. The MxL-1 and MxL-2 models, on the other hand, suggest that composite size may actually have a negative effect on the utility of a zone for 2-10% of the individuals. In other words, depending on the sociodemographic characteristics, between 2% and 10% of the individuals may prefer smaller composite size zones to larger composite size zones. This will be discussed further in the section on unobserved heterogeneity.

7.6.2.2 Zonal Non-Size Attributes

Population density, central district (dummy) and presence of daycare (dummy) are the zonal non-size attributes that were significant in the MNL model estimations. While the central district and daycare dummies drop in significance when feedback and unobserved heterogeneity are incorporated, population density remains statistically highly significant across all the model types. The estimated effects of each of these variables on the utility of a zone are discussed below.

A central district zone is preferred more for non-maintenance shopping activity participation than non-central zones, as indicated by the parameter in Table 12. Among the sociodemographic characteristics, the observed sources of heterogeneity that are statistically significant include gender and age. While females seem to prefer central zones more than males, older individuals prefer central zones lesser than younger individuals. The overall effect of a central zone on the utility is clearly positive according to the MNL models. However, there are several choice occasion-specific constraints that also influence the utility of a central zone. A central zone is less preferred for non-maintenance shopping activities undertaken in the morning, which is reasonable since most people would prefer to avoid the morning traffic in a central district. A central zone is also less preferred when an individual has company in traveling to the non-maintenance shopping location (as captured by the interaction term 'Central District x No. of Accompanying Adults'). This could be capturing different kinds of constraints and group dynamics. For instance, when people shop in groups they might want to experiment with new or unfamiliar locations, or they might prefer to shop close to someone's home in order to make drop-offs convenient. On the other hand, there is an almost equal (to group travel) positive effect on the utility of a central zone when the shopping activity is chained with other activities. Finally, the longer the duration of the proposed non-maintenance shopping activity the higher the utility associated with central zones.

Zones with daycare are preferred less than other zones, as indicated by the parameter in Table 12. Although the reason for this is not easily apparent, this could be the result of high correlation between residential zones (with few shopping opportunities) and the presence of daycare. What is more intuitive is the fact that married people with children less than 16 years of age prefer zones with daycare for non-maintenance shopping[12] (as seen from the MNL-1 estimation results in Table 12, the presence of daycare deducts 0.286 from the utility of a zone for most people, except married individuals who have children, in which case the presence of daycare adds 0.688-0.286 = 0.402, to the utility of a zone). Although the daycare interaction term is observed in all the model types, the interaction term drops in statistical significance when feedback and unobserved heterogeneity are included in the model.

High population density zones are preferred less than other zones by most people. This is reasonable, since high population density zones are primarily residential and the few neighborhood shopping opportunities in such zones are mostly favored only by the residents of the zone.

---

[12] Since the data only contains households with children above 6 years of age, the parameter on the interaction term indicates that individuals with older children may be more familiar with zones containing daycare facilities, perhaps due to past experiences, and are therefore likelier to visit these zones.

7.6.2.3 Impedance Measures

Two impedance measures, distance from home and distance from work/school, were introduced in the model specifications and they were both significant in all the model estimations. The estimated parameters indicate a strong disutility associated with these variables, which is intuitive. Most people prefer to visit locations in the vicinity of their homes, schools and work places. Therefore, zones that are farther away from these locations are less preferred.

The ‘distance from home’ variable was interacted with several sociodemographics and choice occasion-specific constraints, some of which proved to be significant. Females demonstrate a higher dispreference for zones that are farther from home than males, as do nationals over non-nationals. Retired individuals also prefer zones closer to home than others. The more the number of cars owned by an individual the lower his/her dispreference for farther zones, which is intuitive since individuals who own cars are less constrained in their ability to travel. On the other hand, individuals with more number of season tickets for public transportation demonstrate an equally higher dispreference for farther zones. Some of these interactions with sociodemographic characteristics, however, become statistically insignificant when other effects such as feedback and unobserved heterogeneity are included.

Choice occasion-specific constraints also significantly influence the dispreference associated with distance, and these effects remain significant across all the model types. Individuals show a higher dispreference to traveling longer distances for non-maintenance shopping activities undertaken during the weekend compared to a weekday.

This is contrary to our expectations. Perhaps people tend to visit locations closer to their work places during the weekdays, which may be a longer distance from their homes. The other interaction terms are intuitive, however, and indicate that people exhibit a lower dispreference for farther zones when they are accompanied by other people, when they chain the shopping activity with other activities and when they plan shopping activities of longer durations. Similarly, people exhibit a lower dispreference for zones that are farther away if the zones have a larger composite size. In other words, the choice between travel distance and availability of shopping opportunities is a trade-off.

7.6.3 Feedback Effects

A comparison of the MNL models in Table 12 indicates that the effect of past choices on the utility of a zone is highly significant and positive. Therefore, on a specific choice occasion, all else being equal, zones visited in the previous choice occasion are preferred over other zones. This implies a habit persistence or loyalty choice behavior. It is important to include feedback effects in location choice models not only to capture this behavior but also to ensure that all the other parameters are correctly estimated. A comparison of the two MNL models shows that in the absence of feedback several parameters are over-estimated (Heckman, 1981, and Hsiao, 1986, discuss this issue in detail). In the absence of feedback, the effects of past choices are spuriously assigned to the zonal attributes. For instance, central zones and larger composite size zones are assigned a higher utility in the absence of feedback, as evidenced by a comparison of the relative magnitudes of these parameters in Table 12. Also, variables such as the interaction between gender and central district are rendered statistically insignificant when feedback is introduced, indicating that true choice behavior cannot be captured unless feedback effects are also included.

The importance of including feedback is apparent even from the mixed logit models presented in Table 13. The effect of feedback is highly significant and positive, in the absence of which parameters are often over-estimated (compare the columns of parameters under MxL-1 and MxL-2). This is particularly important when estimating unobserved inter-individual heterogeneity, as will be discussed in the following section.

7.6.4 Unobserved Heterogeneity

The mixed logit models presented in Table 13 distinguish between the effects of observed and unobserved sources of heterogeneity, unlike the MNL models that only identify the effects of observed sources of heterogeneity in location choice. As shown in Table 13, all the standard deviations associated with the responses to zonal attributes are highly significant, indicating significant effects of unobserved response heterogeneity. Further, it is also evident from the MxL-2 model that there exists significant unobserved heterogeneity in the state dependence effect, although the net feedback effects are positive for 97.7% of the individuals. It is not surprising, therefore, that the model fit statistics indicate the mixed logit models are a better fit for the data than the corresponding MNL models (see Section 7.6.1).

Consider the composite size measure in the mixed logit model without feedback (MxL-1). For individuals from medium and high income households who do not own any season tickets, the estimated mean and standard deviation of the composite size measure, on a choice occasion when the non-maintenance shopping activity is not part of a trip chain, are 0.346 and 0.297. Or, in other words, the response of such individuals to the composite size measure can be drawn from a normal distribution N(0.346,0.297). This implies that 87.7% of these individuals find larger composite size zones more attractive than smaller composite size zones and the degree of attractiveness varies across individuals. In contrast, the MNL model only captures the average response to composite size. According to the MNL-1 model, all individuals find larger composite size zones more attractive than smaller composite size zones and the degree of attractiveness is fixed at 0.254 per composite size unit. Clearly, the Mixed Logit model captures the mechanism of choice behavior more realistically than the MNL model.

The variances of the unobserved heterogeneity terms provide important information regarding the fraction of variation in the utility associated with a zone for non-maintenance shopping. For instance, consider the Mixed Logit model in equation (24). The variation in utility for this model is given as

$$Var[U_{ijt}] = Var[\delta_1 X_i Z_j] + Var[\delta_2 C_{it} Z_j] + Var[\delta_3 C_{it} X_i Z_j] + Var[\alpha_i Z_j] + Var[\xi_i L_{jt}] + Var[\varepsilon_{ijt}]$$

Eq. 52

The first three terms are the variances due to covariates (that is, the observed sources of inter- and intra-individual heterogeneity), the next two terms are the variances due to unobserved sources of inter-individual response heterogeneity and the last term is the variance due to unobserved sources of intra-individual heterogeneity. Since the last term is not known[13], the contribution of each of the other sources of heterogeneity can only be computed relatively.

[13] The error term is gumbel distributed with location parameter 0 and scale parameter β. The variance of the error term is therefore $\pi^2\beta^2/6$.

Computation based on the MxL-1 and MxL-2 models indicates that the variation in utility explained by unobserved inter-individual response heterogeneity is about 3.75 times that explained by the effects of covariates. A model that incorporates unobserved inter-individual response heterogeneity is therefore a more accurate model. Further, a comparison of the MxL-1 and MxL-2 models shows that it is important to incorporate unobserved heterogeneity in a model with state dependence and vice versa, since each can manifest itself spuriously as the other (see Heckman, 1981, Keane, 1997, Bhat and Castelar, 2002). In addition, ignoring state dependence or unobserved heterogeneity will generally lead to a bias in the effect of the other coefficients in the model (Heckman, 1981, Hsiao, 1986).

Bi-level Mixed Logit models that incorporate intra-individual response heterogeneity in addition to inter-individual response heterogeneity were also estimated as a part of this empirical analysis. The bi-level mixed logit models attempt to capture unobserved (from the analyst's viewpoint) sources of heterogeneity in location choice both across individuals as well as across different choice occasions of an individual. It is intuitive that just as different individuals may react differently to the same situation for no observable reason (attitude differences), an individual may also react differently on different choice occasions for no observable reason (mood dependent reactions). However, unobserved intra-individual response heterogeneity proved to be statistically insignificant in all the bi-level models estimated. It is possible that the interactions of the zonal attributes with choice-occasion specific constraints capture the intra-individual heterogeneity very effectively and therefore the remaining unobserved sources of intra-individual heterogeneity have insignificant effects on the choice of location. On the other hand, it is likelier that this is a data limitation. A larger sample than the 158 individuals used in this application with more choice occasions per individual might be required to estimate unobserved intra-individual heterogeneity.

7.6.5 Spatial Correlation

The SCL and MSCL models presented in Table 14 are essentially the MNL-2 and MxL-2 models with spatial interaction incorporated. Although the SCL and MSCL models appear to be a better fit for the data, they are both rejected since the estimated dissimilarity parameters are significantly greater than 1. The dissimilarity parameter ρ in the SCL and MSCL models is required to lie between 0 and 1 in order to satisfy the GEV conditions. This condition can be traced back to the requirement that the variance of the joint alternatives be identical in the GEV model (see Koning and Ridder, 2003, for details). A smaller dissimilarity parameter indicates high levels of spatial correlation, whereas a dissimilarity parameter of 1 indicates zero spatial correlation. A dissimilarity parameter greater than 1, on the other hand, is not consistent with random utility maximization and should be rejected. There has been some debate over this issue and it is believed by some researchers that the general conditions in determining consistency with utility maximization are too stringent (see, for instance, Kling and Herriges, 1996). In this empirical study it was decided to proceed with the norm and reject models with dissimilarity parameters greater than 1. In effect, the assumption is that a dissimilarity parameter greater than 1 indicates the absence of significant spatial correlation.

Based on the estimated SCL and MSCL models, therefore, it may be concluded that there are no significant spatial correlation effects in the study area (Karlsruhe core city). What this means is that there is no correlation between the unobserved errors in the utilities associated by individual *i* with zones that are adjacent. While the absence of spatial correlation is rare for spatial data, it is possible under certain conditions. The spatial correlation between a pair of zones is dependent on the distance between the zones. So if the zones in the study area are large, the distance between adjacent zones would be correspondingly larger and the strength of correlation would be low. In other words, when the variability in land use is at a scale smaller than the distance between the zones, there is no spatial correlation between the zones. It is also reasonable to expect low spatial correlation if the zonal boundaries are well-defined, in that they completely enclose land-use parcels and the land use changes across zones. In such a case, the zones would be distinct from each other in their attractiveness for different purposes (such as shopping) and spatial correlation would thus be absent.

The core city of Karlsruhe is a fairly small region of area approximately 15.6 sq.km with a mature transportation system and tight land-use control. It is therefore conceivable that the zonal configuration creates clear boundaries between different land-use parcels. It also appears that the goods on offer in the various zones in Karlsruhe are rather distinct (based on discussion with Dr. Kay Axhausen). The non-maintenance shopping opportunities in Karlsruhe are focused on the CBD, which primarily sells fashion and expensive goods, and two minor centers in the east and the west (Durlach and Mühlburg, respectively), which sell goods in the middle price range. Under these conditions it is not unreasonable that the model estimations suggest the absence of spatial correlation in the study area.

7.7 Summary and Policy Implications

The location choice model structure proposed in chapter 4 comprehensively incorporates the various factors that contribute toward heterogeneity in the choice of location, including observed and unobserved sources of inter- and intra-individual heterogeneity, feedback effects and spatial correlation effects. Whether location choice models that are estimated based on this model structure capture all these effects, though, depends entirely on the zonal configuration of the study area and the quality and quantity of the observed choice data.

The empirical application presented here is based on the proposed model structure and uses the non-maintenance shopping activity information from the Mobidrive data, the best panel data source available for the purpose. The model estimations indicate the absence of spatial correlation effects in the study area as well as the absence of unobserved intra-individual heterogeneity. The best model estimated is a Mixed Logit model that incorporates observed and unobserved sources of inter-individual heterogeneity, and feedback effects (MxL-2). While the inability of the model estimations to capture unobserved intra-individual heterogeneity may be a result of the small data size (158 individuals and 745 choice occasions), the absence of spatial correlation is probably due to the zonal configuration of the study area (as discussed in the previous sections).

The model estimations in this study are further handicapped by the nature of the sample data. There is a substantial imbalance in the observed choice toward one of the zones in the CBD (zone 21 chosen on 35% of the choice occasions). Presumably, this zone is the primary source of non-maintenance shopping opportunities in the study area, a fact that is not sufficiently substantiated by the zonal land-use data. According to the zonal land-use data, zone 21 is a central district zone with no daycare facilities and a fair number of shopping opportunities, a description that could match several other CBD zones each of which is observed to be chosen only on 1%-8% of the choice occasions. The only zonal attribute that distinguishes zone 21 from other CBD zones is the number of recreational opportunities, specifically, the number of museums[14]. The model estimations are therefore handicapped by the contrast between the large fraction of travel to zone 21 and the medium level of attractiveness of the zone with respect to relevant zonal attractiveness measures.

Despite the data limitations and the inability of the models to capture unobserved intra-individual heterogeneity and spatial correlation, the models estimated as part of this empirical application are fairly accurate and sensitive to different policy scenarios. The best non-maintenance location choice models estimated in this application are the MNL models, MNL-1 and MNL-2, and the MxL models, MxL-1 and MxL-2. The most commonly applied location choice model in practice is the MNL model based on cross-sectional data with only observed sources of heterogeneity. MNL-1 is a more intelligent

[14] The number of recreational opportunities in a zone is computed as the sum of the number of cinemas, theaters, event halls, golf clubs, gyms, museums, restaurants, parks and outdoor recreational venues in the zone.

model since it is based on panel data and therefore incorporates the effects of choice occasion-specific constraints among the observed sources of heterogeneity. MNL-2 goes a step further and incorporates first order feedback. MxL-2 builds on MNL-2 and estimates unobserved inter-individual heterogeneity, and is the best estimated model. A comparison of the MNL-1, MNL-2 and MxL-2 models in the base case and under different policy scenarios follows.

The predicted fraction of non-maintenance shopping travel to the central districts (or CBD) forms the basis for the comparison of the models. The tests were conducted using only the second choice occasion of each individual, with the observed choice from the first occasion used as feedback. Each of the MNL-1, MNL-2 and MxL-2 models was applied to predict the chosen location for each individual in the base case and policy scenarios. A comparison of the observed choices in the sample used for estimation against the base case predictions for each model indicates that the model MNL-1 over predicts travel to the CBD zones by about 21%. The MNL-2 and MxL-2 models, on the other hand, are comparable and over predict travel to the CBD zones only by a little more than 15%. However, the MxL-2 model performs better in policy analyses than the MNL-2 model as indicated by the scenarios analyzed.

In both the following scenarios, the model predictions in the policy case are compared to the model predictions in the base case rather than the observed choice data. This ensures that the comparison of the models from a policy analysis perspective is not confounded by the differing accuracies of the models in the base case.

In the first scenario, the composite size of non-CBD zones is increased by 25%. This corresponds to a situation where non-CBD zones grow and there are increased shopping and recreational opportunities available in these zones. This potentially generates a large draw away from the CBD zones, and the MNL-1 model correspondingly predicts a 11% drop in the travel to CBD zones. The MNL-2 model, however, takes loyalty and inertial behavior into account and predicts only a 3% drop in the travel to CBD zones. The MxL-1 model, not only accounts for loyalty and inertia but also for the heterogeneity across individuals in loyalty/inertial behavior, and therefore predicts a 7% drop in the travel to CBD zones. Clearly, the MxL-2 model better accounts for the various aspects of choice behavior and a poorer model would result in significantly different policy analysis results.

The second policy scenario is one of increased access to public transport. This is simulated by incrementing the number of season tickets owned by each individual by one. Effectively, in the second policy case, 100% of the sample owns at least one season ticket (the corresponding fraction in the base case is 68%). The model results in Tables 12 and 13 suggest that the number of season tickets owned by an individual influences the individual's sensitivity to zonal composite size and distance. In the base case, individuals who live farther away from the CBD do not always choose a CBD zone since they trade-off composite size against distance. In the policy case, all the individuals have access to public transport and the trade-off between composite size and distance reduces. We would thus expect an increase in travel to the CBD. MNL-1 duly predicts a 2% increase in travel to the CBD. MNL-2, on the other hand, takes loyalty and inertial behavior into account and predicts only a 0.7% increase in travel to the CBD. MxL-2 takes into consideration the heterogeneity in response to composite size and distance and predicts that despite the loyalty/inertial factor unobserved heterogeneity effects would result in a 4% increase in travel to the CBD. The low percentages observed in this analysis are not surprising given the size of the study area. The small size of the study area implies that the tradeoff between composite size and distance will not influence as large a population as may be expected in a larger study area, where distances are a deterrent to a larger fraction of the population.

Many other policy scenarios, including aging of the population with a correspondingly larger retired community, and an increase in auto-ownership, and a reduction in trip chaining, were also tested with the estimated models, with similar results. The MxL-2 model incorporates accuracy and behavioral realism and represents individual choice behavior better than the MNL models, and is clearly better suited for policy analyses.

CHAPTER 8. CONCLUSION

In today's world of exploding travel demand and traffic delays and insufficient scope for infrastructural expansions, urban and transportation planners increasingly rely on the accuracy and behavioral realism of travel demand models to make informed and reliable policy decisions. Accuracy and behavioral realism in the travel demand models also helps establish their credibility outside the modeling community. The focus of this dissertation is to develop a comprehensive, unified, framework for spatial location choice that is both accurate and behaviorally realistic, and can be practically applied by planners and policy makers in the estimation of travel demand. The following sections summarize this dissertation research (section 8.1), and discuss future work and extensions to the research (Section 8.2).

8.1 Summary

The development of accurate and behaviorally realistic travel demand models requires a good understanding of individual travel behavior. An important step toward the better understanding of travel behavior has been the development of the activity-based paradigm, which states that travel is a result of the desire to participate in activities at spatially scattered locations. Activity based methods are thus more accurate and behaviorally realistic than the traditionally applied trip-based methods, and planning organizations are beginning to slowly adopt activity-based modeling systems over the traditional trip-based four-step planning process. Activity-based modeling systems essentially model the activity-travel patterns of individuals, which are characterized by several attributes such as activity purpose, location of activity participation and choice of mode.

Of all the attributes that characterize the activity-travel patterns of individuals, the choice of location of activity participation is one that has received relatively inadequate attention in the literature. On the other hand, the location of activity participation spatially pegs the daily activity-travel patterns of individuals. Accurate predictions of activity location are, therefore, key to effective travel demand management and air quality control strategies. Moreover, an understanding of the factors that influence the choice of location can contribute to more effective land-use and zoning policies.

The choice of location of activity participation and the factors that influence this choice vary with the activity purpose. While the work location for most people is generally fixed in the short term, non-work activity participation is typically characterized by a high degree of spatial-temporal flexibility and discretion. The first objective of this dissertation research was to develop a comprehensive econometric model of location choice for non-work activities that incorporates accuracy and behavioral realism in capturing different kinds of choice behaviors.

The development of an accurate and behaviorally realistic model of location choice for non-work activity participation necessitates a good understanding of the factors influencing the choice process. An extensive survey of the spatial choice modeling literature was conducted to achieve this aim. The key issues associated with understanding location choice behavior were identified, and a comprehensive list of observed and unobserved factors that influence location choice decisions was prepared.

Subsequently, a conceptual framework of location choice decision-making for non-work activity participation was developed that incorporates all the observed and unobserved factors that potentially influence the decision-maker. Finally, the proposed conceptual framework was translated into a general econometric model of location choice for non-work activity participation. The model structure thus developed is comprehensive in its incorporation of the different sources of heterogeneity such as spatial cognition, preference behavior and spatial interaction.

The incorporation of behaviorally realistic concepts, such as spatial cognition and spatial interaction, in the proposed econometric model of location choice is achieved through the relaxation of restrictions that impose inappropriate behavioral assumptions regarding the underlying choice process. This relaxation of behavioral restrictions on the choice model structure leads to analytically intractable choice probability expressions, which necessitate the use of numerical integration techniques such as pseudo-Monte Carlo (PMC) and quasi-Monte Carlo (QMC) simulation techniques to evaluate the multidimensional integrals in the probability expressions.

Simulation techniques have evolved over the years, and the use of QMC sequences for simulation is slowly beginning to replace PMC methods, as the efficiency and faster convergence rates of the low-discrepancy QMC sequences makes them more desirable. There have been several studies comparing the performance of different QMC sequences in the evaluation of a single multidimensional integral. The use of QMC sequences in the simulated maximum likelihood estimation of flexible discrete choice models, which entails the estimation of parameters by the approximation of several multidimensional integrals at each iteration of the optimization procedure, is, however, relatively recent. The second objective of this dissertation research was to experimentally compare the overall performance of the Halton and Faure sequences against each other and against the Latin Hypercube Sampling (LHS) sequence in the context of the simulated likelihood estimation of a mixed logit choice model. The different scrambled versions of QMC sequences were also compared, and the effect of scrambling on the accuracy and efficiency of these sequences was examined. In addition, the efficiency of the QMC sequences generated with and without scrambling across observations was compared. The results of this analysis indicate that the Faure sequence consistently outperforms the Halton sequence. The Random Linear and Random Digit scrambled Faure sequences, in particular, are amongst the most effective QMC sequences for simulated maximum likelihood estimation of the mixed logit model.

This dissertation, therefore, not only proposes a comprehensive econometric model of non-work location choice, but also proposes techniques to improve the efficiency of the simulated maximum likelihood estimation procedure. In combination, this research makes the estimation of accurate and behaviorally realistic non-work location choice models practically feasible, which is demonstrated in the empirical application presented here.

As part of the empirical application, location choice models for non-maintenance shopping were estimated using the Mobidrive data, based on the proposed model structure and applying the most efficient QMC sequence identified for estimation. Although the Mobidrive data is the richest multi-day data source currently available for this purpose, the empirical analysis is limited by the zonal configuration of the study area and the quantity of the data. The core city of Karlsruhe is a fairly small region of area approximately 15.6 sq.km, with a mature public transportation system and tight land-use control. It is therefore conceivable that the zonal configuration creates clear boundaries between different land-use parcels. Moreover, even the rich Mobidrive data only provides a sample of 158 individuals with 745 non-maintenance shopping occasions in all. Consequently, the estimated location choice models predict the absence of spatial correlation and are unable to capture the effects of unobserved intra-individual heterogeneity. Despite these limitations, the best estimated model in this empirical application is superior to commonly applied MNL models of location choice since it is based on repeated choice observations (panel data) and incorporates feedback as well as unobserved inter-individual heterogeneity. A comparison of the estimated models in the base case and different policy scenarios further proves the importance of incorporating accuracy and behavioral realism in travel demand models.

In conclusion, the proposed model structure is comprehensive in its incorporation of unobserved inter- and intra-individual heterogeneity, spatial correlation and feedback effects. In order to exploit the full potential of the proposed model structure, however, it is necessary to use a rich multi-day data source that satisfies the criteria presented in section 7.1. The models thus estimated would present an accurate representation of individual choice behavior, and provide reliable policy implications.

8.2 Extensions and Future Work

Future work and potential extensions to this dissertation research can be directed along two different, and yet related, dimensions. On the one hand, further research can be undertaken on the non-work location choice model toward a better understanding of travel behavior, and on the other hand there are many more avenues to be explored in further improving the efficiency of the simulated maximum likelihood estimation (SMLE) procedure. As travel demand models become more behaviorally realistic, the burden of estimation becomes a very real problem and a more efficient optimization process may be the best solution. A few research ideas along these lines are presented here.

8.1.1 Multi-day Data Collection

The non-work location choice theory and model structure developed in this dissertation is rich and comprehensively incorporates different types of behavior, spatial cognition and learning, various sources of heterogeneity and spatial interaction. However, in order to exploit this model structure to the fullest extent, a rich multi-day data source with a reasonably large sample of individuals is necessary. An examination of the available data sources indicates the scarcity of this type of data. Further, all existing multi-day data sources are European in origin. Research toward the efficient and successful collection of multi-day data would therefore be a significant contribution to behavioral research.

8.2.2 Simultaneous Variety-Seeking and Location Choice Model

The variety-seeking ratio proposed in section 7.3 is intuitive in its ability to capture the degree of heterogeneity in the choice of activity location. Further, a regression of the variety-seeking ratio for non-maintenance shopping against individual socio-economic attributes effectively captures the correlation between individual characteristics and the degree of heterogeneity in spatial choice (see Table 10). It is, therefore, a reasonable assumption that a combined discrete-continuous model of location choice and variety-seeking will lead to a better understanding of the underlying choice behavior. Such a simultaneous model would assume a correlation between the unobserved sources of heterogeneity in the variety-seeking ratio and the choice of location of activity participation, which is very intuitive.

8.2.3 Effects of Trip Chaining

Discretionary (or non-work) activities are typically chained together, and therefore the travel-related decisions for one type of non-work activity can be expected to influence the other non-work activity-travel patterns. Although the location choice model structure proposed in this dissertation is flexible enough to accommodate the effects of trip chaining, trip chaining effects are not examined in detail. This then provides another interesting avenue for further research. Specifically, extended empirical applications with model specifications geared toward including the effects of trip chaining would be of interest.

8.2.4 Flexible Destination Configurations

Transportation analysis zones constitute the choice set in the empirical application presented here. In other words, the implicit assumption is that individuals evaluate zones comprising several non-maintenance shopping options in selecting a destination for the activity participation. The choice of zones in such models has been the subject of debate for many years now (see, for example, Guo and Bhat, 2004). Instead of the transportation analysis zones, it may be more intuitive to consider clusters of shopping opportunities that simulate the choice alternatives faced by the individuals more realistically. Such an empirical analysis would still be based on the comprehensive model structure developed in this research, since the proposed model structure is not limited by the configuration of the choice alternatives.

8.2.5 Extended Comparison of QMC Sequences

The comparison of Quasi-Monte Carlo sequences and scrambling methods presented in this dissertation could be extended further to include the Sobol and Niederreiter sequences, and the optimal scrambling of Halton sequences proposed by Mascagni and Chi (2004). Both the Sobol and Niederreiter sequences have been demonstrated to perform better than the Halton and Faure sequences in the estimation of a single multidimensional integral. It would, therefore, be interesting to examine their performance in the simulated maximum likelihood estimation of a mixed logit model. However, both the Sobol and Niederreiter sequences are somewhat limited from a practical perspective since they must be generated specific to each application context. Since the Halton and Faure sequences are not limited from this perspective, it is useful to examine other scrambling methods for both these sequences in the hope of further efficiency gains. It would also be of interest to examine the effects of other uniform-to-normal transformation techniques, such as Moro's method or Ramberg and Schmeiser approximation.

8.2.6 Joint Comparison of Optimization Techniques & QMC Sequences

Another approach to improving the efficiency of the simulated maximum likelihood estimation (SMLE) procedure is to examine different optimization algorithms for maximizing the simulated maximum likelihood function. One such algorithm proposed in the literature is a trust region-based method that dynamically adapts the number of draws for the computation of the simulated maximum likelihood function on the basis of statistical estimators of the simulation error and simulation bias (Bastin et al., 2006). A research area of further interest is the comparison of different optimization techniques in combination with different QMC sequences on the efficiency of the SMLE procedure.

BIBLIOGRAPHY

Aaker, D.A. and J.M. Jones (1971) Modeling Store Choice Behavior, Journal of Marketing Research, Vol. 8, pp. 38-42.

Aldskogius, H. (1977) A Conceptual Framework and a Swedish Case Study of Recreational Behaviour and Environmental Cognition, *Economic Geography*, Vol. 53(2), pp. 163-183.

Anderson, N. H. (1976) How Functional Measurement can Yield Validated Interval Scales of Mental Qualities, *Journal of Applied Psychology*, Vol. 6, pp. 677-693.

Anooshian, L. J., and P. S. Seibert (1996) Diversity within Spatial Cognition: Memory Processes Underlying Place Recognition, *Applied Cognitive Psychology*, Vol. 10, pp. 281-299.

Axhausen, K.W., Z. Zimmerman, S. Schönfelder, G. Rindsfuser, and T. Haupt (2000) Observing the Rhythms of Daily Life: A Six Week Travel Diary, Paper submitted to Transportation.

Bastin, F., Cirillo, C. and P. Tointe (2006) An Adaptive Monte-Carlo Algorithm for Computing Mixed Logit Estimators, Forthcoming in Computational Management Science, Volume 3(1).

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

Ben-Akiva, M. and J. Bowman (1998) Integration of an Activity-based Model System and a Residential Location Model, *Urban Studies*, Vol. 35(7), pp. 1231-1253.

Ben-Akiva, M. and T. Morikawa (1997) Methods to Estimate Discrete Choice Models with Stated and Revealed Preferences, Presented at 1997 NSF Symposium on Eliciting Preferences, Berkeley, California, July 1997.

Berman, O., N. Fouska and R. C. Larson (1992) Optimal Locations for Discretionary Service Facilities, *Transportation Science*, Vol. 26(3), pp. 201-211.

Beynon, M., Griffiths, B. and D. Marshall (2002) A Prototype Store Choice and Location Modelling System using Dempster-Shafer Theory, *Expert Systems*, Vol. 19, No. 5, pp. 273-284.

Bhat, C. R. (1998) Accommodating Flexible Substitution Patterns in Multi-dimensional Choice Modeling: Formulation and Application to Travel Mode and Departure Time Choice, *Transportation Research*, Vol. 32B, No. 9, pp. 455-466.

Bhat, C.R. (1999) An Analysis of Evening Commute Stop-Making Behavior Using Repeated Choice Observations from a Multi-Day Survey, *Transportation Research Part B*, Vol. 33, No. 7, pp. 495-510.

Bhat, C.R. (2000) A Multi-Level Cross-Classified Model for Discrete Response Variables, *Transportation Research*, Vol. 34B, pp. 567-582.

Bhat, C.R. (2001) Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model, *Transportation Research*, 35B, 677-693.

Bhat, C.R. (2002) Recent Methodological Advances Relevant to Activity and Travel Behavior Analysis, In *Perpetual Motion: Travel Behavior Research Opportunities and Application Challenges*, pp. 381-414, Ed. H.S. Mahmassani, Pergamon, Elsevier Science.

Bhat, C.R. (2003) Simulation Estimation of Mixed Discrete Choice Models using Randomized and Scrambled Halton Sequences, *Transportation Research*, 37B, pp. 837-855.

Bhat, C. R., and S. K. Singh (2000) A Comprehensive Daily Activity Travel Generation Model System for Workers, *Transportation Research*, Vol. 34A, No. 1, pp. 1-22.

Bhat, C.R. and R. Misra (2002) Comprehensive Activity-Travel Pattern Modeling System for Non-Workers with Empirical Focus on the Organization of Activity Episodes, *Transportation Research Record*, Vol. 1777, pp. 16-24.

Bhat, C.R., and S. Castelar (2002) A Unified Mixed Logit Framework for Modeling Revealed and Stated Preferences: Formulation and Application to Congestion Pricing Analysis in the San Francisco Bay Area, *Transportation Research Part B*, Vol. 36, No. 7, pp. 593-616.

Bhat, C.R. and R. Gossen (2004) A Mixed Multinomial Logit Model Analysis of Weekend Recreational Episode Type Choice, Forthcoming *Transportation Research*.

Bhat, C.R., and J.Y. Guo, (2004) A Mixed Spatially Correlated Logit Model: Formulation and Application to Residential Choice Modeling, *Transportation Research Part B*, Vol. 38, No. 2, pp. 147-168.

Bhat, C.R. and S. Srinivasan (2004) A Multidimensional Mixed Ordered-Response Model for Analyzing Weekend Activity Participation, Forthcoming *Transportation Research*.

Bhat, C.R., J. Guo, S. Srinivasan, and A. Sivakumar (2003) Guidebook on Activity-Based Travel Demand Modeling for Planners, Product 4080-P3, prepared for the Texas Department of Transportation.

Birkin, M. and G. P. Clarke (1991) Spatial Interaction in Geography, *Geography Review*, Vol. 4, pp. 16-24.

Borgers, A. and H. Timmermans (1987) Choice Model Specification, Substitution and Spatial Structure Effects, *Regional Science and Urban Economics*, Vol. 17, pp. 29-47.

Bowman, J.L., and M.E. Ben-Akiva (2000) Activity-based Disaggregate Travel Demand Model System with Activity Schedules, *Transportation Research*, Vol. 35A, pp. 1-28.

Box, G.E.P. and M. E. Muller (1958) A Note on the Generation of Random Normal Deviates, *Annals of Math. Stat*., Vol. 29(2), pp. 610–611.

Braaten, E. and G. Weller (1979) An Improved Low-Discrepancy Sequence for Multidimensional Quasi-Monte Carlo Integration, *Journal of Computational Physics*, Vol. 33, pp. 249-258.

Brandeau, M.L. and S.S. Chiu (1994) Location of Competing Facilities in a User-Optimizing Environment with Market Externalities, *Transportation Science*, Vol. 28, No. 2, pp. 125-140.

Bratley, P. and B.L. Fox (1988) Implementing Sobol's Quasi-random Sequence Generator, *ACM Transactions on Mathematical Software*, Vol. 14, pp. 88-100.

Bolduc, D., Fortin, B. and M. Fournier (1996) The Effect of Incentive Policies on the Practice Location of Doctors: A Mutinomial Probit Analysis, *Journal of Labor Economics*, Vol. 14, Issue 4, pp. 703-732.

Brownstone, D. and K. Train (1999) Forecasting New Product Penetration with Flexible Substitution Patterns, *Journal of Econometrics*, Vol. 89, pp. 109-129.

Burnett, P. (1973) The Dimensions of Alternatives in Spatial Choice Processes, *Geographical Analysis*, Vol. 5, pp. 181–204.

Burnett, P. (1976) Behavioral Geography and the Philosophy of the Mind, In *Spatial Choice and Spatial Behavior*, pp. 23-48, Eds. R. Golledge and G. Rushton, Ohio State University Press, Columbus, Ohio.

Burnett, P. (1977) Tests for the Linear Learning Model of Destinations: Application to Shopping Travel by Heterogeneous Population Groups, *Geographical Analysis*, Series B, Vol. 59, No. 2, pp. 95-108.

Burnett, P. (1978) Markovian Models of Movement Within Urban Spatial Structures, *Geographical Analysis*, Vol. 10, No. 2, pp. 142-153.

Cadwallader, M. (1975) A Behavioral Model of Consumer Spatial Decision Making, *Economic Geography*, Vol. 51, pp. 339–49.

Cadwallader, M. (1995) Interaction Effects in Models of Consumer Spatial Behavior, *Applied Geography*, Vol. 15, No. 2, pp. 135-145.

Carey, H.C. (1859) Principles of Social Science, Philadelphia: Lippincott.

Cesario, F. J. (1973) A Note on the Entropy Model of Trip Distribution, *Transportation Research*, Vol. 27, pp. 331-333.

Daskin, M. S., Snyder, L. V., and R. T. Berter (2003) Facility Location in Supply Chain Design, Working Paper No. 03-010, Department of Industrial Engineering and Management Sciences, Northwestern University, Illinois.

Dellaert, B.G.C., Arentze, T.A., Bierlaire, M., Borgers, A.W.J. and H.J.P. Timmermans (1998) Investigating Consumers' Tendency to Combine Multiple Shopping Purposes and Destinations, *Journal of Marketing Research*, Vol. 35, pp. 177-188.

Drezner, Z., Mehrez A. and G. O. Wesolowsky (1991) The Facility Location Problem with Limited Distances, *Transportation Science*, Vol. 25, No. 3, pp. 183-187.

Dunn, R. and N. Wrigley (1985) Beta-logistic Models of Urban Shopping Center Choice, *Geographical Analysis*, Vol. 17, No. 2, pp. 95-113.

Eymann, A. and G. Ronning (1997) Microeconometric Models of Tourists' Destination Choice, *Regional Science and Urban Economics*, Vol. 27, pp. 735-761.

Faure, H. (1982) Discrépance de Suites Associées à un Système de Numeration (En Dimension s), *Acta Arith*, Vol. 41, pp. 337-351.

Feather, P.M. (1994) Sampling and Aggregation Issues in Random Utility Model Estimation, *American Journal of Agricultural Economics*, Vol. 76, pp. 772-780.

Ferguson, M. R. and P. S. Kanaroglou (1998) Representing the Shape and Orientation of Destinations in Spatial Choice Models, *Geographical Analysis*, Vol. 30, No. 2, pp. 119-137.

Fischer, M. M. and M. Reisman (2002) A Methodology for Neural Spatial Interaction Modeling, *Geographical Analysis*, Vol. 34, No. 3, pp. 207-228.

Fisk, C. And G. R. Brown (1975) A Note on the Entropy Formulation of Distribution Models, *Operations Research Quarterly*, Vol. 26, pp. 755-758.

Fotheringham, A. S. (1983) A New Set of Spatial Interaction Models: The Theory of Competing Destinations, *Environment and Planning A*, Vol. 15, pp. 15-36.

Fotheringham, A. S. (1988) Consumer Store Choice and Choice Set Definition, *Marketing Science*, Vol. 7, pp. 299-310.

Fotheringham, A. S. (1991) Statistical Modeling of Spatial Choice: An Overview. In *Spatial Analysis in Marketing: Theory, Methods and Applications*, pp. 95-118, Eds. A. Ghosh and C. Ingene, Greenwich, CT: JAI Press.

Fotheringham, A. S. and A. Curtis (1999) Regularities in Spatial Information Processing: Implications for Modelling Destination Choice, *The Professional Geographer*, Vol. 51, pp. 227-239.

Fotheringham, A.S., Nakaya, T., Yano, K., Openshaw, S., and Y. Ishikawa (2001) Hierarchical Destination Choice and Spatial Interaction Modelling: A Simulation Experiment, *Environment and Planning A*, Vol. 33, pp. 901-920.

Fox, B.L. (1986) Implementation and Relative Efficiency of Quasi-random Sequence Generators, *ACM Transactions on Mathematical Software*, Vol. 12, pp. 362-376.

Gannon, C.A. (1972) Intra-Urban Industrial Location Theory and Inter-Establishment Linkages, Presented at the Second Annual Meeting of the North-East Section of the Regional Science Association, Pennsylvania State University, University Park, Pennsylvania.

Gärling, T., Böök, A. and E.Lindberg (1984) Cognitive Mapping of Large Scale Environments: The Interrelationship of Action Plan, Acquisition and Orientation, *Environment and Behavior*, Vol. 16, pp. 3-34.

Golledge, R. G. And R. Briggs (1973) Decision Processes and Locational Behavior, *High Speed Ground Transportation*, Vol. 7, pp. 81-100.

Golledge, R.G. and A.N. Spector (1978) Comprehending the Urban Environment: Theory and Practice, *Geographical Analysis*, Vol. 10, No. 4, pp. 403-426.

Golledge, R. G. and G. Zannaras (1973) Cognitive Approaches to the Analysis of Human Spatial Behavior, In *Environment and Cognition*, pp. 59-94, edited by W.H. Ittelson, New York: Seminar Press.

González-Benito, Ó. (2002) Overcoming Data Limitations for Store Choice Modelling. Exploiting Retail Chain Choice Data by Means of Aggregate Logit Models, *Journal and Retailing and Consumer Services*, Vol. 9, pp. 259-268.

Goodchild, M.F. (1978) Spatial Choice in Location-Allocation Problems: The Role of Endogenous Attraction, *Geographical Analysis*, Vol. 10, No. 1, pp. 65-72.

Gould, P. and R. White (1974) Mental Maps, Harmondsworth, UK: Penguin.

Goulias, K.G. (2002) Multilevel Analysis of Daily Time Use and Time Allocation to Activity Types Accounting for Complex Covbariance Structures using Correlated Random Effects, *Transportation*, Vol. 29, pp. 31-48.

Guo, J. Y. (2004) Addressing Spatial Complexities in Residential Location Choice Models, Ph.D. Dissertation, University of Texas at Austin.

Guo, J.Y., and C.R. Bhat (2004) Modifiable Areal Units: Problem or Perception in Modeling of Residential Location Choice?, *Transportation Research Record*, Vol. 1898, pp. 138-147.

Hajivassiliou, V.A. and P.A. Ruud (1994) Classical Estimation Methods for LDV Models using Simulation. In *Handbook of Econometrics, IV*, pp. 2383-2441, Eds. Engle, R.F. and D.L. McFadden, Elsevier, New York.

Hajivassiliou, V.A., McFadden, D.L. and P.A. Ruud (1996) Simulation of Multivariate Normal Rectangle Probabilities and their Derivatives: Theoretical and Computational Results, *Journal of Econometrics*, Vol. 72, pp. 85-134.

Halperin, W., Gale, N., Golledge, R. and L. Hubert (1983) Exploring Entrpreneurial Cognitions of Retail Environments, *Economic Geography*, Vol. 59, pp. 3-15.

Halton, J.H. (1960) On the Efficiency of Certain Quasi-random Sequences of Points in Evaluating Multidimensional Integrals, *Numer. Math*., Vol. 2, pp. 84-90.

Hamed, M. M., and F. L. Mannering (1993) Modeling Travelers' Postwork Activity Involvement: Toward a New Methodology, *Transportation Science*, Vol. 27, No. 4, pp. 381-394.

Handy, S. L. (1992) Regional Versus Local Accessibility: Neo-Traditional Development and its Implications for Non-Work Travel, *Built Environment*, Vol.18, No. 4, pp. 253-267.

Hansen, E. R. (1987) Industrial Location Choice in São Paulo, Brazil – A Nested Multinomial Logit Model, *Regional Science and Urban Economics*, Vol. 17, pp. 89-108.

Hanson, S. and P. Hanson (1981) The Travel Activity Patterns of Urban Residents: Dimensions and Relationships to Socio-demographic Characteristics, *Economic Geography*, Vol. 57, pp. 332-347.

Hanson, S. and J. Huff (1988) Systematic Variability in Repetitious Travel, *Transportation*, Vol. 15, pp. 111-135.

Harris, B. and A. G. Wilson (1978) Equilibrium Values and Dynamics of Attractiveness Terms in Production-Constrained Spatial-Interaction Models, *Environment and Planning A*, Vol. 10, pp. 371-388.

Haynes, K. E. and A. S. Fotheringham (1984) *Scientific Geography Series 2. Gravity and Special Interaction Models*, Sage, Beverly Hills, CA.

Heckshер E F (1919) The Effect of Foreign Trade on the Distribution of Income, *Ekonomisk Tidskrift*, Vol 7, No. 21, pp. 497-579.

Hensher, D. A. (2001) Measurement of the Valuation of Travel Time Savings, *Journal of Transport Economics and Policy*, Vol. 35, No. 1, pp. 71-98.

Horowitz, J. (1980) A Utility Maximizing Model of the Demand for Multi-destination Non-work Travel, *Transportation Research*, Vol. 14B, pp. 369-386.

Horton, F. E. and D. R. Reynolds (1969) An Investigation of Individual Activity Spaces: A Progress Report, *Proceedings of the Association of American Geographers*, Vol. 1, pp. 70-75.

Hsu, C. and Y. Hsieh (2004) Travel and Activity Choices Based on an Individual Accessibility Model, *Papers in Regional Science*, Vol. 83, pp. 387-406.

Huff, D. L. (1963) A Probabilistic Analysis of Consumer Spatial Behavior, In *Emerging Concepts in Marketing,* pp. 443-461, Ed. W.S. Decker, Chicago, American Marketing Association.

Hunt, J. D. (1996) A Stated Preference Examination of Time of Travel for a Recreational Trip, *Journal of Advanced Transportation*.

Isard, W. (1960) Methods of Regional Analysis, MIT Press, Cambridge.

Johnson, E. and J. Payne (1985) Effort and Accuracy in Choice, *Management Science*, Vol. 31, pp. 395-414.

Jones, P. M. (1977) New Approaches to Understanding Traveller Behavior: The Human Activity Approach, Resource Paper, Proceedings of the Third International Conference on Behavioral Travel Modelling, Tanunda, South Australia, Australia, Pergamon Press.

Kanaroglou, P.S. and M.R. Ferguson (1996) Discrete Spatial Choice Models for Aggregate Destinations, *Journal of Regional Science*, Vol. 36, No. 2, pp. 271-290.

Kemperman, A., Borgers, A., Oppewal, H. And H. Timmermans (2000) Consumer Choice of Theme Parks: A Conjoint Choice Model of Seasonality Effects and Variety-Seeking Behavior, *Leisure Sciences*, Vol. 22, pp. 1-18.

Kemperman, A., Borgers, A. and H. Timmermans (2002) Incorporating Variety-Seeking and Seasonality in Stated Preference Modeling of Leisure Trip Destination Choice: A Test of External Validity, Presented at the 81st Annual Meeting of the Transportation Research Board, Washington, D.C., January 2002.

Kemperman, A., Ponjé, M. And H. Timmermans (2004) Analyzing Heterogeneity and Substitution in Trip Making Propensity to Urban Parks: A Mixed Logit Model, Presented at the 83rd Annual Meeting of the Transportation Research Board, Washington, D.C., January 2004.

Kitamura, R. (1984) Incorporating Trip Chaining into Analysis of Destination Choice, *Transportation Research B*, Vol. 18, pp. 67-82.

Kitamura, R., Kostyniuk, L.P. and K. Ting (1979) Aggregation in Spatial Choice Modeling, *Transportation Science*, Vol. 13, No. 4, pp. 325-342.

Kocis, L. and W.J. Whiten (1997) Computational Investigations of Low-Discrepancy Sequences, *ACM Transactions on Mathematical Software*, Vol. 23(2), pp. 266-294.

Laporte, G., Louveaux, F.V. and L. van Hamme (1994) Exact Solution to a Location Problem with Stochastic Demands, *Transportation Science*, Vol. 28, No. 2, pp. 95-103.

Lee, L-F. (1992) On Efficiency of Methods of Simulated Moments and Maximum Simulated Likelihood Estimation of Discrete Choice Models, *Econometric Theory*, Vol. 8, pp. 518-552.

Leonardi, G. and Y. Y. Papageorgiou (1992) Conceptual Foundations of Spatial Choice Models, *Environment and Planning A*, Vol. 24, pp. 1393-1408.

Lerman, S. R. (1983) Random Utility Models of Spatial Choice, In *Optimization and Discrete Choice in Urban Systems*, Proceedins of the International Symposium on New Directions in Urban Systems Modeling at the University of Waterloo, Canada, pp. 200-217, Eds. Hutchinson, B. G., Nijkamp, P. and M. Batty.

Leszczyc, P. and H. Timmermans (1996) An Unconditional Competing Risk Hazard Model of Consumer Store-Choice Dynamics, *Environment and Planning A*, Vol. 28, 357-368.

Levin, I. P. and J. J. Louviere (1979) Functional Measurement Analysis of Spatial and Travel Behavior, In *Advances in Consumer Research*, Association for Consumer Research.

Long, W. and R Uris (1971) Distance, Intervening Opportunities, City Hierarchy, and Air Travel, *Annals of Regional Science*, Vol. 5, pp. 152-161.

Louviere, J. J. (1984) Using Discrete Choice Experiments and Multinomial Logit Models to Forecast Trial in a Competitive Retail Environment: A Fast Food Restaurant Illustration, *Journal of Retailing*, Vol. 60, pp. 81-107.

Louviere, J. J. and G. G. Woodworth (1983) Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data, *Journal of Marketing Research*, Vol. 20, pp. 350-367.

MacKay, D.B. (1973) Measuring Shopping Patterns, *Geographical Analysis*, Vol. 5, pp. 329-337.

Mascagni, M. and H. Chi (2004) On the Scrambled Halton Sequence, *Monte Carlo Methods and Applications*, Vol. 10. No. 3-4, pp. 435-442.

Matoušek, J. (1998) On the L2-Discrepancy for Anchored Boxes, *Journal of Complexity*, Vol. 14, pp. 527-556.

McFadden, D. (1978) Modelling the Choice of Residential Location, In *Spatial Interaction Theory and Residential Location*, pp. 75-96, Ed. A. Korlquist, North Holland, Amsterdam.

McFadden, D. and K. Train (2000) Mixed MNL Models for Discrete Response, *Journal of Applied Econometrics*, Vol. 15(5), pp. 447-470.

McKay, M.D., Conover, W.J. and R.J Beckman (1979) A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code, *Technometrics*, Vol. 21, pp. 239-245.

Meyer, R. (1979) Thgeory of Destination Choice-Set Formation Under Informational Constraint, *Transportation Research Record*, Vol. 750, pp. 6-12.

Miller, E. J. and M. E. O'Kelly (1983) Estimating Shopping Destination Choice Models from Travel Diary Data, *Professional Geographer*, Vol. 35(4), pp. 440-449.

Miyamoto, K., Vichiensan, V., Shimomura, N., and A. Paez (2004) Discrete Choice Model with Structuralized Spatial Effects for Location Analysis, Presented at the 83rd Conference of the Transportation Research Board, Washington, D.C.

Morokoff, W.J. and R.E. Caflisch (1994) Quasi-Random Sequences and their Discrepancies, *SIAM Journal of Scientific Computation*, Vo1. 5(6), pp. 1251-1279.

Morokoff, W.J. and R.E. Caflisch (1995) Quasi-Monte Carlo Integration, *Journal of Computational Physics*, Vol. 122, pp. 218-230.

Morris, J.G. and J. P. Norback (1980) A Simple Approach to Linear Facility Location, *Transportation Science*, Vol. 14, pp. 1-8.

Odland, J. (1981) A Household Production Approach to Destination Choice, *Economic Geography*, Vol. 57, No. 3, pp. 257-269.

Ohlin B (1933) International and Interregional Trade, Harvard University Press, Cambridge, MA.

Ökten, G. and W. Eastman (2004) Randomized Quasi-Monte Carlo Methods in Pricing Securities, *Journal of Economic Dynamics & Control*, Vol. 28(12), pp. 2399-2426.

Onaka, J. and W. Clark (1983) A Disaggregate Model of Residential Mobility and Housing Choice, *Geographical Analysis*, Vol. 15, pp. 287-304.

Oppewal, H., Louviere, J. and H. Timmermans (1994) Modeling Hierarchical Conjoint Processes with Integrated Choice Experiments, *Journal of Marketing Research*, Vol. 31, pp. 92-105.

Orpana, T. and J. Lampinen (2003) Building Spatial Choice Models from Aggregate Data, *Journal of Regional Science*, Vol. 43, No. 2, pp. 319-347.

Owen, A.B. (1997) Scrambled Net Variance for Integrals of Smooth Functions, *The Annals of Statistics*, Vol. 25, pp. 1541-1562.

Owen, A.B. (1998) Latin Super Cube Sampling for Very High Dimensional Simulations, *ACM Transactions on Modeling and Computer Simulation*, Vol. 8, pp. 71-102.

Park, Y.H., Rhee, S.B. and Bradlow, E.T. (2003) An Integrated Model for Who, When, and How Much in Internet Auctions, Working Paper, Department of Marketing, Wharton.

Parsons, G. R. and A.B. Hauber (1998) Spatial Boundaries and Choice Set Definition in a Random Utility Model of Recreation Demand, *Land Economics*, Vol. 74, Issue 1, pp. 32-48.

Pellegrini, P. A. and A. S. Fotheringham (1999) Inter-metropolitan Migration and Hierarchical Destination Choice: A Disaggregate Analysis from the US PUMS, *Environment and Planning A*, Vol. 31, pp. 1093-1118.

Pellegrini, P. A. and A. S. Fotheringham (2002) Modelling Spatial Chioce: A Review and Synthesis in a Migration Context, *Progress in Human Geography*, Vol. 26(4), pp. 487-510.

Pendyala, R.M., T. Yamamoto, and R. Kitamura (2002) On the Formulation of Time-Space Prisms To Model Constraints on Personal Activity-Travel Engagement, *Transportation*, Vol. 29(1), pp. 73–94.

Perl, J. and P. Ho (1990) Public Facilities Location under Elastic Demand, *Transportation Science*, Vol. 24, No. 2, pp. 117-136.

Pipkin, J.S. (1979) Problems in the Psychological Modelling of Revealed Destination Choice, In *Philosophy in Geography*, pp. 309-328, Eds. Gale, S. and G. Olsson, Reidel Publishing Company, Dordrecht, Holland.

Pipkin, J.S. (1981) The Concept of Choice and Cognitive Explanations of Spatial Behavior, *Economic Geography*, Vol. 57, No. 4, Studies in Choice, Constraints, and Human Spatial Behaviors, pp. 315-331.

Pozsgay, M.A., and C.R. Bhat (2002) Destination Choice Modeling for Home-Based Recreational Trips: Analysis and Implications for Land-Use, Transportation, and Air Quality Planning, *Transportation Research Record*, Vol. 1777, pp. 47-54.

Press, W.H., Teukolsky, S.A. and M. Nerlove (1992) Numerical Recipes in C: The Art of Scientific Computing. Cambridge University Press, Massachusetts.

Ramberg, J.S. and B.W. Schmeiser (1972) An Approximate Method for Generating Symmetric Random Variables. *Communications of the ACM*, Vol. 5, pp. 987-990.

Ravenstein, E.G. 1885. The Laws of Migration (1), *Journal of the Royal Statistical Society*, Vol. 48, pp. 167-235.

Recker, W. and H. Schuler (1981) Destination Choice and Processing Spatial Information: Some Empirical Tests with Alternative Constructs, *Economic Geography*, Vol. 57, pp. 373-383.

Revelt, D. and K. Train (2000) Customer-Specific Taste Parameters and Mixed Logit: Household’s Choice of Electricity Supplier, Economics Working Papers E00-274, Department of Economics, University of California, Berkeley.

Roy, J. R. (1981) A Note on the Consistency of Choice Units in Entropy Models, *Transportation Research Part B*, Vol. 15, pp. 159-164.

Roy, J. R. And P. F. Lesse (1981) On Appropriate Microstate Descriptions in Entropy Modelling, *Transportation Research Part B*, Vol. 15, pp. 85-96.

Rust, R. T. and N. Donthu (1995) Capturing Geographical Localized Misspecification Error in Retail Store Choice Models, *Journal of Marketing Research*, Vol. 32, pp. 103-110.

Sandor, Z. and K. Train (2004) Quasi-random Simulation of Discrete Choice Models, *Transportation Research*, Vol. 38B, pp. 313-327.

Sarkar, P.K. and M.A. Prasad (1987) A Comparative Study of Pseudo and Quasi Random Sequences for the Solution of Integral Equations, *Journal of Computational Physics*, Vol. 68, pp. 66-88.

Schneider, M. (1959) Gravity Models and Trip Distribution Theory, *Regional Science Association, Papers and Proceedings*, Vol. 5, pp. 51-56.

Sermons, M. W. and F. S. Koppelman (1998) Factor Analytic Approach to Incorporating Systematic Taste Variation into Models of Residential Location Choice, *Transportation Research Record*, Vol. 1617, pp. 194-202.

Sermons, M. W. and F. S. Koppelman (2001) Representing the Differences Between Female and Male Commute Behavior in Residential Location Choice Models, *Journal of Transport Geography*, Vol. 9, pp. 101-110.

Shaw, J.E.H. (1988) A Quasirandom Approach to Integration in Bayesian Statistics, *The Annals of Statistics*, Vol. 16(2), pp. 895-914.

Sheppard, E.S. (1979) Spatial Interaction and Geographic Theory, In Philosophy in Geography, pp. 361-378, Eds. Gale, S. and G. Olsson, Reidel Publishing Company, Dordrecht, Holland.

Shönfelder, S. and K.W. Axhausen (2004) Structure and Innovation of Human Activity Spaces, Full length manuscript for the International Symposium on Transportation and Traffic Theory, ISTTT16, October 2004.

Shukla, V., and P. Waddell (1991) Firm Location and Land Use in Discrete Urban Space: A Study of the Spatial Structure of Dallas-Fort Worth, *Regional Science and Urban Economics*, Vol. 21, No. 2, pp. 225-51.

Slater, P.B. (1992) Equilibrium and Nonequilibrium Statistical Thermodynamical Approaches to Modeling Spatial Interaction Dynamics, *Environment and Planning A*, Vol. 24, pp. 441-446.

Smith, T.R. (1978) Uncertainty, Diversification, and Mental Maps in Spatial Choice Problems, *Geographical Analysis*, Vol. 10, No. 2, pp. 120-141.

Sobol, I.M. (1967) The Distribution of Points in a Cube and the Approximate Evaluation of Integrals, *USSR Computational Math. And Math. Phys*., Vol. 7, pp. 86-112.

Stewart, J.Q. (1941) An Inverse Distance Variation for Certain Social Influences, *Science*, Vol. 93, pp. 89-90

Stouffer, S. A. (1940) Intervening Opportunities: A Theory Relating Mobility and Distance, *American Sociological Review*, Vol. 5, pp. 845-867.

Tan, K.S. and P.P Boyle (2000) Applications of Randomized Low Discrepancy Sequences to the Valuation of Complex Securities, *Journal of Economic Dynamics & Control*, Vol. 24, pp. 1747-1782.

Tezuka, S. (1995) Uniform Random Numbers: Theory and Practice, Kluwer Academic Publishers, Dordrecht.

Thill, J. and A. Wheeler (2000) Tree Induction of Spatial Choice Behavior, *Transportation Research Record*, Vol. 1719, pp. 250-257.

Timmermans, H., Borgers, A. and P. van der Waerden (1992) Mother Logit Analysis of Substitution Effects in Consumer Shopping Destination Choice, *Journal of Business Research*, Vol. 24, pp. 177-189.

Train, K. (1998) Recreation Demand Models with Taste Differences Over People, *Land Economics*, Vol. 74, pp. 230-239.

Train, K. (1999) Halton Sequences for Mixed Logit, Working Paper No. E00-278, Department of Economics, University of California, Berkeley.

Tuffin, B. (1996) On the use of Low-Discrepancy Sequences in Monte Carlo Methods, *Monte Carlo Methods and Applications*, Vol. 2, pp. 295-320.

Van der Corput, J.G. (1935a) Verteilungsfunctionen I Nederl. Akad. Wetensch, Proc. 38, pp. 813-820.

Van der Corput, J.G. (1935b) Verteilungsfunctionen II Nederl. Akad. Wetensch, Proc. 38, pp. 1058-1066.

Waddell, P. (1993) Exogenous Workplace Choice in Residential Location Models: Is the Assumption Valid?, *Geographical Analysis*, Vol. 25, pp. 65-82.

Waddell, P. (1996) Accessibility and Residential Location: The Interaction of Workplace, Residential Mobility, Tenure, and Location Choices, Presented at the Lincoln Land Institute TRED Conference. http://www.odot.state.or.us/tddtpan/modeling.html

Wang, X. and F.J. Hickernell (2000) Randomized Halton Sequences, *Mathematical and Computer Modelling*, Vol. 32,pp. 887-899.

Wesolowsky, G.O. (1973) Location in continuous space, *Geographical Analysis*, Vol. 5, pp. 95-112.

Wheeler, J. O. and F. P Stutz (1971) Spatial Dimensions of Urban Social Travel, *Annals of Association of American Geographers*, Vol. 61, 371-386.

Wilson, A. G. (1967) A Statistical Theory of Spatial Distribution Models, *Transportation Research*, Vol. 1, pp. 253-269.

Wilson, A. G. (1971) A Family of Spatial Interaction Models and Associated Developments, *Environment and Plannijng A*, Vol. 3, pp. 1-32.

Wilson, A. G. (1974) Urban and Regional Models in Geography and Planning, John Wiley, Chichester, Sussex.

Xue, Y. and D. E. Brown (2003) A Decision Model for Spatial Site Selection by Criminals: A Foundation for Law Enforcement Decision Support, *IEEE Transactions on Systems, Man, and Cybernetics*, Part C Vol. 33 (1), pp. 78-85.

VITA

Aruna Sivakumar was born in Secunderabad, Andhra Pradesh, India on December 8, 1977, the daughter of Sivakumar Balasubramaniam and Anandavalli Sivakumar. She did her schooling over multiple cities in South India including Hyderabad, Trivandrum and Chennai. After completing her secondary education at Vidya Mandir Senior Secondary School, Chennai, she joined the Indian Institute of Technology, Madras, (IITM) in 1995. She received the degree of Bachelor of Technology (B.Tech.) in Civil Engineering from IIT Madras in August 1999, at which point she was accepted for graduate studies at the University of Texas at Austin. She arrived in Austin, Texas, USA on August 9, 1999, and in August 2001 she received the degree of Master of Science in Engineering (MSE) at the University of Texas at Austin. She began her doctoral study in Transportation Engineering at the University of Texas at Austin in September 2001.

Permanent Address: #14, Raman Avenue
Cart Track Road, Velachery
Chennai 600 042
Tamil Nadu, India

This dissertation was typed by the author.