ABSTRACT

Title of Dissertation:	CONTINUOUS CHOICE MODELS FOR TIME-OF-DAY CHOICE MODELING APPLICATIONS
	Sepehr Ghader, Doctor of Philosophy, 2019
Dissertation directed by:	Professor Lei Zhang, Department of Civil and Environmental Engineering

I propose a series of tools to model travelers' time-of-day choice in continuous time. The models discussed in this dissertation can help advancing time-of-day modeling of trips or activities and produce demand with fine time resolution. These models are a good fit for dynamic traffic assignment and they can be applied for policy evaluation, travel management, and real-time applications. I first present the Continuous Logit (CL) model as the originator of a variety of discrete and continuous choice models and shed light on the relationship between some of the available choice models and CL by showing how these models can be seen as approximations to the CL. I also demonstrate how different approximation techniques can lead to new forms of choice models. I conduct Monte Carlo experiments to study the magnitude of error in the approximated models. These experiments can help the reader better understand the implications of various approximation and discretization schemes for time-of-day modeling.

Due to the limits of CL in modeling correlations, I introduce and formulate the AutoRegressive Continuous Logit (ARCL) as a novel continuous class of choice models capable of representing correlations across alternatives in the continuous spectrum. I formulate this model by considering two approaches: combining a discrete-time autoregressive process of order one with the CL model, and combining a continuous-time autoregressive process with the CL model. ARCL is the only Random Utility Maximization-based continuous choice model, besides the Continuous Cross-Nested Logit (CCNL), able to handle correlations across alternatives in the continuous spectrum.

I extend the continuous time-of-day modeling to multi-dimensional case by introducing a framework to model the joint choice of arrival to an activity and departure from the activity. Each choice is modeled in continuous time using CCNL. I use Copula to capture the correlation between the two dependent choices. Copula can model the correlation structure without knowing the actual bivariate distribution function. With its multidimensionality and ability to capture different sorts of correlations and model demand in fine time resolution, the introduced framework can provide a sufficient tool for the time-of-day component of various travel demand models.

CONTINUOUS CHOICE MODELS FOR TIME-OF-DAY CHOICE MODELING APPLICATIONS

by

Sepehr Ghader

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2019

Advisory Committee:

Professor Lei Zhang, Chair
Professor Anna Alberini
Professor Cinzia Cirillo
Professor Ali Haghani
Professor Benjamin Kedem, Dean’s Representative

ACKNOWLEDGEMENT

I would like to express the deepest appreciation to my advisor and committee chair Professor Lei Zhang, for his continuous support and vast knowledge. Without his guidance and persistent help, writing this dissertation would not have been possible. I would also like to thank my committee members, Professor Anna Alberini, Professor Cinzia Cirillo, Professor Ali Haghani, and Professor Benjamin Kedem for their time, support, and valuable comments.

This dissertation work would not have been possible without the help and guidance of my colleagues, Dr. Carlos Carrion, Dr. Liang Tang, and Arash Asadabadi. I am truly grateful to them, and all of those with whom I have had the pleasure to work during the past couple of years at the University of Maryland.

I would also like to thank National Transportation Center and Maryland Transportation Institute at the University of Maryland for their financial support.

Nobody has been more important to me in the pursuit of my studies than the members of my family. I would like to thank my parents and my sister, whose love and guidance are with me in whatever I pursue.

TABLE OF CONTENTS

ACKNOWLEDGEMENT ii

TABLE OF CONTENTS iii

LIST OF TABLES vi

LIST OF FIGURES vii

1. INTRODUCTION 1

1.1 OVERVIEW 1

1.2 OBJECTIVES 3

1.3 CONTRIBUTIONS 5

1.4 ORGANIZATION 7

2. LITERATURE REVIEW 9

3. REVISITING THE CONTINUOUS LOGIT 15

3.1 CONTINUOUS LOGIT MODEL 15

3.1.1 BACKGROUND 15

3.1.2 PROPERTIES, SPECIFICATION AND IDENTIFICATION 20

3.2 APPROXIMATING THE CONTINUOUS LOGIT FOR TIME-OF-DAY CHOICE MODELING 25

3.2.1 APPROXIMATIONS IN THE FORM OF DISCRETE CHOICE MODELS 26

3.2.2 APPROXIMATIONS IN THE FORM OF CONTINUOUS CHOICE MODELS 29

3.3 MONTE CARLO EXPERIMENTS OF APPROXIMATIONS 31

3.3.1 QUANTIFYING THE APPROXIMATION ERROR 35

3.3.2 QUANTIFYING THE BIAS OF THE MAXIMUM LIKELIHOOD ESTIMATOR DUE TO THE APPROXIMATION ERROR 41

3.4 ANALYTICAL BOUNDS FOR THE APPROXIMATION ERROR OF THE DERIVED MODELS 43

3.5 SUMMARY AND DISCUSSION 49

4. AUTOREGRESSIVE CONTINUOUS LOGIT 52

4.1 DISCRETE-TIME AUTOREGRESSIVE CONTINUOUS LOGIT MODEL . 53

4.1.1 FORMULATION 53

4.1.2 IDENTIFICATION 58

4.1.3 MONTE CARLO EXPERIMENT 60

4.1.4 COMPARISON WITH CONTINUOUS LOGIT AND CONTINUOUS CROSS-NESTED LOGIT 65

4.2 CONTINUOUS-TIME AUTOREGRESSIVE CONTINUOUS LOGIT MODEL 70

4.2.1 FORMULATION 70

4.2.2 RELATION TO THE DISCRETE-TIME AUTOREGRESSIVE CONTINUOUS LOGIT MODEL 79

4.2.3 MONTE CARLO EXPERIMENT 80

4.3 SUMMARY AND DISCUSSION 84

5. COPULA-BASED CONTINUOUS CROSS-NESTED LOGIT 86

5.1 FORMULATION 87

5.1.1 COPULA 87

5.1.2 CONTINUOUS CROSS NESTED LOGIT MODEL 89

5.1.3 UTILITY FUNCTION SPECIFICATION 91
5.1.4 ESTIMATION 93
5.1.5 APPROXIMATION 94
5.2 MONTE CARLO EXPERIMENT 95
5.3 EMPIRICAL APPLICATION 100
5.3.1 DATA 101
5.3.2 TRAVEL TIME REGRESSION 102
5.3.2 UTILITY SPECIFICATION 106
5.3.3 COPULA SELECTION 107
5.3.4 MODEL ESTIMATION RESULTS 110
5.3.5 COMPARISON WITH OTHER MODELS 111
5.4 SUMMARY AND DISCUSSION 113
6. CONCLUDING REMARKS 115
REFERENCES 119

LIST OF TABLES

Table 1. True parameters for generating the synthetic data 34

Table 2. Monte Carlo estimates of the statistical bias of MLE due to the approximation error 43

Table 3. True parameters for generating the synthetic data - discrete-time autoregressive continuous logit model 63

Table 4. Estimates of the parameters using the synthetic data - discrete-time autoregressive continuous logit model 65

Table 5. Computation time comparison 68

Table 6. True parameters for generating the synthetic data - continuous-time autoregressive continuous logit model 81

Table 7. Estimates of the parameters using the synthetic data - continuous-time autoregressive continuous logit model 82

Table 8. True parameters for generating the synthetic data in the Monte Carlo experiment 98

Table 9. Estimates of the parameters using the synthetic data – copula-based continuous cross-nested logit 100

Table 10. Estimated parameters for arrival travel time regression 104

Table 11. Estimated parameters for departure travel time regression 105

Table 12. Estimation results for HBW car tours in Washington D.C.-Baltimore area 111

Table 13. Comparison of different models 112

LIST OF FIGURES

Figure 1. Examples of trigonometric functions with $K = 4$ 24

Figure 2. Probability density function of the continuous logit for the simulated data 38

Figure 3. Approximating error $\xi t \in \delta k - 1, \delta k$ for the simulated data set 39

Figure 4. Approximating error $\xi(t)$ for the simulated data set.................................. 40

Figure 5. Effects of the parameters on the AR(1) process 56

Figure 6. Discrete-time autoregressive continuous logit (period size: 10 minutes) vs. continuous logit and continuous cross-nested logit. 69

Figure 7. Continuous-time autoregressive continuous logit (Euler-Maruyama Partitions: 72) vs. continuous logit. 83

Figure 8. Approximation error in PDF evaluation using midpoint rule with 25 subintervals. 95

Figure 9. Travel time profiles for arrival and departure 106

Figure 10. Comparing the data with the draws from different copulas 109

1. INTRODUCTION

1.1 OVERVIEW

Time-of-day choice is a fundamental choice dimension of any trip or tour. The modeling of this choice dimension is required for the analysis of transportation system performance and the evaluation of travel demand management policies, especially those policies targeted at the temporal dimension of demand. In addition, time-of-day choice models are essential components for the integration of agent or activity-based travel demand models with the Dynamic Traffic Assignment (DTA) models.

In the transportation research literature, several econometric models have been applied to study the time-of-day choice of travelers. I group these models into two categories: (1) models considering time-of-day choice as a discrete dependent variable; and (2) models considering time-of-day choice as a continuous dependent variable. The main drawbacks with the models in the first category are the lack of a theoretical foundation to aid in the decision of the analyst with regards to choosing a suitable temporal resolution, and the forecasting of these models only into aggregate intervals. Continuous time-of-day models are preferred to discrete models because time is a continuous variable and the time-of-day choice is a continuous choice in nature. Discretizing a time variable involves putting adjacent time points in different intervals. As a result, the model treats them differently, while they might be considered the same by the travelers. Besides, there is no robust rule for setting the intervals. Furthermore, model applications may need point prediction which is not possible in models with a discrete time variable.

The models in the second category are further divided into two general types: (a) Hazard Models, and (b) Continuous choice models consistent with the Random Utility Maximization (RUM). The main disadvantage with the hazard models is the lack of behavioral theory foundation in contrast to the random utility models (both discrete choice models and continuous choice models). In time-of-day choice modeling, continuous choice models consistent with RUM include: the continuous logit and the continuous cross-nested logit. These models both treat time as continuous and have a behavioral theoretical foundation. The main difference between the continuous logit and the continuous cross-nested logit is that the latter allows for correlation between similar timing decisions in the continuous spectrum.

In recent years, decision making in transportation has shifted from long-term infrastructure investments to short-term systems management and operations solutions. Despite all the advancement in travel demand modeling, models for time-of-day choice are greatly simplified in terms of specifications and applications. Travel demand modeling literature shows that time-of-day modeling is one of the major weaknesses of current travel demand models. Timing and scheduling models should catch-up with advancements in DTA. Many researchers and practitioners now use DTA for traffic assignment (Zhang, et al., 2018), which requires a travel demand with fine time resolution. The continuous time-of-day models are able to predict the demand in small intervals; in fact, they are able to give point prediction of the demand, which makes them a good fit for DTA assignment. Integrating continuous time-of-day models with DTA makes them suitable tools for travel management and real-time applications. Modeling travel demand in continuous form or small intervals can also benefit modeling effects of time-dependent policies such as variable speed limit, congestion pricing, and dynamic message signs. Significant improvements in computation power and continuous advancements in big data are also encouraging researchers to develop more advanced time-of-day models. Travel-time data, speed data, and location data being gathered minute-by-minute from devices all around the world, enable the estimation of models for minute-by-minute travel demand prediction that would be suitable tools for real-time applications. Given the aforementioned applications and data availability, the relevant modeling tools and theory for time-of-day modeling should be further developed.

1.2 OBJECTIVES

This dissertation has three key objectives. The first key objective is to revisit continuous logit, the simplest continuous choice model consistent with RUM used for time-of-day modeling, in order to investigate the errors corresponding to the use of various approximation and discretization schemes for time-of-day modeling. This dissertation seeks to illustrate how some of the simple discrete choice models (such as multinomial logit) are in fact approximations to the continuous logit and show how new discrete and continuous choice models can be derived from the continuous logit using different approximation schemes. I perform numerical experiments and present analytical proofs to quantify the approximation error and statistical bias related to the use of the discrete and continuous choice models derived from the continuous logit. The results can show how small changes in the approximation scheme can lead to significant improvements in the approximation error.

Continuous logit has Independence from Irrelevant Alternative (IIA) property. It cannot capture the correlation among alternatives. The second key objective is to introduce and formulate a novel class of continuous choice model consistent with RUM that is able to model various types of correlation across alternatives. This class of models, called autoregressive logit, is derived from combining continuous logit and an autoregressive process. This dissertation seeks to introduce autoregressive continuous logit to the reader, and show its properties through numerical experiments. The autoregressive continuous logit is a RUM model besides the continuous cross-nested logit that can capture correlation across alternatives in continuous time.

The continuous cross-nested logit and the autoregressive logit can both capture correlation across alternatives, but their application is limited to the one-dimensional choice situation. A multi-dimensional framework that can capture correlation across dependent time-of-day choices is required for proper time-of-day modeling in an activity-based or tour-based modeling context. The third key objective is to introduce a complete framework for time-of-day modeling with a continuous time variable that can be used in activity-based models and replace the current simplified time-of-day components of these models. Scheduling decision for activities and tours usually involve more than one time-of-day choice, i.e. choice of arrival to the activity and choice of departure from the activity. Consequently, the introduced framework needs to consider multi-dimensional choices, and account for the correlation among dependent choices in addition to correlation among alternatives of each choice, in continuous time. This study seeks to introduce a new framework called the copula-based continuous cross-nested logit for this purpose, explore its properties using Monte Carlo experiments, and showcase its real-world application through an example.

1.3 CONTRIBUTIONS

When dealing with mode choice, destination choice, or route choice, the decision variable is discrete in nature. However, time-of-day choice is different; in that any time-point in the continuous spectrum of time can be chosen by a traveler and there is no obvious discrete set of alternatives. Choice models treating time as a discrete variable assume decision makers choose time intervals instead of time points. One can question the validity of this assumption. How should the intervals be formed? What is the magnitude of error caused by this assumption? This dissertation shows that some of the choice models used for time-of-day modeling are in fact approximations to the continuous logit and evaluates the validity of the discrete-time assumption by focusing on the approximation error. The first main contribution of this study to the literature is investigating the error corresponding to the use of various discrete and continuous choice models, shown to be approximations of the continuous logit, for modeling time-of-day choice. This study is the first to demonstrate how new types of continuous and discrete choice models can be derived from continuous logit through numerical approximation methods, and to quantify approximation error and statistical bias of the Maximum Likelihood Estimation (MLE) for the derived models through numerical experiments and analytical proofs.

The second main contribution of this study is the introduction of a novel class of continuous choice models consistent with the RUM that can relax continuous logit's

IIA property and model various types of correlation across alternatives. This class of models, referred to as autoregressive continuous logit, is derived from the combination of continuous logit and an autoregressive process. The autoregressive continuous logit model is formulated, with applications to time-of-day choice modeling in mind, by considering two approaches: combining a discrete-time autoregressive process of order one (i.e., a linear stochastic difference equation) with the continuous logit model, which leads to the discrete-time autoregressive continuous logit; and combining a continuous-time autoregressive process (i.e., a linear stochastic differential equation), known in the stochastic process literature as Ornstein-Uhlenbeck process, with the continuous logit model, which leads to the continuous-time autoregressive continuous logit.

The third contribution of this study is extending the time-of-day modeling with continuous time variable to the multi-dimensional case and introducing the copula-based continuous cross-nested logit framework. This novel modeling framework is the only framework in the literature able to consider both the correlation between two choices and the correlation between alternatives of each choice (all possible correlations; both on a tour, and in a day) in continuous time. Modeling two choices together makes this framework compatible with the mainstream of activity-based models. The fully continuous, multi-dimensional modeling framework introduced in this dissertation can act as a substitute for simpler time-of-day components of many activity-based models to make them suitable tools for evaluating time-dependent policies and real-time applications.

1.4 ORGANIZATION

The second chapter is dedicated to the literature review. I reviewed the time-of-day modeling literature and summarized the methods used for time-of-day modeling.

In the third chapter of this dissertation, I present analytical and numerical results for the application of the continuous logit model for time-of-day choice modeling. The third chapter includes a brief background of the continuous logit model, derivations of models based on approximations to the continuous logit model for time-of-day choice modeling, a discussion of Monte Carlo experiments of the derived models, and a discussion of analytical results for bounding the approximation error of the derived models. The continuous logit has IIA property, and it fails to capture the correlation across alternatives.

In the fourth chapter, I focus on the correlation across alternatives for a one-dimensional choice of arrival to the activity, or departure from the activity, and formulate the discrete-time autoregressive continuous logit model and the continuous-time autoregressive continuous logit model. The fourth chapter includes formulation of the discrete-time autoregressive continuous logit model and the continuous-time autoregressive continuous logit model for time-of-day choice modeling, a discussion of the results of Monte Carlo experiments to illustrate these models numerically, and a comparison between the autoregressive continuous logit, the continuous logit, and the continuous cross-nested logit. While continuous cross-nested logit and autoregressive continuous logit can both capture the correlation across alternatives for a single choice, capturing correlation across different choices is also essential. Time-of-day decisions making usually consists of a series of dependent choice situation, for instance, the choice of arrival time to the activity and the departure time from the activity. Modeling the dependence of such choices is essential in tour-based or activity-based travel demand modeling.

In the fifth chapter, I extend the modeling framework into the multi-dimensional case, and introduce the copula-based continuous cross-nested logit to model the joint choice of arrival to the activity and departure from the activity. The fifth chapter includes formulation of the copula-based continuous cross-nested logit framework, numerical analysis of the framework using Monte Carlo simulation, and an empirical application of the model to showcase how it can be applied in real word applications.

The sixth chapter is the concluding chapter that summarizes the key findings and discusses the future directions.

2. LITERATURE REVIEW

In the transportation research literature, several econometric models have been applied to study the time-of-day choice of travelers. I group these models into two categories: (1) models considering time as a discrete variable; and (2) models considering time as a continuous variable.

The models in the first category discretize time into a set of finite alternatives that are suitable to the application of discrete choice models consistent with the RUM framework (Ben-Akiva and Lerman, 1985, Train, 2009). Discrete choice models for the time-of-day follow the work of Small (1982) (Small, 1982). In the research literature, the level of discretization (e.g. splitting time of 24 hours into 24 alternatives representing 1-hour intervals) varies widely as well as the time period (e.g. AM peak hour vs. the entire day) of analysis. Researchers have considered alternatives in the temporal resolution of 5-minutes intervals to 15-minutes intervals (Abkowitz, 1981, Hendrickson and Plank, 1984, Small, 1982), and others have considered alternatives in the temporal resolutions of 30-minutes to 1-hour (Ben-Akiva and Abou-Zeid, 2013, Popuri, et al., 2008, Zeid, et al., 2006). In addition, researchers have also applied these models to multi-dimensional time-of-day choice modeling by considering arrival time choice (i.e. arrival to the activity) and departure time choice (i.e. departure from the same activity) in an activity scheduling setting (Ben-Akiva and Abou-Zeid, 2013, Popuri, et al., 2008, Zeid, et al., 2006). Moreover, researchers have also accommodated the issue of correlation between alternatives in time-of-day choice modeling through the use of Generalized Extreme Value models (GEV; e.g. nested logit in Chin (1990)

(Chin, 1990), ordered generalized extreme value in Small (1987) (Small, 1987), and Mixture GEV models such as error component logit in Bhat (1998), De Jong, et al. (2003), and Hess, et al. (2007) (Bhat, 1998, De Jong, et al., 2003, Hess, et al., 2007)) introduced by McFadden (1987) (McFadden, 1978). For instance, the multinomial logit is a simple type of GEV models that assumes error terms are Independent Identically distributed (IID) Gumbel, resulting in no correlation among the error terms. An example can be seen in Zeid, et al. (2006) (Zeid, et al., 2006). The nested logit is another type of GEV models, which divides the alternatives into nests. The Error terms for alternatives in the same nest are correlated, while alternatives in different nests have independent error terms. An example for work trips can be seen in Chin (1990) (Chin, 1990). Another type of GEV models similar to the nested logit is the ordered generalized extreme value model introduced by Small (1987) (Small, 1987), which is used with ordered alternatives. An example of the ordered generalized extreme value can be found in Bhat (1998) (Bhat, 1998). The last aforementioned type of GEV models is the mixed logit, which has been known since Cardel and Dunbar (1980) and Bolduc and Ben-Akiva (1991) as a highly flexible yet practical model type (Bolduc and Ben-Akiva, 1991, Cardell and Dunbar, 1980). In the literature, mixed logit models are in two forms: the error components logit and the random coefficient logit. According to McFadden and Train (2000), the error components logit can approximate any type of discrete choice models based on the RUM as closely as one pleases (McFadden and Train, 2000). Some examples of the mixed logit model can be seen in Bhat (1998), De Jong (2003), and Borjesson (2008) (Bhat, 1998, Börjesson, 2008, De Jong, et al., 2003). Another widely-used type of discrete choice models is the multinomial probit, which assumes normally distributed error terms. The multinomial probit model is able to consider a complete variance-covariance matrix, at the expense of evaluating high-dimensional multivariate-normal integrals for the choice probabilities. The multinomial probit has been used to some extent in the literature, such as Liu and Mahmassani (1998) in which a multinomial probit model is estimated by exposing constraints on the covariance matrix (Liu and Mahmassani, 1998). However, application of the multinomial probit still requires considerable computational power.

The main drawbacks of the discretization in the models of the first category are the lack of a theoretical foundation to aid in the decision of the analyst with regards to choosing a suitable temporal resolution, and the forecasting of these models only into aggregate intervals. The models in the second category are further divided into two general types: (a) Hazard Models, and (b) Continuous choice models consistent with the RUM. Several examples of Hazard models include Wang (1996) (parametric hazard modeling of activity start times), Bhat (1996) and Bhat and Steed (2002) (nonparametric hazard modeling for shopping trips), Komma and Srinivasan (2008) (nonparametric hazard modeling for commute trips), and Gadda, et al. (2009) (Bayesian estimation techniques for hazard modeling) (Bhat, 1996, Bhat and Steed, 2002, Gadda, et al., 2009, Komma and Srinivasan, 2008, Wang, 1996). The main disadvantage with the models in this category is the lack of behavioral theory foundation in contrast to the RUM models (both discrete choice models and continuous choice models). In time-of-day choice modeling, continuous choice models consistent with the RUM include: the continuous logit (Lemp and Kockelman, 2010, Lemp, 2009) and the continuous cross-nested logit (Lemp, 2009, Lemp, et al., 2010). These models both treat time as continuous and have a behavioral theoretical foundation. The main difference between the continuous logit and the continuous cross-nested logit is that the latter allows for correlation between similar timing decisions in the continuous spectrum.

Besides the econometric models, the literature review suggests that there are other ways to model the temporal component for travel demand modeling such as: post-processing technique of applying hourly factors; link-based or trip-based adjustments to address the problem of projected demand-exceeding capacity; the Equilibrium Scheduling Theory (EST); and rule-based models.

Applying hourly factors is the most basic approach for estimating volumes for hourly analysis. The factors are widely used because of their simplicity and their ability in providing a rough estimate of peak hour traffic volume. However, the hourly factor method is a static process, not able to allow any types of temporal or geographical changes. In addition, the factors are not sensitive to policy changes, congestion level, or capacity constraints. The Maryland Statewide Transportation Model (MSTM) version 1.0, similar to many other trip-based models, uses this method with four time periods, namely morning peak, midday, afternoon peak, and night (Costinett, et al., 2009). The state of Maryland is now developing an activity-based model for the state.

Link-based and trip-based methods are other ways for time-of-day modeling. They use the capacity of the links and do not allow demand to exceed capacity during the peak hour by shifting demand to the shoulders of the peak. Link-based methods are more realistic than hourly factors, and they are sensitive to congestion; however, they lack behavioral assumptions. Furthermore, the continuity of flow is not guaranteed. Trip-based methods are preferred to link-based, since they can keep the continuity of flow. They revise trip tables in order to reduce trips on the links on which demand exceeds capacity.

The EST uses direct equilibration of simple models of supply and demand. These models are based on Vickrey's bottleneck model (Hyman, 1997, Vickrey, 1969). In EST, Vickrey's model is extended in a number of aspects, such as consideration of heterogeneous users. The theory can be generalized to be applied to transportation networks. It can also be integrated with DTA. The positive aspect of EST is its ability to model continuous time. The biggest negative feature is being deterministic and assuming that there is no unmeasured interpersonal variation. The other negative issue is that the effects of socio-economics and demographics can only be seen in preferred arrival time (PAT) estimation. One example of EST is the heterogeneous arrival and departure time, based on the equilibrium scheduling theory (HADES) discussed by van Vuren, et al. (1999) (van Vuren, et al., 1999). The conclusion of this study states that HADES is the final stage of EST development, and further research should be focused toward discrete choice models.

Most of the aforementioned models were based on rational behavior, assuming travelers are able to identify all their feasible alternatives, measure all their attributes and choose accordingly to maximize their utility. Rule-based models avoid this assumption of rationality, and try to model how travelers actually make decisions through learning, knowledge, and searching. One example is the positive model of departure time choice by Xiong and Zhang (2013) (Xiong and Zhang, 2013, Zhang, 2007). Rule-based models usually need a large amount of data for training. Their main focus is usually improving the predictive accuracy, not explaining the travel behavior.

3. REVISITING THE CONTINUOUS LOGIT

This chapter presents analytical and numerical results for the application of the continuous logit model for time-of-day choice modeling. Firstly, it is shown that the continuous logit model is the originator of the discretization scheme of the multinomial logit as it is currently applied to time-of-day choice modeling. Secondly, new discrete choice models and new continuous choice models are derived through the application of simple numerical integration rules to the continuous logit model. Thirdly, the approximation error of these derived models is studied through Monte Carlo experiments. These experiments also study the presence of bias due to the approximation error. Also, analytical results are presented with regards to the bounding of the error of the derived models. The chapter is organized as follows: a brief background of the continuous logit model including its properties, specification and identification; derivations of models based on approximations to the continuous logit model for time-of-day choice modeling; a discussion of Monte Carlo experiments of the derived models; a discussion of analytical results for bounding the approximation error of the derived models; and conclusions.

3.1 CONTINUOUS LOGIT MODEL

3.1.1 BACKGROUND

The continuous logit is a continuous choice model that accommodates for a continuous dependent variable within the behavioral framework of the RUM. This model was initially formulated by McFadden (1976) and formalized by Ben-Akiva, et al. (1985)

and Ben-Akiva and Watanatada (1981) (Ben-Akiva, et al., 1985, Ben-Akiva and Watanatada, 1981, McFadden, 1976). It was recently applied to time-of-day choice modeling by Lemp and Kockelman (2010) and Lemp (2009) using Bayesian estimation methods (Lemp and Kockelman, 2010, Lemp, 2009). Previously, the continuous logit was only used in spatial choice modeling by Ben-Akiva, et al. (1985) and Ben-Akiva and Watanatada (1981) (Ben-Akiva, et al., 1985, Ben-Akiva and Watanatada, 1981).

The probability density function (PDF) of the continuous logit for a continuous variable t, where t ∈ [a, b] and both a and b are constants, is given by,

$$f(t) = \frac{g(t)e^{(\mu V(t))}}{\int_a^b g(z)e^{(\mu V(z))}dz}, \tag{1}$$

where the $g(\cdot)$ function is equivalent to the availability matrices in the discrete choice models. These matrices control the availability of alternatives to the individuals in the sample. For continuous choice models, the $g(t)$ function may be a step function (i.e., allow 1 or 0 values for availability of alternatives) or a density function of alternatives for a given value of *t*. This $g(t)$ is included in the formulation by Ben-Akiva, et al. (1985) and Ben-Akiva and Watanatada (1981) (Ben-Akiva, et al., 1985, Ben-Akiva and Watanatada, 1981). For the purposes of time-of-day choice modeling, the function $g(t)$, where t represents a point in a time period, may be used as a step function (0 or 1 values) for restricting the alternatives available for the departure time from an activity to be after the arrival time to the same activity in tours. The μ is a positive constant that has the same function as the scale in the multinomial logit model. The function $V(\cdot)$ is the systematic utility function for the continuous logit model. Attributes specified along with coefficients in the systematic utility function must be specified in terms of functions of t. For example, the travel cost function $c(t)$ for a given time point t could be represented as a step function. Also, note that t represents a continuous random variable, which is the choice of an individual.

The probability distribution or cumulative distribution function (CDF) of the continuous logit for a continuous variable *t*, where $t \in [a, b]$ is given by,

$$F(t) = Prob[T \leq t] = \frac{\int_a^t g(s)e^{(\mu V(s))}ds}{\int_a^b g(z)e^{(\mu V(z))}dz} \tag{2}$$

Also note that it is clear that the probability of $t \in [c, d] \subseteq [a, b]$ is given by,

$$Prob[c \leq T \leq d] = \frac{\int_c^d g(s)e^{(\mu V(s))}ds}{\int_a^b g(z)e^{(\mu V(z))}dz} \tag{3}$$

It should be noted that $f\ (t) = 0$ for values of *t* such that $t \notin [a, b]$, and thus $F(t) = 0$ for values of *t* such that $t < a$, and $F(t) = 1$ for values of *t* such that $t \geq b$.

The log-likelihood function for a random sample of size *N* for the continuous logit model, where *n* is an index for an individual in the sample, is given by,

$$LL(\beta) = \sum_{n=1}^{N} \ln\left(\frac{g_n(t_n)e^{(\mu V(t_n;\beta))}}{\int_a^b g_n(z)e^{(\mu V(z;\beta))}dz}\right), \tag{4}$$

where β is a vector of parameters to be estimated. This log-likelihood function may be maximized to obtain the Maximum Likelihood estimates (ML; see Cramer (1986) for details about this estimator (Cramer, 1986)) as well as the covariance matrix to obtain the standard errors of the parameters. It should be noted that the variable t_n

is indexed by n to indicate that t_n represents the observed value of the variable t for an individual n in the sample. Also, the function $g_n(t)$ is indexed by n to represent the ability to accommodate for individual-specific availability restrictions. This model may also be estimated using Bayesian methods, but in this dissertation, all the models are estimated using the MLE.

Readers familiar with the multinomial logit (see Ben-Akiva and Lerman (1985) for more information about multinomial logit (Ben-Akiva and Lerman, 1985)) will note the similarities between the multinomial logit and the continuous logit. Here, I just present some basic results for comparison.

The probability mass function of the multinomial logit for a generic alternative i, usually referred as probably choice function, is given by

$$f(i) = Prob[I = i] = \frac{A(i)e^{\mu V(i)}}{\sum_{j=1}^{J} A(j)e^{\mu V(j)}} \tag{5}$$

Clearly, a discrete random variable I represents the choice of an individual, where I only takes discrete values representing the alternatives with their associated probabilities. The probabilities are those given in the probability mass function. In Ben-Akiva and Lerman (1985), it is shown that a vector of latent continuous random variables, referred as random utilities, are defined in order to obtain the derivation of the multinomial logit model (Ben-Akiva and Lerman, 1985). The main idea behind the derivation is that for each discrete alternative i to be chosen, its associated random utility must be greater than the maximum of the random utilities of the other alternatives. In mathematical terms, $I = i$ if $U_i > \max_{j \in \{1,2,\ldots,J\}, j \neq i} U_j$. Also, the function

$A(.)$ is defined as the availability matrix, in which a row represents the alternatives available to an individual in the sample. The elements of this matrix are 1 and 0. It should be noted that J represents the number of alternatives, and thus the choice set is simply defined as the set $\{1,2,\dots,J\}$.

Also, if the alternatives follow an order (e.g. time alternatives), the discrete probability distribution of the multinomial logit for alternative i is given by,

$$F(i) = Prob[I \leq i] = \sum_{z \leq i} \frac{A(z)e^{\mu V(z)}}{\sum_{j=1}^{J} A(j)e^{\mu V(j)}} \tag{6}$$

The log-likelihood function for a random sample of size N for the multinomial logit model, where n is an index for an individual in the sample, is given by

$$ll(\beta) = \sum_{n=1}^{N} \ln\left(\frac{A(z)e^{\mu V(i_n;\beta)}}{\sum_{j=1}^{J} A(j)e^{\mu V(j;\beta)}}\right) \tag{7}$$

Similar to t_n, I define i_n to mean the observed chosen alternative by the individual n in the sample. Multinomial logit models are typically estimated using the MLE.

Lastly, the set of integration for the introduced continuous logit model is for one single variable t, where $t \in R$ and $t \in [a,b]$. This set of integration may be extended depending on the modeling needs of the analyst. For example, an analyst considering the departure time choice from an activity, and the arrival time choice for the same activity may model this as an ordered pair (t_{dp}, t_{arr}), where t_{dp} is the choice of the departure time, and t_{arr} is the choice of the arrival time. Moreover, these variables are defined for a 24 hour time period such that $t_{dp} \in [0,24]$, and $t_{arr} \in [0,24]$. The Cartesian product of both sets (i.e. $[0,24] \times [0,24]$) defines the set of integration. Thus, the probability density function for the continuous logit of this example is given by,

$$f(t_{dp}, t_{arr}) = \frac{g(t_{dp},t_{arr})e^{\mu V(t_{dp},t_{arr})}}{\int_0^{24}\int_0^{24} g(s,z)e^{\mu V(s,z)}dsdz} \tag{8}$$

For my purposes, I only consider modeling the departure time choice from a location for a trip, or the arrival time choice to a location for a trip. Also, I will not consider the $g(.)$ functions, or in other words, I assume that they are always 1, for all values of the choice variable t in $[a, b]$.

3.1.2 PROPERTIES, SPECIFICATION AND IDENTIFICATION

In this section, I discuss briefly some of the properties, the specification requirements, and the theoretical identification of the continuous logit model. Also, the continuous logit model shares similar features (i.e. properties, specification, and identification) with the multinomial logit model, and Ben-Akiva and Lerman (1985) discusses them at length for the multinomial logit model (Ben-Akiva and Lerman, 1985).

3.1.2.1 Properties

The continuous logit model shares the following similar limiting cases with the multinomial logit model (Ben-Akiva and Lerman, 1985):

- $\lim_{\mu \to 0} \frac{e^{\mu V(t)}}{\int_a^b e^{(\mu V(z))}dz} = \frac{1}{b-a}$

In this case, the density function *f (t)* of the continuous logit as $\mu \rightarrow 0$ becomes the probability density function of the continuous uniform distribution defined in the interval [a, b]. In the case of the multinomial logit, the probability mass function of this model becomes the probability mass function of the discrete uniform distribution for the same number of alternatives.

- $$\lim_{\mu\to\infty} f(t) = \lim_{\mu\to\infty} \frac{e^{\mu V(t)}}{\int_a^b e^{(\mu V(z))} dz} = \lim_{\mu\to\infty} \frac{1}{\int_a^b e^{(\mu V(z) - \mu V(t))} dz}$$

To calculate the limit, fix *t*, partition the integral at *t*, and apply the mean value theorem (Binmore, 1982):

$$\lim_{\mu\to\infty} \frac{1}{(t-a)e^{(\mu V(z_1) - \mu V(t))} + (b-t)e^{(\mu V(z_2) - \mu V(t))}}$$

Where $z_1 \in (a, t)$ and $z_2 \in (t, b)$. The cases for this limit are as follows,

$$\lim_{\mu\to\infty} f(t) = \begin{cases} \infty & V(t) > V(z_1) \wedge V(t) > V(z_2) \\ 0 & V(t) < V(z_1) \wedge V(t) < V(z_2) \\ 0 & V(t) < V(z_1) \wedge V(t) > V(z_2) \\ 0 & V(t) > V(z_1) \wedge V(t) < V(z_2) \end{cases}$$

The probability density function collapses for *t*, such that $\forall z \in [a, b], V(t) \geq V(z)$ as $\mu \rightarrow \infty$, and the cumulative distribution function becomes $F(\infty) = 1$ at *t*. This result is similar to the Dirac delta function and the unit step function (Khuri, 2004). The probabilistic choice model becomes deterministic. It should be noted that I do not consider pathological cases for the functional form of $V(\cdot)$. A similar result holds for the multinomial logit.

Another important property of the continuous logit that is shared with the multinomial logit is IIA. For example, the ratio of the probability of $t \in [c, d] \subseteq [a, b]$ to the probability of $t \in [e, f] \subseteq [a, b]$ is given by,

$$\frac{Prob[c \le T \le d]}{Prob[e \le T \le f]} = \frac{\int_c^d e^{(\mu V(s))} ds}{\int_e^f e^{(\mu V(z))} dz} \tag{9}$$

Clearly, the ratio of these probabilities does not depend on the continuous choice set [a, b].

3.1.2.2 Specification

The specification of a continuous logit model begins with the definition of the interval $[a, b]$. This is the continuous choice set. For the departure time choice of a trip and the arrival time choice of a trip for the entire day, this interval is simply defined as [0, 24]. For the covariates, the continuous logit model requires continuously time-varying covariates. These covariates may be smooth functions, continuous functions, and even step functions. Lemp and Kockelman (2010), Lemp (2009), Popuri, et al. (2008), and Zeid, et al. (2006) describe a methodology for obtaining travel time and travel time variability as time-varying functions through the use of regression models (Lemp and Kockelman, 2010, Lemp, 2009, Popuri, et al., 2008, Zeid, et al., 2006). These regression models are estimated using the Ordinary Least Squares (OLS). In addition, socio-economic covariates may also be included in the specification of the continuous logit model.

For the time-of-day choice modeling of tours, Ben-Akiva and Abou-Zeid (2013) introduced a specification for the $V\ (\cdot)$ function (systematic utility function) that is summarized as follows (Ben-Akiva and Abou-Zeid, 2013),

$$V(t_a, t_d) = V^a(t_a) + V^d(t_d) + V^{dur}(t_d - t_a), \tag{10}$$

where t_a and t_d are the arrival time choice to the activity and the departure time choice from the same activity, respectively. V^a , V^d, and V^{dur} are the systematic utility functions for arrival, departure, and duration, respectively.

They defined the following components for $V^d(t_d)$, which are the same components for $V^a(t_a)$:

$$V^d(t_d) = \sum_{r=1}^{R} z_r s^d(t_d) + \beta_{tt} TT(t_d) + \cdots \tag{11}$$

$$s^d(t_d) = \sum_{k=1}^{K} \left[\beta_{2k}^d \sin\left(\frac{2k\pi t_d}{24}\right) + \beta_{2k-1}^d \cos\left(\frac{2k\pi t_d}{24}\right)\right] \tag{12}$$

$s^d(t_d)$ is a trigonometric function that captures the cyclicality of the time-of-day choices for the 24-hour day (The utility value should be equal in $t = 0$ and $t = 24$). It also makes the utility function flexible and enables it to show various trends over the day. The β^d are parameters to be estimated. K indicates the number of frequencies used in s^d, usually selected based on cross validation (see Friedman, et al. (2001) for more information about cross validation (Friedman, et al., 2001)) after trying different values and comparing the goodness-of-fit results. Increasing K makes the function more flexible. Figure 1 shows how the trigonometric function can take different forms, with $K = 4$. Also, the $s^d(t_d)$ function is interacted with a vector of socioeconomics covariates (z) to capture the observed heterogeneity of the travelers' time-of-day preferences. R is the length of the socioeconomic covariates vector (including intercept). Other variables may be added, such as the travel time function ($\beta_{tt} TT(t_d)$, where β_{tt} is a parameter to be estimated, based on the methodology of previously mentioned literature (Lemp and Kockelman, 2010, Lemp, 2009, Popuri, et al., 2008, Zeid, et al., 2006)).

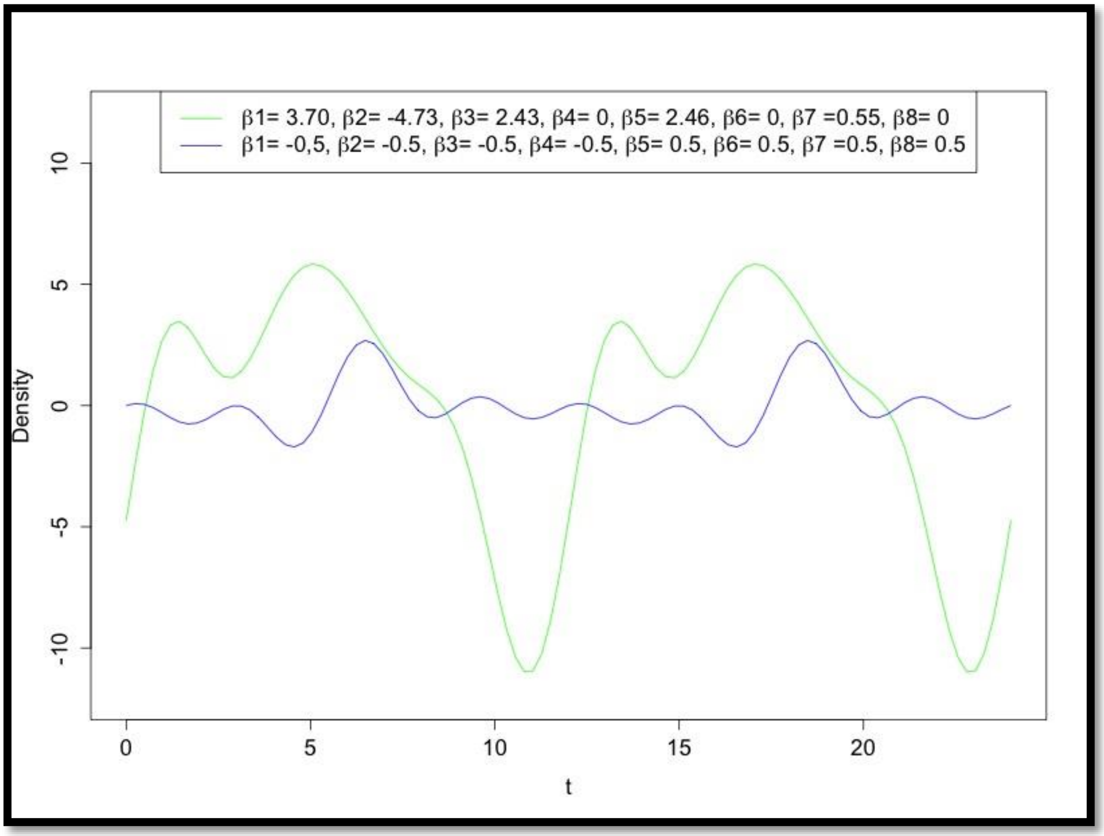


Figure 1. Examples of trigonometric functions with $K = 4$

For this chapter, I focus on the time-of-day choice modeling of a trip instead of a tour, and therefore adopt the systematic utility function of Ben-Akiva and Abou-Zeid (2013), such that $V\ (t_d\ ) = V^d\ (t_d)$ for departure time choice modeling and $V\ (t_a\ ) = V^a\ (t_a)$ for arrival time choice modeling (Ben-Akiva and Abou-Zeid, 2013). The term $V^{dur}\ (t_d - t_a)$ is not used for time-of-day choice modeling of trips. Readers are referred to Ben-Akiva and Abou-Zeid (2013) for details (Ben-Akiva and Abou-Zeid, 2013).

3.1.2.3 Identification

The identification of the continuous logit model follows the standard practice of choice models as described in Ben-Akiva and Lerman (1985) and Train (2009) (Ben-Akiva and Lerman, 1985, Train, 2009). This practice as applied to the continuous logit is summarized as follows:

- No constant functions: Constant functions, or mathematically, $f\ (t)\ = k$, in the systematic utility function $V\ (t)$ makes the coefficients associated with them unidentifiable. Also, socioeconomic variables cannot be added directly to $V\ (t)$ as these will be constant functions as well. They should be interacted with other time-varying functions, or at minimum, introduced as step functions.
- The scale must be normalized: The μ term cannot be identified for linear functional forms of $V\ (t)$, and thus it must be set to an arbitrary value. Typically, the μ is set to 1. This also holds when multiplying the systematic utility function $V\ (t)$ by a positive constant k that must be estimated.

3.2 APPROXIMATING THE CONTINUOUS LOGIT FOR TIME-OF-DAY CHOICE MODELING

In this section, the continuous logit is shown to be the originator of a variety of discrete and continuous choice models. These models are derived from the application of simple numerical integration rules (trapezoidal, midpoint and Simpson), and thus they are approximations to the continuous logit model. Furthermore, these derived models are discussed in the context of time-of-day choice modeling.

3.2.1 APPROXIMATIONS IN THE FORM OF DISCRETE CHOICE MODELS

These models are obtained by partitioning the set of integration $[a, b]$ of the continuous logit into intervals, typically of equal length, but unequal length is also acceptable. Mathematically, divide $[a, b]$ into K intervals of length Δ_k where k is the index of an interval defined as $k = 1, \dots, K$. For intervals of equal length, $\Delta = \frac{b-a}{K}$. The endpoints of the interval are as follows: $\delta_0 = a$, $\delta_1 = a + \Delta_1$, $\delta_2 = a + \Delta_1 + \Delta_2$, and so on until $\delta_K = b$. For intervals of equal length, $\delta_k = a + k\Delta$.

Now, I proceed to compute the probability of $t \in [\delta_{k-1} \leq T \leq \delta_k]$ using the partitions of the set of integration,

$$Prob\big[T \in [\delta_{k-1}, \delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] = \frac{\int_{\delta_{k-1}}^{\delta_k} e^{\mu V(s)} ds}{\sum_{h=1}^{K} \int_{\delta_{h-1}}^{\delta_h} e^{\mu V(z)} dz} \qquad (13)$$

At this point, I have simply applied the definition of probability for the continuous logit, and the properties of the integral, and thus I have not introduced any approximations to the model. The approximation occurs when the integrals over the intervals are replaced with numerical integration rules. I consider three widely known and simple numerical integration rules: trapezoidal, midpoint, and Simpson (Burden and Faires, 2001). These rules are discussed subsequently. Furthermore, these models are considered discrete choice models because they approximate the probability (i.e. CDF) of t being in an interval and produce a probability mass function. Therefore, these models predict the time-of-day choices as aggregate intervals.

3.2.1.1 Midpoint-based Multinomial Logit

For this approximation, I replace the integral of each interval in equation (13) by its approximation using the midpoint rule. The midpoint rule approximates a definite integral as follows: $\int_c^d f(s)ds \approx f(\frac{c+d}{2})(d-c)$. Thus,

$$Prob\big[T \in [\delta_{k-1}, \delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] = \frac{e^{\mu V\left(\frac{\delta_{k-1}+\delta_k}{2}\right)}\Delta_k}{\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right)}\Delta_h} \quad (14)$$

For the case of equal length Δ, the probability of $t \in [\delta_{k-1}, \delta_k]$ is,

$$Prob\big[T \in [\delta_{k-1}, \delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] = \frac{e^{\mu V\left(\frac{\delta_{k-1}+\delta_k}{2}\right)}\Delta}{\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right)}\Delta} = \frac{e^{\mu V\left(\frac{\delta_{k-1}+\delta_k}{2}\right)}}{\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right)}} \quad (15)$$

This equation (15) is the discretized multinomial logit, which is widely used in the literature for time-of-day choice modeling (Ben-Akiva and Abou-Zeid, 2013, Popuri, et al., 2008, Zeid, et al., 2006). Moreover, Ben-Akiva and Abou-Zeid 2013 addresses the issue of intervals of unequal length by recognizing the link between the discretized multinomial logit and the continuous logit, and thus rewriting equation (14) as follows (Ben-Akiva and Abou-Zeid, 2013),

$$Prob\big[T \in [\delta_{k-1}, \delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] = \frac{e^{\mu V\left(\frac{\delta_{k-1}+\delta_k}{2}\right)+\ln(\Delta_k)}}{\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right)+\ln(\Delta_h)}} \quad (16)$$

Equation (16) is derived by applying the mean value theorem (Binmore, 1982), and evaluating the systematic utilities at what is called the "mid-point". The reason for writing mid-point in quotes is that the function must be evaluated at a value m such that $\int_c^d f(s)ds = f(m)(d-c)$. Clearly, m may not necessarily be the midpoint, and thus it may be that $f(m) \neq f(\frac{c+d}{2})$. Therefore, equation (16) is an approximation to the continuous logit. In addition, Ben-Akiva and Lerman (1985) refers to the term $\ln(\Delta_k)$ as a size variable (Ben-Akiva and Lerman, 1985).

3.2.1.2 Trapezoidal-based Multinomial Logit

For this approximation, the integral of each interval in equation (13) is replaced using the trapezoidal rule. The trapezoidal rule approximates a definite integral as follows: $\int_c^d f(s)ds \approx \big(f(c)+f(d)\big)\,(\frac{d-c}{2})$. Thus,

$$Prob\big[T \in [\delta_{k-1},\delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] = \frac{\left(e^{\mu V(\delta_{k-1})}+e^{\mu V(\delta_k)}\right)(\frac{\Delta_k}{2})}{\sum_{h=1}^{K}\left(e^{\mu V(\delta_{h-1})}+e^{\mu V(\delta_h)}\right)(\frac{\Delta_h}{2})} \qquad (17)$$

Equation (17) may also be written as follows,

$$Prob\big[T \in [\delta_{k-1},\delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] = \frac{e^{\mu V(\delta_{k-1})+\ln(\frac{\Delta_k}{2})}+e^{\mu V(\delta_k)+\ln\left(\frac{\Delta_k}{2}\right)}}{\sum_{h=1}^{K} e^{\mu V(\delta_{h-1})+\ln(\frac{\Delta_h}{2})}+e^{\mu V(\delta_h)+\ln\left(\frac{\Delta_h}{2}\right)}}$$

(18)

For the case of equal length Δ, the probability of $t \in [\delta_{k-1},\delta_k]$ is,

$$Prob\big[T \in [\delta_{k-1},\delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] = \frac{e^{\mu V(\delta_{k-1})}+e^{\mu V(\delta_k)}}{\sum_{h=1}^{K} e^{\mu V(\delta_{h-1})}+e^{\mu V(\delta_h)}} \qquad (19)$$

3.2.1.3 Simpson-based Multinomial Logit

For this approximation, the integral of each interval in equation (13) is replaced using the Simpson rule. The Simpson rule approximates a definite integral as follows: $\int_c^d f(s)ds \approx \left(f(c) + 4f\left(\frac{c+d}{2}\right) + f(d)\right)(\frac{d-c}{6})$ Thus,

$$Prob\big[T \in [\delta_{k-1}, \delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] =$$

$$\frac{\left(e^{\mu V(\delta_{k-1})} + 4e^{\mu V\left(\frac{\delta_{k-1}+\delta_k}{2}\right)} + e^{\mu V(\delta_k)}\right)(\frac{\Delta_k}{6})}{\sum_{h=1}^{K}\left(e^{\mu V(\delta_{h-1})} + 4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right)} + e^{\mu V(\delta_h)}\right)(\frac{\Delta_h}{6})} \tag{20}$$

Equation (20) may also be written as follows,

$$Prob\big[T \in [\delta_{k-1}, \delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] =$$

$$\frac{e^{\mu V(\delta_{k-1}) + \ln(\frac{\Delta_k}{6})} + 4e^{\mu V\left(\frac{\delta_{k-1}+\delta_k}{2}\right) + \ln(\frac{\Delta_k}{6})} + e^{\mu V(\delta_k) + \ln(\frac{\Delta_k}{6})}}{\sum_{h=1}^{K} e^{\mu V(\delta_{h-1}) + \ln(\frac{\Delta_h}{6})} + 4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right) + \ln(\frac{\Delta_h}{6})} + e^{\mu V(\delta_h) + \ln(\frac{\Delta_h}{6})}} \tag{21}$$

For the case of equal length Δ, the probability of $t \in [\delta_{k-1}, \delta_k]$ is,

$$Prob\big[T \in [\delta_{k-1}, \delta_k]\big] = Prob[\delta_{k-1} \leq T \leq \delta_k] =$$

$$\frac{e^{\mu V(\delta_{k-1})} + 4e^{\mu V\left(\frac{\delta_{k-1}+\delta_k}{2}\right)} + e^{\mu V(\delta_k)}}{\sum_{h=1}^{K} (e^{\mu V(\delta_{h-1})} + 4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right)} + e^{\mu V(\delta_h)})} \tag{22}$$

3.2.2 APPROXIMATIONS IN THE FORM OF CONTINUOUS CHOICE MODELS

The models in this section are derived using the same partitioning logic and notation explained in the previous section. The key difference is that I proceed to compute the probability density function (PDF) using the partitions of the set of integration as follows,

$$f(t) = \frac{e^{\mu V(t)}}{\sum_{h=1}^{K} \int_{\delta_{h-1}}^{\delta_h} e^{\mu V(z)} dz} \tag{23}$$

This equation (23) applies the properties of the integral, and it does not introduce any approximations to the model. Similar to the previous sub-section, the approximation occurs when the integrals over the intervals are replaced using numerical integration rules. I consider the following rules as previously mentioned: trapezoidal, midpoint, and Simpson (Burden and Faires, 2001). Moreover, these models are considered *continuous choice models* because they approximate the probability density function (i.e. PDF) instead of the probability distribution (i.e. CDF). Therefore, these models preserve the continuous feature of the continuous logit model and are able to produce point predictions of the time-of-day choices. Also, the *size variables* in these models do not cancel even for the cases with equal length Δ.

3.2.2.1 Midpoint-based Continuous Logit

For this model, the integral of each interval in equation (23) is replaced by its approximation using the midpoint rule. The midpoint rule approximates a definite integral as follows: $\int_c^d f(s) ds \approx f(\frac{c+d}{2})(d-c)$. Thus,

$$f(t) = \frac{e^{\mu V(t)}}{\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right)} \Delta_h} = \frac{e^{\mu V(t)}}{\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2}\right) + \ln(\Delta_h)}} \tag{24}$$

3.2.2.2 Trapezoidal-based Continuous Logit

In this approximation, the integral of each interval in equation 23 is replaced by its trapezoidal rule equivalent. The trapezoidal rule for a definite integral is stated as: $\int_c^d f(s)ds \approx \big(f(c) + f(d)\big)\,(\frac{d-c}{2})$. Thus,

$$f(t) = \frac{e^{\mu V(t)}}{\sum_{h=1}^{K} e^{\mu V(\delta_{h-1}) + \ln\left(\frac{\Delta_h}{2}\right)} + e^{\mu V(\delta_h) + \ln\left(\frac{\Delta_h}{2}\right)}} \tag{25}$$

1.2.2.3 Simpson-based Continuous Logit

For this approximation, the integral of each interval in equation (23) is approximated using the Simpson rule. The Simpson rule for a definite integral is as follows: $\int_c^d f(s)ds \approx \left(f(c) + 4f\left(\frac{c+d}{2}\right) + f(d)\right)(\frac{d-c}{6})$. Thus,

$$f(t) = \frac{e^{\mu V(t)}}{\sum_{h=1}^{K} e^{\mu V(\delta_{h-1}) + \ln\left(\frac{\Delta_h}{6}\right)} + 4e^{\mu V\left(\frac{\delta_{h-1} + \delta_h}{2}\right) + \ln\left(\frac{\Delta_h}{6}\right)} + e^{\mu V(\delta_h) + \ln\left(\frac{\Delta_h}{6}\right)}} \tag{26}$$

3.3 MONTE CARLO EXPERIMENTS OF APPROXIMATIONS

The derived models are studied using Monte Carlo experiments (Hammersley, 2013, Johnston and DiNardo, 1972). I examine the following research problems: the approximation error of the derived models; and the bias of the MLE due to the approximation error. I discuss these research problems in detail in the subsequent subsections. Here, I present the common elements of the experiments.

I assume the following systematic utility function for all the simulated data sets,

$$V(t;\beta) = \beta_{tt}TT(t) + \sum_{h=1}^{4}(\beta_{2h}\sin(\frac{2h\pi}{24}t) + \beta_{2h-1}cos\left(\frac{2h\pi}{24}\ t\right)) \quad (27)$$

This systematic utility function follows the specification of Ben-Akiva and Abou-Zeid (2013) for time-of-day choice modeling for a trip (Ben-Akiva and Abou-Zeid, 2013). The true parameters were chosen from the estimates of the continuous logit model from the table (5.2) in Lemp (2009), with the exception of the parameter of the travel time function $TT(t)$ (Lemp, 2009). This parameter was set to 0.50 because the 95% confidence interval of the travel time estimate in the model of (Lemp, 2009)contains 0. The continuous logit of Lemp (2009) is a model of the departure time choice for trips from home to work. I also assume that the synthetic travelers are departing from home to work. The values of the true parameters are presented in ***Table 1***.

The travel time function $TT(t)$ follows the travel speed regression model and uses the estimates from the table (1) in Popuri, et al. (2008) (Popuri, et al., 2008). The dependent variable $m(t)$ of the regression model of Popuri, et al. (2008) is $m(t) = \ln(\frac{reported\ speed}{free\ flow\ speed})$. Therefore, the predictions for a given time t are the multiplier $m(t)$ for the free flow speed of the origin-destination pair of a traveler. In addition, the covariates are interacted with the delay, which is defined as $1 - \frac{peak\ network\ speed}{free\ flow\ speed}$. The specification of the travel speed regression model used for the travel time function $TT(t)$ is

$$m(t) = \alpha_0 + \alpha_{dist}\, ln(distance) + \alpha_{del}\, delay + \sum_{h=1}^{4}\left(\alpha_{2h}\, e^{hsin\left(\frac{2\pi}{24}t\right)}\, delay + \alpha_{2h-1}\, e^{hcos\left(\frac{2\pi}{24}t\right)}\, delay\right) + \sum_{h=1}^{4}(\alpha_{2h+8}\, e^{hsin\left(\frac{4\pi}{24}t\right)}\, delay + \alpha_{2h+7}\, e^{hcos\left(\frac{4\pi}{24}t\right)}\, delay) \tag{28}$$

The travel time function is defined as

$$TT(t) = \frac{distance}{free\ flow\ speed \times e^{m(t)}} \tag{29}$$

I assume that the travel distance (unit in miles) is distributed as a Gamma distribution with a shape parameter of 2, and a scale parameter of 4. The Gamma distribution was chosen as studies have shown that it is a good fit for travel distances (Ben-Akiva and Watanatada, 1981). Also, I assume that the free flow speed is distributed as a discrete uniform distribution with the following outcomes: 40, 50, 60, 70, 80, and 90. The unit used is miles per hour (MPH). For the peak network speed, I assume also a discrete uniform distribution with the following outcomes: 15, 20, 25, 30, 35, and 40. The unit is miles per hour (MPH). The true parameters for the travel time function $TT(t)$ are also presented in ***Table 1***.

Table 1. True parameters for generating the synthetic data

Systematic Utility	
μ	1.00
β_{tt}	-0.50
β_1	-4.73
β_2	3.70
β_3	-0.51
β_4	2.43
β_5	2.40
β_6	2.46
β_7	1.34
β_8	0.55
Travel time $TT(t)$	
α_0	0.0317
α_{dist}	0.0072
α_{del}	-1.4865
α_1	7.0780
α_2	-11.2967
α_3	-5.6571
α_4	12.5130
α_5	1.8769
α_6	-5.6582
α_7	-0.2359
α_8	0.9182
α_9	2.3598
α_{10}	-1.9606
α_{11}	-0.7701
α_{12}	1.7074
α_{13}	0.1102
α_{14}	-0.5433
α_{15}	0.0018
α_{16}	0.0630

I do not estimate the parameters of the travel time function $TT(t)$. They are assumed as constants and they are incorporated into the travel time function $TT(t)$, which is a covariate of the systematic utility function $V(t)$. This is the practice followed for time-of-day choice modeling in the research literature (Ben-Akiva and Abou-Zeid, 2013, Lemp, 2009, Popuri, et al., 2008, Zeid, et al., 2006). In addition, I only simulate cross-sectional data sets. In other words, one observation is one synthetic traveler. For each synthetic traveler, I simulate (i.e. draw from the respective distribution) the travel distance, peak network speed, and free flow speed. The departure time choices of the synthetic travelers are generated using the *acceptance-rejection* method (Casella and Berger, 2002) as applied to the probability density function of the continuous logit model evaluated at the true parameters in Table 1. Furthermore, the integral of the continuous logit model is computed numerically using an adaptive quadrature based on the 21-point Gauss-Kronrod quadrature. This quadrature is discussed in Piessens, et al. (2012) (Piessens, et al., 2012), and available in the R statistical package (RCore, 2012). I set the error tolerance of the quadrature to an order of magnitude of 10^{-10}. Lastly, the Monte Carlo experiments are coded in R (RCore, 2012).

3.3.1 QUANTIFYING THE APPROXIMATION ERROR

In this subsection, I present numerical results from Monte Carlo experiments to illustrate the magnitude of the approximation error in calculating the probability of $t \in [\delta_{k-1}, \delta_k]$ for the derived discrete choice models, and the probability density function at t for the derived continuous choice models with respect to the continuous logit model.

I define the approximation error for $t \in [\delta_{k-1}, \delta_k]$ for the derived discrete choice models as follows,

$$\xi_{t \in [\delta_{k-1}, \delta_k]} = \left| \frac{\int_{\delta_{k-1}}^{\delta_k} e^{(\mu V(s;\beta))} ds}{\int_a^b e^{(\mu V(z;\beta))} dz} - \widetilde{Prob}(t \in [\delta_{k-1}, \delta_k]; \beta) \right| \tag{30}$$

For the derived continuous choice models, the approximation error for t is defined as follows,

$$\xi(t) = \left| \frac{e^{\mu V(t;\beta)}}{\int_a^b e^{\mu V(z;\beta)} dz} - \tilde{f}(t;\beta) \right| \tag{31}$$

For both $\xi_{t \in [\delta_{k-1}, \delta_k]}$ and $\xi(t)$, β is the vector of values of the true parameters in Table 1. The function $|.|$ is the absolute value function (Binmore, 1982). The term $\widetilde{Prob}(t \in [\delta_{k-1}, \delta_k]; \beta)$ represents one of the approximations in the form of the derived discrete choice models. The term $\tilde{f}(t;\beta)$ is one of the approximations in the form of the derived continuous choice models.

For these numerical experiments, I simulate a data set of 1000 observations using the acceptance-rejection method applied to the continuous logit model using an adaptive quadrature based on the 21-point Gauss-Kronrod quadrature with an error tolerance of 10^{-10}. I generate the dataset using the true parameters in Table 1. For the approximations in the form of discrete choice models, I use equation (30) and evaluate both the approximations and the continuous logit at the true parameter vector β. I also follow the same procedure for the approximations in the form of the continuous choice models, except that I use equation (31) to compute the approximation error. The reason I evaluate both the approximations and the continuous logit model at the true parameters is to separate the approximation error from any other sources of discrepancy such as statistical bias. I use the same data set to compute all the approximation errors.

The probability density function for the continuous logit model is presented in Figure 2. The median and interquartile range are computed for the probability density function across the 1000 observations. Each observation/synthetic agent has a different probability density function because I simulate for each observation/synthetic agent his/her travel distance, free flow speed, and peak network speed. I do not simulate any heterogeneity (i.e. variation in preferences of parameter across a population density function and/or across socioeconomic variables) in the generated dataset.

The approximation error $\xi_{t\in[\delta_{k-1},\delta_k]}$ of the discrete choice models for the simulated data set is presented in Figure 3, and the approximation error $\xi(t)$ for the continuous choice models is presented in Figure 4. For a given derived model, the 90th percentile of the approximation error across the simulated data for a t (i.e. continuous choice models) and for a $t \in [\delta_{k-1}, \delta_k]$ (i.e discrete choice model) is used to signify that 90% of the errors are below the value shown in the figures. The figures indicate that the approximation errors for the same number of partition intervals (e.g. 30 minutes intervals) are several orders of magnitude greater in the discrete choice models compared to the continuous choice models. For example, the midpoint-based continuous logit has approximation errors in the order of 10^{-7} for 30 minutes intervals in contrast to the midpoint-based multinomial logit also for 30 minutes intervals that has approximation errors in the order of 10^{-3}. This trend is observed between the other discrete choice models and their continuous choice model counterparts. Another result, which is expected (Burden and Faires, 2001), is that the Simpson-based continuous logit and Simpson-based multinomial logit models have the smallest approximation errors compared to the other models in their respective classes. The Simpson-based continuous logit has the smallest approximation compared to all the models. The Simpson-based multinomial logit has approximation errors in the order of 10^{-4}, and the Simpson-based continuous logit has approximation errors in the order of 10^{-7}. In addition, as expected (Burden and Faires, 2001), the approximation errors decrease as the number of partition intervals increases. Also, as expected (Burden and Faires, 2001), the greatest approximation errors occur at the points of the probability density function where its curvature changes rapidly.

In summary, the continuous choice models not only offer the ability to preserve the continuous dependent variable in its correct form, but they also offer much less approximation error for the same number of partition intervals than their discrete choice model counterparts based on these numerical experiments.

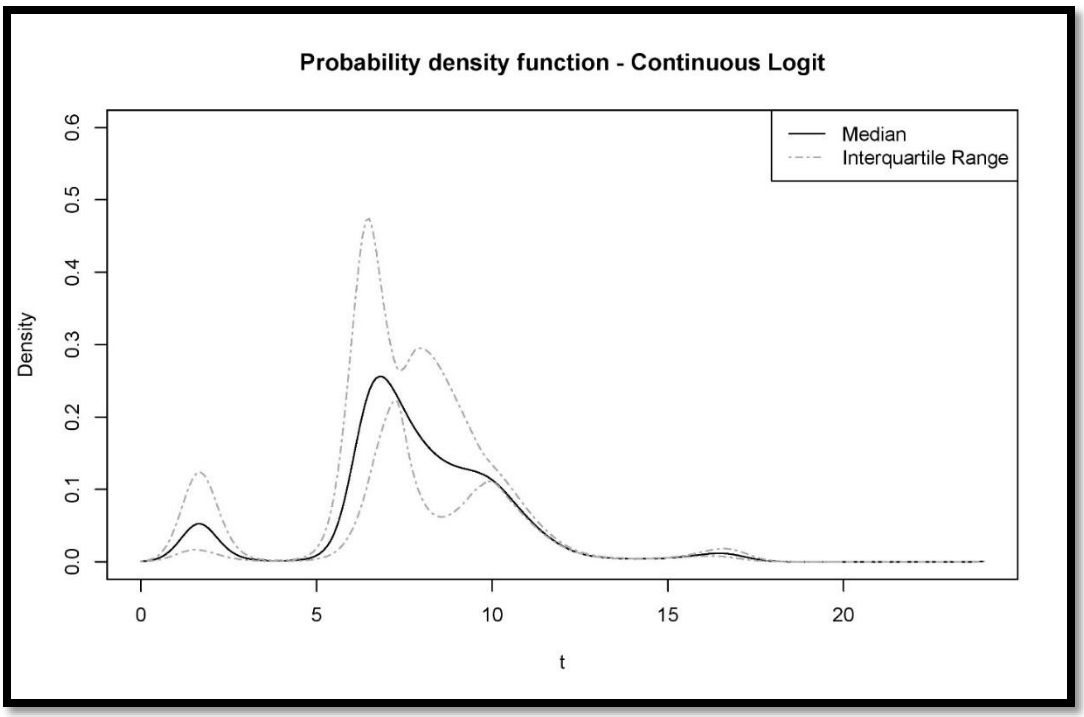
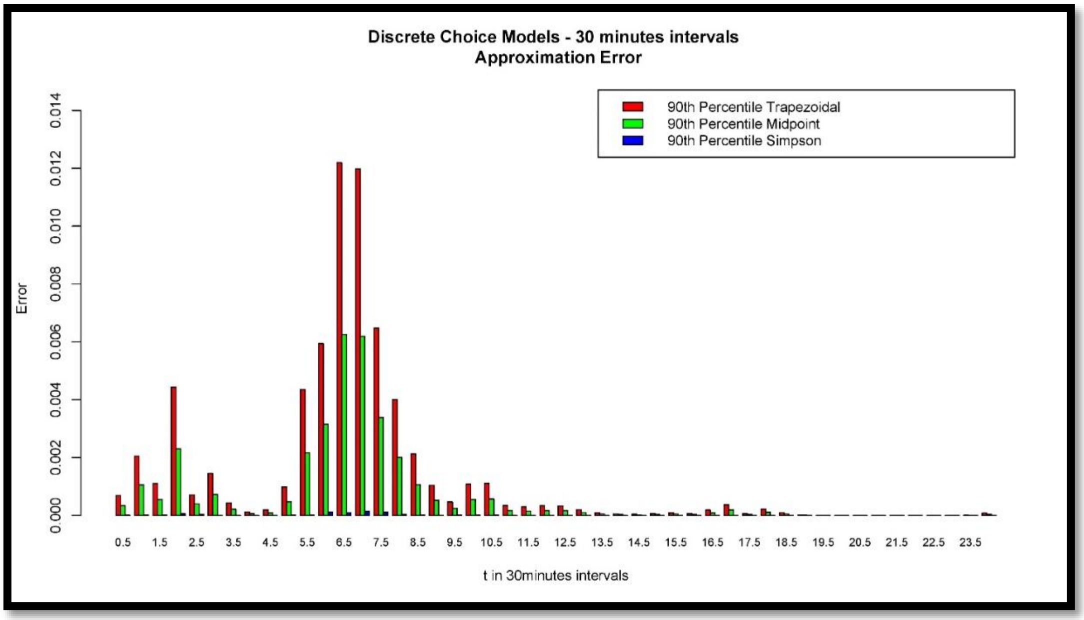


Figure 2. Probability density function of the continuous logit for the simulated data



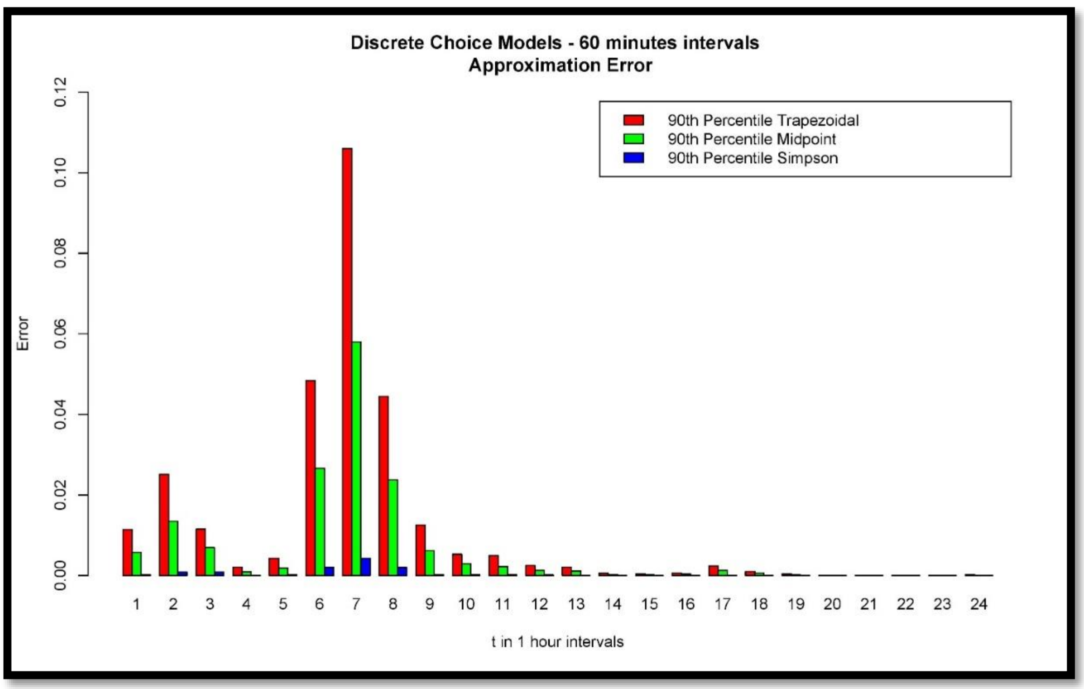


Figure 3. Approximating error $\xi_{t\in[\delta_{k-1},\delta_k]}$ for the simulated data set

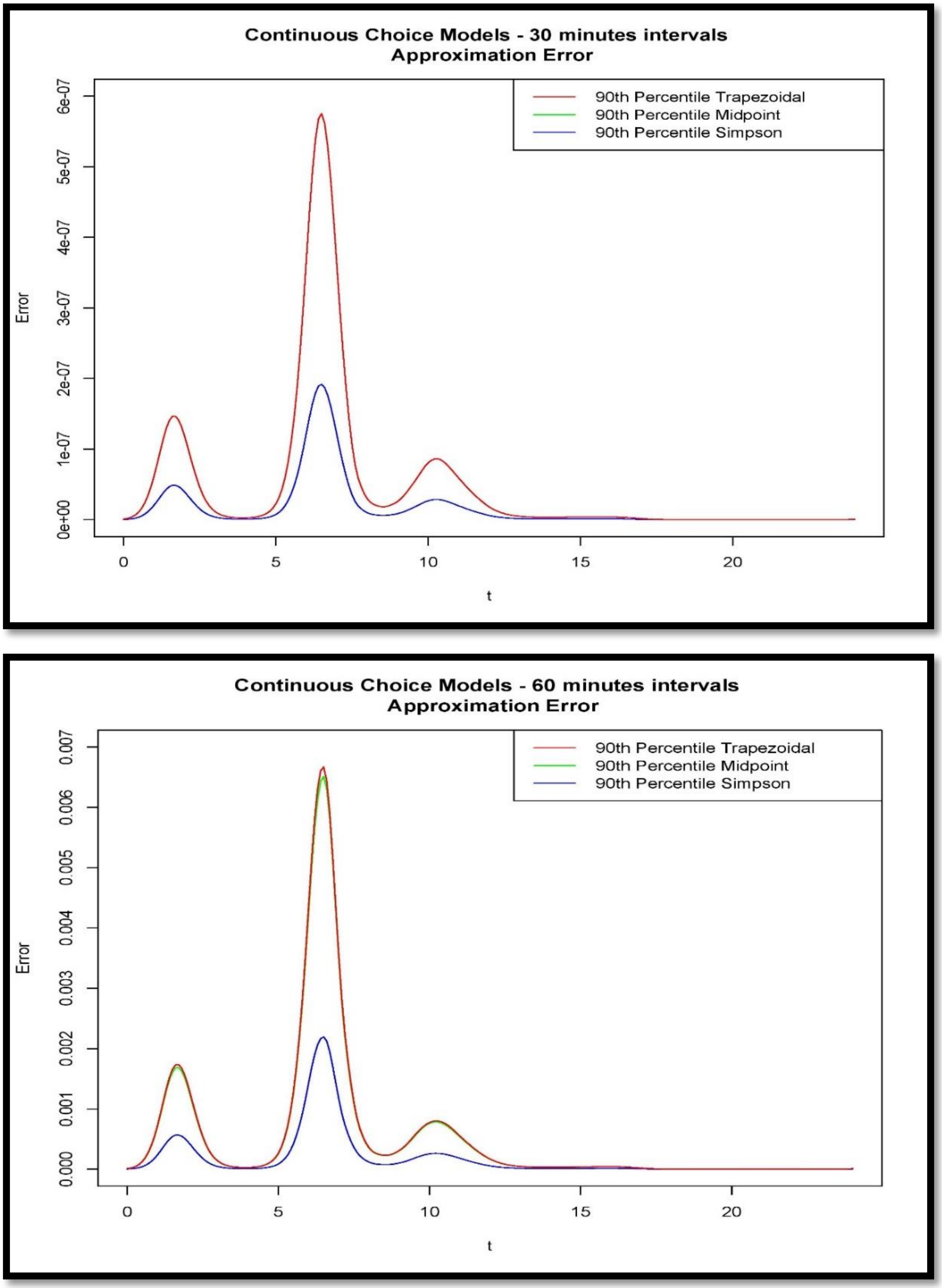


Figure 4. Approximating error $\xi(t)$ for the simulated data set

3.3.2 QUANTIFYING THE BIAS OF THE MAXIMUM LIKELIHOOD ESTIMATOR DUE TO THE APPROXIMATION ERROR

In this sub-section, I present numerical results about the bias of the MLE due to the approximation error as defined in the previous sub-section. In principle, there are four possible cases to consider for the presence of bias due to the approximation error: (1) The approximation error introduces bias, but this bias becomes smaller as the approximation error becomes smaller; (2) The approximation error introduces bias, but this bias becomes smaller as the sample size increases; (3) a combination of both cases 1 and 2; (4) The approximation error does not introduce any bias. Readers are referred to Casella and Berger (2002), Cramer (1986), and Johnston and DiNardo (1972) for details about sampling theory, MLE, and Monte Carlo experiments (Casella and Berger, 2002, Cramer, 1986, Johnston and DiNardo, 1972).

For these numerical experiments, I simulate one dataset with a sample size of 500 observations for a total of 100 replications. The 100 replications allow constructing the sampling distributions of the estimators. They also allow calculating the confidence intervals to control for errors of the Monte Carlo simulation. Also, it is possible to obtain the confidence intervals for $\hat{\beta} - \beta_{true}$, where β_{true} is the vector of true values of the parameters. I calculate the confidence intervals (with $\alpha = 0.025$) to test the existence of statistical bias. For each replication, I estimate the β parameters of two derived models: the midpoint-based multinomial logit, and the midpoint-based continuous logit. In addition, I estimate the β parameters for the following lengths of intervals for the partitions: 15 minutes, 30 minutes, and 60 minutes. Furthermore, I simulate the datasets using the acceptance-rejection method applied to the continuous logit model using an adaptive quadrature based on the 21-point Gauss-Kronrod quadrature with an error tolerance of 10^{-5}. I generate these data sets using the true parameters in ***Table 2***.

The numerical results are presented in ***Table 2***. At each cell, the top value shows $E(\hat{\beta} - \beta_{true})$ (measure of bias) and the bottom values shows the 95% confidence interval for $\hat{\beta} - \beta_{true}$. The results does not show any evidence for the presence of statistical bias in the two models. This is because the Monte Carlo estimate of the confidence intervals for $\hat{\beta} - \beta_{true}$ contain 0 for almost all of the β parameters. For the midpoint-based multinomial logit, the results suggest that the confidence interval becomes smaller as the approximation error becomes smaller. For the midpoint-based continuous logit, the results of the statistical bias do not seem to follow such a clear cut trend. In fact, the midpoint-based continuous logit with Δ of 60min has the biggest confidence interval, but the smallest confidence interval is for Δ of 30min. Furthermore, the results show that the confidence intervals for the continuous logit models are smaller than their multinomial logit counterparts. These are preliminary results, and further research need to be done in this area.

Table 2. Monte Carlo estimates of the statistical bias of MLE due to the approximation error

		Midpoint-based Multinomial Logit			Midpoint-based Continuous Logit		
$Parameter$	True Parameter	15 min interval	30 min interval	60 min interval	15 min interval	30 min interval	60 min interval
$E(\widehat{\beta_{tt}} - \beta_{tt})$	-0.500	0.002	0.011	0.049	-0.001	-0.002	0.002
CI	-	[-0.017, 0.020]	[-0.011, 0.033]	[0.007, 0.083]	[-0.017, 0.013]	[-0.018, 0.009]	[-0.011, 0.022]
$E(\widehat{\beta_1} - \beta_1)$	-4.73	0.125	0.139	0.696	-0.185	-0.070	-0.071
CI	-	[-1.492, 0.629]	[-0.590, 0.925]	[-0.789, 1.594]	[-1.587, 0.505]	[-1.135, 0.786]	[-1.188, 0.681]
$E(\widehat{\beta_2} - \beta_2)$	3.70	0.152	-0.065	-0.397	0.195	0.071	-0.045
CI	-	[-0.704, 1.386]	[-0.932, 0.799]	[-1.228, 1.074]	[-0.637, 1.402]	[-0.833, 1.086]	[-0.816, 1.332]
$E(\widehat{\beta_3} - \beta_3)$	-0.510	0.012	0.027	0.080	0.013	0.008	-0.024
CI	-	[-0.183, 0.268]	[-0.118, 0.334]	[-0.234, 0.337]	[-0.219, 0.209]	[-0.157, 0.291]	[-0.278, 0.143]
$E(\widehat{\beta_4} - \beta_4)$	2.43	0.150	-0.094	-0.518	0.199	0.071	-0.061
CI	-	[-0.721, 1.408]	[-1.068, 0.713]	[-1.514, 1.011]	[-0.637, 1.403]	[-0.939, 1.044]	[-0.842, 1.129]
$E(\widehat{\beta_5} - \beta_5)$	2.40	0.070	-0.057	-0.337	0.103	0.041	-0.083
CI	-	[-0.468, 0.745]	[-0.547, 0.421]	[-0.855, 0.521]	[-0.486, 0.819]	[-0.509, 0.572]	[-0.660, 0.454]
$E(\widehat{\beta_6} - \beta_6)$	2.46	0.051	-0.091	-0.406	0.080	0.024	-0.040
CI	-	[-0.299, 0.713]	[-0.605, 0.272]	[-0.887, 0.208]	[-0.276, 0.761]	[-0.451, 0.397]	[-0.397, 0.575]
$E(\widehat{\beta_7} - \beta_7)$	1.34	0.033	-0.045	-0.226	0.051	0.015	-0.054
CI	-	[-0.235, 0.410]	[-0.348, 0.183]	[-0.616, 0.160]	[-0.234, 0.455]	[-0.280, 0.304]	[-0.300, 0.295]
$E(\widehat{\beta_8} - \beta_8)$	0.550	-0.016	-0.028	-0.092	-0.012	-0.007	0.000
CI	-	[-0.197, 0.019]	[-0.222, 0.060]	[-0.268, 061]	[-0.240, 0.113]	[-0.174, 0.073]	[-0.129, 0.101]

3.4 ANALYTICAL BOUNDS FOR THE APPROXIMATION ERROR OF THE DERIVED MODELS

In this section, I present derivations of analytical bounds for the *relative error* of the continuous choice models: the trapezoidal-based continuous logit, the midpoint-based continuous logit, and the Simpson-based continuous logit. I follow the notation and equations from the previous sections.

The relative error is defined for the continuous choice models as

$$\Xi(t) = \left|\frac{f(t;\beta) - \tilde{f}(t;\beta)}{f(t;\beta)}\right| \quad (32)$$

The relative error $\Xi(t)$ is the absolute error $\xi(t)$ relative to the exact value. Intuitively, it provides an idea of how negligible the absolute error is. I present the analytical bounds derivations for the relative error of the continuous choice models in the following theorems,

Theorem 1: Error bound for the midpoint-based continuous logit

Suppose that $f(t;\beta)$ is the probability density function of the continuous logit model, defined and continuous on the interval $[a.b]$, and that $\tilde{f}(t;\beta)$ is the probability density function of the midpoint-based continuous logit model, also defined and continuous on the interval $[a,b]$. Also, suppose that the first and the second derivatives of the $e^{\mu V(t;\beta)}$ function are defined and continuous on the interval $[a,b]$. Then

$$\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta M_2(b-a)}{24\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}} \leq \frac{\Delta^2 M_2}{24 e^{\mu V(c;\beta)}} \tag{33}$$

Where, $M_2 = Max_{t\in[a,b]}\left|\frac{d^2(e^{\mu V(t;\beta)})}{dt^2}\right|$, and also $V(c;\beta)$ is a constant function such that $\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)} \geq \sum_{h=1}^{K} e^{\mu V(c;\beta)}$.

Theorem 1 proof

Notice that, $f(t) = \frac{e^{\mu V(t,\beta)}}{\sum_{h=1}^{K}\int_{\delta_{h-1}}^{\delta_h} e^{\mu V(x,\beta)}dx}$ and that $\tilde{f}(t) = \frac{e^{\mu V(t,\beta)}}{\Delta\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2},\beta\right)}}$. Also, notice that $(t;\beta) - \tilde{f}(t;\beta) = f(t;\beta)\frac{\Delta\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)} - \sum_{h=1}^{K}\int_{\delta_{h-1}}^{\delta_h} e^{\mu V(x;\beta)}dx}{\Delta\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}}$.

Also for the midpoint rule for a function $f(x)$ and its first and second derivative defined and continuous on $[a, b]$, $\exists\ \epsilon \in (a, b)$ for which, $\int_a^b f(x)dx = \sum_{h=1}^{K} \int_{\delta_{h-1}}^{\delta_h} f(x)dx = \sum_{h=1}^{K} |\Delta f\left(\frac{\delta_{h-1}+\delta_h}{2}\right)| + \frac{1}{24}\Delta^2(b-a)(\frac{d^2 f(x)}{dx^2}\ |_{\ x=\epsilon}\ )$. Readers are referred to Burden and Faires (2001), Mathews and Fink (2004), and Süli and Mayers (2003) for details (Burden and Faires, 2001, Mathews and Fink, 2004, Süli and Mayers, 2003).

Thus, $f(t;\beta) - \tilde{f}(t;\beta) = f(t;\beta)\frac{\frac{1}{24}\Delta^2(b-a)(\frac{d^2(e^{\mu V(t;\beta)})}{dt^2}\ |_{\ t=\epsilon}\ )}{\Delta\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}}$.

By applying absolute value, substituting with M_2 and rearranging, $\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta M_2(b-a)}{24\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}}$

By substituting with $\sum_{h=1}^{K} e^{\mu V(c;\beta)} = \frac{b-a}{\Delta} e^{\mu V(c;\beta)}$,

$$\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta M_2(b-a)}{24\sum_{h=1}^{K} e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}} \leq \frac{\Delta^2 M_2}{24 e^{\mu V(c;\beta)}}\ .$$

<u>Theorem 2: Error bound for the trapezoidal-based continuous logit</u>

Suppose that $f(t;\beta)$ is the probability density function of the continuous logit model, defined and continuous on the interval $[a.b]$, and that $\tilde{f}(t;\beta)$ is the probability density function of the trapezoidal -based continuous logit model, also defined and continuous on the interval $[a, b]$. Also, suppose that the first and the second derivatives of the $e^{\mu V(t;\beta)}$ function are defined and continuous on the interval $[a, b]$. Then

$$\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta M_2(b-a)}{6\sum_{h=1}^{K}\left[e^{\mu V(\delta_{h-1};\beta)}+e^{\mu V(\delta_h;\beta)}\right]} \leq \frac{\Delta^2 M_2}{12e^{\mu V(c;\beta)}} \quad (34)$$

Where, $M_2 = Max_{t\in[a,b]}|\frac{d^2(e^{\mu V(t;\beta)})}{dt^2}|$, and also $V(c)$ is a constant function such that $\sum_{h=1}^{K}\left[e^{\mu V(\delta_{h-1};\beta)} + e^{\mu V(\delta_h;\beta)}\right] \geq \sum_{h=1}^{K} 2e^{\mu V(c;\beta)}$.

<u>Theorem 2 Proof</u>

Notice that, $f(t) = \frac{e^{\mu V(t,\beta)}}{\sum_{h=1}^{K}\int_{\delta_{h-1}}^{\delta_h} e^{\mu V(x,\beta)}dx}$ and that $\tilde{f}(t) = \frac{e^{\mu V(t,\beta)}}{\frac{\Delta}{2}\sum_{h=1}^{K}[e^{\mu V(\delta_{h-1},\beta)}+e^{\mu V(\delta_h,\beta)}]}$.

Also, notice that

$$f(t;\beta) - \tilde{f}(t;\beta) = f(t;\beta)\frac{\frac{\Delta}{2}\sum_{h=1}^{K}[e^{\mu V(\delta_{h-1};\beta)}+e^{\mu V(\delta_h;\beta)}]-\sum_{h=1}^{K}\int_{\delta_{h-1}}^{\delta_h} e^{\mu V(x;\beta)}dx}{\frac{\Delta}{2}\sum_{h=1}^{K}[e^{\mu V(\delta_{h-1};\beta)}+e^{\mu V(\delta_h;\beta)}]} .$$

Also for the trapezoidal rule for a function $f(x)$ and its first and second derivative defined and continuous on $[a,b]$, $\exists\ \epsilon \in (a,b)$ for which, $\int_a^b f(x)dx = \sum_{h=1}^{K}\int_{\delta_{h-1}}^{\delta_h} f(x)dx = \sum_{h=1}^{K}[\frac{\Delta}{2}(f(\delta_{h-1}) + f(\delta_h))] - \frac{1}{12}\Delta^2(b-a)(\frac{d^2f(x)}{dx^2}\ |_{\ x=\epsilon})$. Readers are referred to Burden and Faires (2001), Mathews and Fink (2004), and Süli and Mayers (2003) for details (Burden and Faires, 2001, Mathews and Fink, 2004, Süli and Mayers, 2003).

Thus, $f(t;\beta) - \tilde{f}(t;\beta) = f(t;\beta)\frac{\frac{1}{12}\Delta^2(b-a)(\frac{d^2(e^{\mu V(t;\beta)})}{dt^2}\ |_{\ t=\epsilon})}{\frac{\Delta}{2}\sum_{h=1}^{K}[e^{\mu V(\delta_{h-1};\beta)}+e^{\mu V(\delta_h;\beta)}]}$.

By applying absolute value, substituting with M_2 and rearranging, $\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta M_2(b-a)}{6\sum_{h=1}^{K}\left[e^{\mu V(\delta_{h-1};\beta)}+e^{\mu V(\delta_h;\beta)}\right]}$.

By substituting with $\sum_{h=1}^{K} 2e^{\mu V(c;\beta)} = \frac{b-a}{\Delta} 2e^{\mu V(c;\beta)}$,

$$\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta M_2(b-a)}{6\sum_{h=1}^{K}\left[e^{\mu V(\delta_{h-1};\beta)}+e^{\mu V(\delta_h;\beta)}\right]} \leq \frac{\Delta^2 M_2}{12e^{\mu V(c;\beta)}} .$$

Theorem 3: Error bound for the Simpson-based continuous logit

Suppose that $f(t;\beta)$ is the probability density function of the continuous logit model, defined and continuous on the interval $[a.b]$, and that $\tilde{f}(t;\beta)$ is the probability density function of the Simpson -based continuous logit model, also defined and continuous on the interval $[a,b]$. Also, suppose that the first and the second derivatives of the $e^{\mu V(t;\beta)}$ function are defined and continuous on the interval $[a,b]$. Then

$$\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta^3 M_4(b-a)}{480\sum_{h=1}^{K}\left[e^{\mu V\left(\delta_{h-1};\beta\right)}+4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}+e^{\mu V(\delta_h;\beta)}\right]} \leq \frac{\Delta^4 M_4}{2880e^{\mu V(c;\beta)}} \tag{35}$$

Where, $M_4 = Max_{t\in[a,b]}\left|\frac{d^4(e^{\mu V(t)})}{dt^4}\right|$, and also $V(c)$ is a constant function such that $\sum_{h=1}^{K}\left[e^{\mu V(\delta_{h-1};\beta)} + 4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)} + e^{\mu V(\delta_h;\beta)}\right] \geq \sum_{h=1}^{K} 6e^{\mu V(c;\beta)}$.

Theorem 3 Proof

This proof follows the same logic as that of Theorem 1 and Theorem 2, and thus I simplify the steps.

$f(t;\beta) - \tilde{f}(t;\beta) =$

$$f(t;\beta)\frac{\frac{\Delta}{6}\sum_{h=1}^{K}[e^{\mu V(\delta_{h-1};\beta)}+4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}+e^{\mu V\left(\delta_h;\beta\right)}]-\sum_{h=1}^{K}\int_{\delta_{h-1}}^{\delta_h} e^{\mu V(x;\beta)}dx}{\frac{\Delta}{6}\sum_{h=1}^{K}[e^{\mu V(\delta_{h-1};\beta)}+4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}+e^{\mu V\left(\delta_h;\beta\right)}]}$$

For the Simpson rule for a function $f(x)$ and its first and second derivative defined and continuous on $[a,b], \exists\, \epsilon \in (a,b)$ for which, $\int_a^b f(x)dx = \sum_{h=1}^{K}\int_{\delta_{h-1}}^{\delta_h} f(x)dx = \sum_{h=1}^{K}\frac{\Delta}{6}[f(\delta_{h-1}) + 4f\left(\frac{\delta_{h-1}+\delta_h}{2}\right) + f(\delta_h)] - \frac{1}{2880}\Delta^4(b - a)(\frac{d^4 f(x)}{dx^4} \mid_{x=\epsilon})$. Readers are referred to Burden and Faires (2001), Mathews and Fink (2004), and Süli and Mayers (2003) for details (Burden and Faires, 2001, Mathews and Fink, 2004, Süli and Mayers, 2003).

By applying absolute value, substituting with $M_4$and rearranging, $\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta^3 M_4(b-a)}{480\sum_{h=1}^{K}\left[e^{\mu V(\delta_{h-1};\beta)}+4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}+e^{\mu V(\delta_h;\beta)}\right]}$.

By substituting with $\sum_{h=1}^{K} 6e^{\mu V(c;\beta)} = \frac{b-a}{\Delta}6e^{\mu V(c;\beta)}$,

$$\left|\frac{f(t;\beta)-\tilde{f}(t;\beta)}{f(t;\beta)}\right| \leq \frac{\Delta^3 M_4(b-a)}{480\sum_{h=1}^{K}\left[e^{\mu V\left(\delta_{h-1};\beta\right)}+4e^{\mu V\left(\frac{\delta_{h-1}+\delta_h}{2};\beta\right)}+e^{\mu V(\delta_h;\beta)}\right]} \leq \frac{\Delta^4 M_4}{2880e^{\mu V(c;\beta)}} .$$

The main results from theorems 1-3 are summarized as follows,

- For the analytical bounds of the continuous choice models, it is clear that the bounds approach 0 as the $\Delta \to 0$. This is an expected but important result. It establishes that we can decrease the error of the approximation by choosing a sufficiently small Δ for our purposes. In the section of the approximation error, a Δ of 30 minutes led to absolute errors in the order of 10^{-7} for the continuous choice models.

- The midpoint-based continuous logit is more accurate for the same Δ than the trapezoidal-based continuous logit. However, this does not preclude that the actual errors are similar such as those observed in previous sections.
- The Simpson-based continuous logit offers the least amount of relative error compared to the midpoint-based continuous logit and the trapezoidal-based continuous logit. For example, the analytical bound of the Simpson-based continuous logit for a Δ of 30 minutes (i.e.0.5^4) reduces the bound by 0.0625. The analytical bound of both the trapezoidal-based continuous logit and the midpoint-based continuous logit only reduces by 0.25.

3.5 SUMMARY AND DISCUSSION

In this chapter, I reviewed the continuous logit model and discussed its specification for time-of-day choice modeling. Also, I showed that the continuous logit, albeit rarely used in its original form, is used in the form of the discretized multinomial logit model for time-of-day choice modeling. In addition, I derived other discrete and continuous choice models through the application of simple numerical integration rules: trapezoidal, midpoint, and Simpson. These models are approximations to the continuous logit model. Therefore, the approximation error of these derived models was studied using Monte Carlo experiments. Discrete choice models such as multinomial logit have been widely applied for time-of-day modeling, but these models assume discrete intervals for time, a variable continuous in nature. I showed that using the multinomial logit for time-of-day choice is, in fact, using an approximation of the continuous logit. I investigated the consequences of such approximations/ simplifications. The chapter quantified the approximation error and the statistical bias of the MLE for the approximated/simplified models through numerical experiments and analytical proofs. I tested different approaches to model time in order to find the best practices to minimize the error. I showed how small changes in the approximation scheme can lead to smaller approximation error and MLE bias.

The numerical results of these experiments indicate that the continuous choice models not only offer the ability to preserve the continuous dependent variable in its correct form, but also offer much less approximation error for the same number of partition intervals than their discrete choice model counterparts. Furthermore, analytical bounds were derived for the continuous choice models, and it was shown analytically that the Simpson-based continuous logit offers the least amount of relative error compared to the other two continuous choice models: the midpoint-based continuous logit and the trapezoid-based continuous logit. Also, it was shown as expected that the relative error (hence the approximation error) of the continuous choice models tends to zero as the number of partition intervals increases to infinity. Lastly, the numerical experiments did not show any sign of the MLE statistical bias due to the approximation error based on preliminary results. In terms of the practical contribution of this chapter, the derived continuous choice models are promising candidates for their use in time-of-day choice modeling. They offer four key features: (1) point predictions instead of predictions in aggregated intervals; (2) approximation error reduces significantly for partitions of 30 minutes intervals and other smaller intervals based on the numerical experiments; (3)

computational time of evaluation of the log-likelihood function is similar to the multinomial logit model; (4) microsimulation is possible through the use of the acceptance-rejection method (Casella and Berger, 2002). Some of the findings in this chapter are published in Ghader, et al. (2016) (Ghader, et al., 2016). The next steps are: (1) studying the correlation of alternatives in the continuous spectrum; (2) deriving models for time-of-day choice modeling of tours. The next two chapters are dedicated to these items.

4. AUTOREGRESSIVE CONTINUOUS LOGIT

In this chapter, I formulate the autoregressive continuous logit model as a novel continuous choice model capable of representing correlations across alternatives in the continuous spectrum. I formulate this model by considering two approaches: combining a discrete-time autoregressive process of order one (i.e., a linear stochastic difference equation) with the continuous logit model, which leads to the discrete-time autoregressive continuous logit; and combining a continuous-time autoregressive process (i.e., a linear stochastic differential equation), known in the stochastic process literature as Ornstein-Uhlenbeck process with the continuous logit model, which leads to the continuous-time autoregressive continuous logit. The OU process is also known as the continuous time AR(1) process, and thus it has a clear link to the AR(1) process. The discrete-time approach is formulated for comparison, and showing its ability to approximate the continuous-time version of the autoregressive continuous logit. The autoregressive continuous logit is the only RUM-based continuous choice model, besides the continuous cross-nested logit, able to handle correlations across alternatives in the continuous spectrum. For both approaches, I study their properties numerically. I also compare both approaches to highlight their relationship with each other. I also discuss the differences between the introduced model, the continuous logit, and the continuous cross-nested logit

This chapter is organized as follows: formulation of the discrete-time autoregressive continuous logit model and the continuous-time autoregressive continuous logit model for time-of-day choice modeling, including a discussion of the results of Monte Carlo experiments to illustrate these models numerically; and conclusions.

4.1 DISCRETE-TIME AUTOREGRESSIVE CONTINUOUS LOGIT MODEL

4.1.1 FORMULATION

Define a continuous logit model for an interval $[a, b]$, where both a and b are constants, and normalize the model by setting the scale μ to 1. Define an autoregressive process of order 1 (i.e., AR(1)), which is represented by $(\alpha_k : k = 1{:}K)$. Next, partition the set of integration $[a, b]$ of the continuous logit into intervals of equal length without loss of generality. Mathematically, divide [a, b] into K intervals of length Δ, where $\Delta = \frac{b-a}{K}$. The endpoints of these intervals are as follows: $\delta_k = a + k\Delta$. The conditional probability density function of this model is as follows:

$$f(t|\alpha_k : k = 1{:}K) = \frac{e^{V(t)+h_\alpha(t)}}{\int_a^b e^{V(z)+h_\alpha(z)} dz} \tag{36}$$

$$h_\alpha = \begin{cases} \alpha_1, & t \in (\delta_0, \delta_1) \\ \alpha_2, & t \in (\delta_1, \delta_2) \\ & \dots \\ \alpha_K, & t \in (\delta_{K-1}, \delta_K) \end{cases} \tag{37}$$

Notice that I have defined the piecewise function $h_\alpha(\cdot)$ for notational convenience in equation (37). The autoregressive process of order one (i.e., AR(1) process) follows the standard definition of the time series econometrics literature (Enders, 2004, Johnston and DiNardo, 1972),

$$\alpha_k = \rho\alpha_{k-1} + \epsilon_k \tag{38}$$

$$\alpha_1 = N(0, \frac{\sigma_\epsilon^2}{1-\rho^2}) \quad (39)$$

$$\epsilon_k = N(0, \sigma_\epsilon^2) \quad (40)$$

Equation (38) is the linear stochastic difference equation defining the AR(1) process. The initial value of this stochastic difference equation is given in equation (39), and the ϵ_t are independent and identically distributed disturbances defined in equation (40). This process is characterized by two parameters, ρ and σ_ϵ^2. ρ describes the correlation between the adjacent values of the process and σ_ϵ^2 is the source of variability in the process. Figure 5 illustrates the effects of the two parameters on the process. The first process has a large ρ (ρ should be in $(-1,1)$ for the process to be stationary) and a small σ_ϵ^2. This leads to an almost constant value for the process. Increasing the σ_ϵ^2 in the second process adds variability to the process, but the values still remain correlated (the process follows a trend). The third process has a moderate ρ and a moderate σ_ϵ^2. The last process has a small ρ and a large σ_ϵ^2, which is similar to white noise.

I operationalize the model by formulating the following unconditional probability density function,

$$f(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{e^{V(t)+h_\alpha(t)}}{\int_a^b e^{V(z)+h_\alpha(z)} dz} N(\alpha_1, \alpha_2, \dots \alpha_k | 0, \Sigma_\alpha)\, d\alpha_1 d\alpha_2 \dots d\alpha_k \quad (41)$$

In equation (41), $N(\cdot\, | 0, \Sigma_\alpha\,)$ is the multivariate normal density with a mean vector equal to the zero vector, and a covariance matrix equal to Σ_α . This covariance matrix is the covariance of an AR(1) process, which is defined for any two arbitrary periods, t and $t - s$, where s is a positive integer and $s < t$, as follows:

$$\sigma_\alpha(k, k-s) = \frac{\sigma_\epsilon^2}{1-\rho^2}\rho^s \tag{42}$$

In equation (42), $\sigma_\alpha\,(k, k-s)$ represents a generic element of the Σ_α covariance matrix. Equation (43) represents an arbitrary utility function in period k, in which $\xi(t)$ is a Gumbel variate with location zero and scale equal to 1. It is assumed similar to Lemp, et al. (2010) that the Gumbel variates across the continuous spectrum are independently and identically distributed (Lemp, et al., 2010).

$$U(t;k) = V(t;\beta) + \alpha_k + \xi(t) \tag{43}$$

Notice that the covariance for utility functions, $U\,(t'\,;\,k)$ and $U\,(t''\,;\,k\,-\,s)$ in two arbitrary periods, k and $k - s$, are

$$Cov\big(U(t';k), U(t'';k-s)\big) = \begin{cases} \frac{\sigma_\epsilon^2}{1-\rho^2} & t' \in [\delta_{k-1}, \delta_k] \wedge t'' \in [\delta_{k-1}, \delta_k] = 0 \\ \frac{\sigma_\epsilon^2}{1-\rho^2}\rho^s & t' \in [\delta_{k-1}, \delta_k] \wedge t'' \in [\delta_{k-s-1}, \delta_{k-s}] \wedge s \neq 0 \end{cases} \tag{44}$$

Lastly, it should be noted that for a period of 24 hours, the covariance matrix must be adjusted to consider the absolute distance between two periods, such that inter-period correlations between times before midnight and after midnight are accounted for properly.

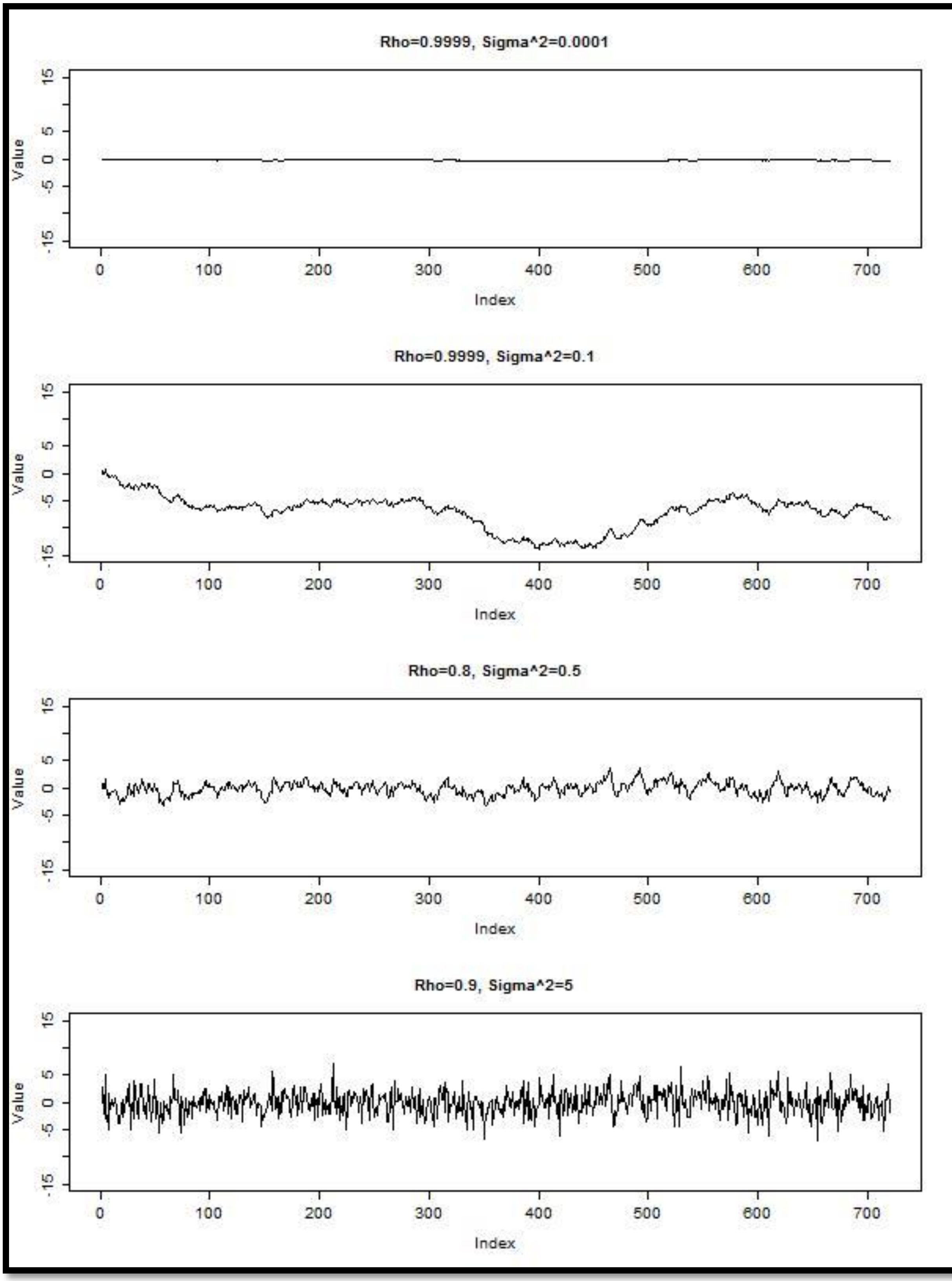


Figure 5. Effects of the parameters on the AR(1) process

Another feature of this discrete-time autoregressive continuous logit model is that the IIA property does not hold in contrast to the continuous logit model. This is easily shown by calculating the ratio of the density function for two arbitrary points in the set [a, b].

The log-likelihood function for a random sample of size *N* for the discrete-time autoregressive continuous logit model, where n is an index for an individual in the sample, is given by:

$$LL(\beta, \Sigma_\alpha) = \sum_{n=1}^{N} ln(\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{e^{V(t;\beta)+h_\alpha(t)}}{\int_a^b e^{V(z;\beta)+h_\alpha(z)} dz} N(\alpha_1, \alpha_2, \dots \alpha_k | 0, \Sigma_\alpha)\, d\alpha_1 d\alpha_2 \dots d\alpha_k)$$

(45)

where β and Σ_α are parameters to be estimated. Theoretically, this log-likelihood function may be maximized to obtain the maximum likelihood estimates (Cramer, 1986), as well as the covariance matrix to obtain the standard errors of the parameters. In practice, the maximization of this log-likelihood is difficult due to the multiple integrals. Typically, this is solved for high dimensional integration sets using other approaches, such as maximum simulated likelihood (Train, 2009). Here, I adopt maximum simulated likelihood with scrambled Sobol draws (Morokoff and Caflisch, 1994, Morokoff and Caflisch, 1995), as these perform better than pseudo-random draws and Halton draws for high dimensions. These are implemented for the multiple integrals of the α terms. An adaptive quadrature scheme (Gander and Gautschi, 2000) is used for the continuous logit kernel (i.e., integral of equation (36)) of this model. Furthermore, it should be noted that the variable t_n is indexed by *n* to indicate that t_n represents the observed value of the variable t for an individual n in the sample.

4.1.2 IDENTIFICATION

I formulated the discrete-time autoregressive continuous logit model with two parameters (ρ and σ^2) to be more general, as each parameter has an important meaning. However, extensive simulation analysis to understand the identification of the proposed model numerically following the guidance of papers such as Chiou and Walker (2007), Walker and Ben-Akiva (2002), and Walker (2001) showed that the Hessian matrix is singular (possibly pointing to collinearity between the parameters) for the cases where both σ^2 and ρ are unconstrained (Chiou and Walker, 2007, Walker and Ben-Akiva, 2002, Walker, 2001). I argue that the reason for this behavior is the lack of identification due to both parameters being unconstrained. However, extensive simulations showed that when ρ or σ^2 is fixed, then the Hessian matrix is well behaved and nonsingular. In general, the singularity of the Hessian may be due to a deficiency in the data or identification issues in the model as discussed in Cramer (1986) (Cramer, 1986). My extensive simulations showed that the issue is only present when both parameters are unconstrained and gets fixed once σ^2 or ρ are constrained.

The conditions of theoretical identification for continuous choice models are largely unexplored in the research literature. Papers such as Ben-Akiva and Watanatada (1981) and Lemp and Kockelman (2010), in addition to the third chapter of this dissertation, provide guidelines inspired by the discrete choice modeling literature such as setting the scale equal to 1, and avoiding the use of constant functions (Ben-Akiva and Watanatada, 1981, Lemp and Kockelman, 2010). However, the theoretical conditions for continuous choice models with mixing distributions unlike their discrete choice model counterparts (Walker and Ben-Akiva, 2002, Walker, 2001, Walker, et al., 2007) have not been delineated in the literature.

I can tackle the identification by fixing the σ^2. The model with a fixed σ^2can still capture inter-period and intra-period correlations through ρ. The intra-period correlation and the inter-period correlation are both obtained from equation (44) (the first case of the bracket for the intra-period, and the second case for the inter-period) by dividing the covariances by the variances (involve the variance of the Gumbel). Clearly, it can be seen that both correlation functions depend on the ρ parameter and the σ^2 parameters; thus, even if one parameter is fixed, both correlations are estimable. The ranges of intra-period and inter-period correlations allowed by the model is (0.126,1) and (-1,1) respectively, with σ^2 fixed to 0.5. Therefore, fixing one of the parameters does not lead to a significant loss of the flexibility of the model.

Lastly, the motivation for the introduced model is for its application in time-of-day choice modeling for activity-based models. For predictive accuracy, the analyst is able to leverage cross-validation to pick a value for the σ^2 that may allow for better predictive performance of the model (Friedman, et al., 2001). Alternatively, profile likelihood-based approaches may be followed to pick a better value for the σ^2 parameter by concentration of the likelihood function, if required (Pawitan, 2001). However, as it was acknowledged, the analyst is not able to compute standard errors for both ρ and σ^2 and these standard errors are only available for one of them along with the parameters in the systematic utility function. This discussion also applies to the continuous-time autoregressive continuous logit model introduced in the next section.

4.1.3 MONTE CARLO EXPERIMENT

In this section, I present the Monte Carlo experiment to test the proposed discrete-time autoregressive continuous logit model. I assume the following systematic utility function for the synthetic data set:

$$\mu V(t;\beta) = \mu(\ \beta_n TT(t) + \sum_{h=1}^{4} \left[\beta_{2h}\, sin\left(\frac{2h\pi}{12}t\right) + \beta_{2h-1}\, cos\left(\frac{2h\pi}{12}t\right)\right]) \qquad (46)$$

This systematic utility function follows the specification of Ben-Akiva and Abou-Zeid (2013) for time-of-day choice modeling for a trip (Ben-Akiva and Abou-Zeid, 2013). The true parameters were chosen based on estimates of the continuous logit model from table 5.2 in Lemp (2009), with the exception of the parameter of the travel time function $TT\ (t)$ (Lemp, 2009). This parameter was set to -0.50, because the 95% confidence interval of the travel time estimate in the model of Lemp (2009) contains 0. Also, the μ parameter is set to 1 for identification. In addition, the ρ parameter was set to a value of 0.8, and the σ_ϵ^2 was set to 0.5 and fixed for this preliminary analysis. The continuous logit of Lemp (2009) is a model of the departure time choice for the trip from home to work. I also assume that the synthetic travelers are departing from home to work. The values of the true parameters are presented in ***Table 3***. I also assumed a period of 12 hours (i.e., $[0, 12]$) in contrast to Lemp (2009), and thus I adjusted the systematic utility functions to reflect this. I also assumed 72 periods (i.e., K = 72) with a period size of 10 minutes.

The travel time function $TT\ (t)$ follows the travel speed regression model and uses the estimates from table 1 of Popuri, et al. (2008) (Popuri, et al., 2008). The dependent variable $m(t)$ of the regression model is $m(t) = ln(\frac{reported\ speed}{free\ flow\ speed})$. Therefore, the predictions for a given time t are the multiplier $m(t)$ for the free flow speed of the origin-destination pair of a traveler. In addition, the covariates are interacted with the delay, which is defined as $(1 - \frac{peak\ network\ speed}{free\ flow\ speed})$. The specification of the travel speed regression model used for the travel time function *TT (t)* is:

$$m(t) = \alpha_0 + \alpha_{dist}\, ln(distance) + \alpha_{delay} delay + \sum_{h=1}^{4}\left[\alpha_{2h} e^{hsin\left(\frac{2\pi}{12}t\right)} delay + \alpha_{2h-1} e^{h\,cos\left(\frac{2\pi}{12}t\right)} delay\right] + \sum_{h=1}^{4}\left[\alpha_{2h+8} e^{hsin\left(\frac{4\pi}{12}t\right)} delay + \alpha_{2h+7} e^{h\,cos\left(\frac{4\pi}{12}t\right)} delay\right]$$

(47)

The travel time function is defined as

$$TT(t) = \frac{distance}{free\ flow\ speed\ .e^{m(t)}} \qquad (48)$$

I assume that the travel distance (unit in miles) is distributed as a Gamma distribution with a shape parameter of two, and a scale parameter of four. The Gamma distribution was chosen because studies have shown that it is a good fit for travel distances (Ben-Akiva and Watanatada, 1981). Also, I assume that the free flow speed is distributed as a discrete uniform distribution with the following outcomes: 40, 50, 60, 70, 80, and 90. The unit used is miles per hour (MPH). For the peak network speed, I assume a discrete uniform distribution with the following outcomes: 15, 20, 25, 30, 35, and 40. The unit is miles per hour (MPH). The true parameters for the travel time function $TT\ (t)$ are presented in ***Table 3***.

I do not estimate the parameters of the travel time function $TT\ (t)$. They are assumed as constants and they are incorporated into the travel time function $TT\ (t)$, which is a covariate of the systematic utility function $V(t)$. This is the practice followed for time-of-day choice modeling in the research literature. In addition, I only simulate a cross-sectional data set. In other words, one observation is one synthetic traveler. For each synthetic traveler, I simulate (i.e., draw from the respective distribution) the travel distance, peak network speed, and free flow speed. The departure time choices of the synthetic travelers are generated using the acceptance-rejection method (Casella and Berger, 2002), as applied to the probability density function of the discrete-time autoregressive continuous logit model, evaluated at the true parameters in ***Table 3***.

Table 3. True parameters for generating the synthetic data - discrete-time autoregressive continuous logit model

Systematic Utility V(t)	
μ	1.00
βtt	-0.50
$\beta 1$	-4.73
$\beta 2$	3.70
$\beta 3$	0.00
$\beta 4$	2.43
$\beta 5$	0.00
$\beta 6$	2.46
$\beta 7$	0.00
$\beta 8$	0.55
ρ	0.80
σ_{ϵ}^{2}	0.50
Travel time TT(t)	
$\alpha 0$	0.0317
$\alpha dist$	0.0072
αdel	-1.4865
$\alpha 1$	7.0780
$\alpha 2$	-11.2967
$\alpha 3$	-5.6571
$\alpha 4$	12.5130
$\alpha 5$	1.8769
$\alpha 6$	-5.6582
$\alpha 7$	-0.2359
$\alpha 8$	0.9182
$\alpha 9$	2.3598
$\alpha 10$	-1.9606
$\alpha 11$	-0.7701
$\alpha 12$	1.7074
$\alpha 13$	0.1102
$\alpha 14$	-0.5433
$\alpha 15$	0.0018
$\alpha 16$	0.0630

For this experiment, I generated a sample size of 200 observations for the preliminary analysis of the model. Furthermore, the integral of the discrete-time autoregressive continuous logit model is computed numerically using scrambled Sobol draws (Morokoff and Caflisch, 1994, Morokoff and Caflisch, 1995) as these perform better than random draws and Halton draws for high dimensions. For the single integral of the continuous logit kernel (i.e., integral of equation (36)) of this model, an adaptive quadrature scheme is used. This quadrature is available in the R statistical package (RCore, 2012). Lastly, the Monte Carlo experiment was coded in R (RCore, 2012).

The results of the Monte Carlo experiment are presented in Table 4. In summary, the discrete-time autoregressive continuous logit model was able to recover the true parameters for this sample of size 200, but additional experiments are required to study the model. The T-Stat (null) values correspond to the t-test with hypothesis $Parameter = 0$. The T-Stat (true value) values correspond to the t-test with hypothesis $Parameter = True\ parameter$. We can see that all null t-tests are rejected with $\alpha = 0.05$, and none of the true value t-tests can be rejected with $\alpha = 0.05$. As previously mentioned, either σ_ϵ^2 or ρ needs to be fixed for the identification. I fixed σ_ϵ^2 in this experiment.

Table 4. Estimates of the parameters using the synthetic data - discrete-time autoregressive continuous logit model

Parameters V(t)	True values	Estimates	Std. Error	T-Stat (null)	T-Stat (true value)
μ	1.00	-	-	-	-
βtt	-0.50	-0.508	0.216	-2.350	-0.040
β1	-4.73	-4.598	1.456	-3.157	0.090
β2	3.70	3.585	1.494	2.399	-0.076
β3	0.00	-	-	-	-
β4	2.43	2.328	1.105	2.106	-0.091
β5	0.00	-	-	-	-
β6	2.46	2.774	0.586	4.731	0.536
β7	0.00	-	-	-	-
β8	0.55	0.669	0.328	2.038	0.363
ρ	0.80	0.827	0.120	6.853	0.226
σ_ϵ^2	0.50	-	-	-	-
$LL(\hat{\beta}, \hat{\Sigma}\alpha)$		-247.292			
Sample size		200			
Number of Periods		72 (10 minutes)			

I introduced the discrete-time autoregressive continuous logit model and showed how it is formulated. The auto-regressive component used in this model was a discrete AR(1) process, which is why I call the model discrete-time autoregressive logit. I also did some numerical experiments to show the model properties. The next subsection compares the introduced model with the other continuous choice models available in the literature.

4.1.4 COMPARISON WITH CONTINUOUS LOGIT AND CONTINUOUS CROSS-NESTED LOGIT

Comparing equation (1) and equation (36) shows us that the autoregressive continuous logit density goes to the continuous logit density if the stochastic process ($h_\alpha(t)$) goes to zero. As previously shown, the σ_ϵ^2 is the source of variation in the stochastic process; therefore, the process stays around zero if $\sigma_\epsilon^2 \to 0$. As a results, the autoregressive continuous logit model becomes similar to the continuous logit model when $\sigma_\epsilon^2 \to 0$

(the first graph in Figure 6). Increasing the value of σ_ϵ^2 shifts the autoregressive logit density away from the continuous logit density. If ρ is small, the stochastic process loses its autoregressive characteristic and oscillates around zero; therefore, increasing σ_ϵ^2 with small ρ leads to an autoregressive continuous logit density that oscillates around the continuous logit density (the last graph in Figure 6). If the ρ is relatively large in comparison with σ_ϵ^2, the autoregressive continuous logit density can take a different form and shifts away from the continuous logit without oscillating around it (the middle graphs in Figure 6).

It is also worthwhile to compare the autoregressive continuous logit model with the continuous cross-nested logit model that is also able to capture correlation in continuous spectrum. I only briefly describe this model for the sake of the comparison here, as the continuous cross-nested logit model is described in details in the next chapter. The readers can refer to Lemp (2009) and Lemp, et al. (2010) to learn more about the continuous cross-nested logit model (Lemp, 2009, Lemp, et al., 2010).

The continuous cross-nested logit can be seen as the counterpart for the cross-nested logit in the continuous spectrum. In the continuous cross-nested logit, each time point t_i in the time horizon is an alternative. Each time point also represents a nest, which contains alternatives from $t_i - h$ to $t_i + h$. The minimum time interval between uncorrelated alternatives is described by h . Alternative t_i may belong to several nests; the degree to which they belong to nest t_j is described by the allocation parameter $n(t_i, t_j)$. n should be positive, and it is normalized in a way that for each t_i , its integral over the range of possible nests is equal to 1. The probability density function of each time point t_k in this model is

$$:P(t_k) = \frac{\int_{t_k+h}^{t_k-h}[n(t_k,w)y(t_k)]^{\rho}(\int_{w-h}^{w+h}[n(r,w)y(r)]^{\rho}dr)^{\frac{1}{\rho}-1}dw}{\int_0^{24}(\int_{q-h}^{q+h}[n(r,q)y(r)]^{\rho}dr)^{\frac{1}{p}}dq} \quad (49)$$

Where $y_j = e^{v(j)}$, V_j is the systematic utility of alternative j, α and h are previously defined. ρ is called the inclusive value, which describes the amount of correlation in the nests. $\rho_m \geq 1$ should be true in order to be consistent with the random utility theory. n can take various forms. A simple one most often used in the literature, which makes calculation less cumbersome, is a triangular function:

$$n(t_i, t_j)\begin{cases} \frac{h-|t_i-t_j|}{h^2} & if\ |t_i - t_j| \leq h \\ 0 & otherwise \end{cases} \quad (50)$$

The continuous cross-nested logit model assumes alternatives are correlated because they share nests. The ρ is usually assumed to be similar for all nests, which leads to the same value of correlation between any two points at a given distance. This correlation shrinks as the distance between the points increase until it get zero when the points get h time units away from each other. The value of the correlation shrinks in the introduced model as well (equation (42)) and tends to zero, but never reaches zero. Figure 6 also shows the density of the continuous cross-nested logit model and its comparison with the continuous logit and the autoregressive continuous logit. It can be seen that the increase of ρ increases the peaks of the density and the increase of the h does not seem to have any significant effect on the density. The autoregressive continuous logit seems to be more flexible in the shape of its density when changing the parameters. Due to error assumptions, the continuous cross nested logit model is more suitable if the errors follow a nested structure. The autoregressive continuous logit model is more suitable if the errors have a trend and follow a time series.

Lastly, I should compare the computation time of the autoregressive continuous logit model with the continuous logit and the continuous cross-nested logit. Looking at equation (36) we can see that conditional on the stochastic process, the computation required for the introduced model is in the same order of the computation required for the continuous logit. However, getting the unconditional density in the maximum simulated likelihood process requires simulating the stochastic process many times (R times) and computing the conditional density for each simulated process. Consequently, the computation time required for the likelihood computation in the autoregressive continuous logit model is about R times larger than the time required for the continuous logit model (R depends on the accuracy needed and also period size, I used 500 here). Table 5 shows the computation time required for one likelihood evaluation in different models. Even though the autoregressive continuous logit model has multiple integrals and the continuous cross-nested logit has a double integral, the multiple integrals of the autoregressive model only depends on the stochastic process which can be easily simulated for the application of the maximum simulated likelihood.

Table 5. Computation time comparison

Model	Computation Time
Continuous Logit	0.015 secs
Continuous Cross-Nested Logit	8.78 secs
Autoregressive Continuous Logit (n=100)	5.17 secs
Autoregressive Continuous Logit (n=500)	25.67 secs
Autoregressive Continuous Logit (n=1000)	50.91 secs

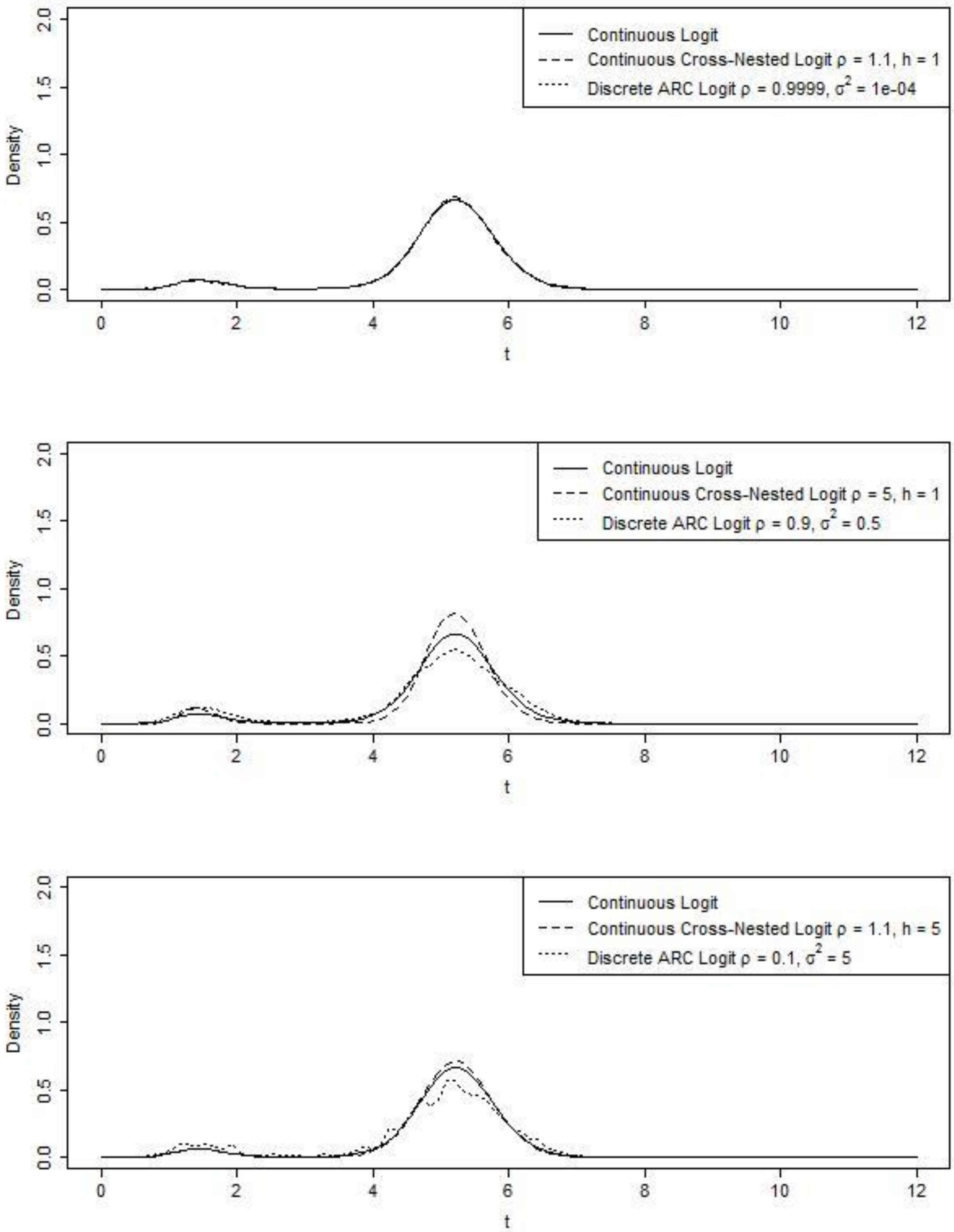


Figure 6. Discrete-time autoregressive continuous logit (period size: 10 minutes) vs. continuous logit and continuous cross-nested logit.

It should be noted that only the estimation of the discrete-time autoregressive continuous logit model is cumbersome, due to the multiple integral. On the other hand, the microsimulation of this model is similar to the microsimulation of the continuous logit model once the draws of the α terms are sampled from their respective multivariate normal density function.

4.2 CONTINUOUS-TIME AUTOREGRESSIVE CONTINUOUS LOGIT MODEL

4.2.1 FORMULATION

The continuous-time auto-regressive logit model is defined similarly to its discrete-time case, except that the stochastic term is derived from a linear stochastic differential equation and I assume it follows an Ornstein- Uhlenbeck (OU) process (Grigoriu, 2002, Shreve, 2004, Øksendal, 2003). This stochastic process may be seen as the continuous version of the AR(1) process. Similar to the previous discrete-time case, I define a continuous logit model for an interval $[a, b]$, where both a and b are constants, and also normalize the model by setting the scale μ to 1. I also define an OU process represented by $\alpha(t) : t \in [a, b]$. Equation (51) is the linear stochastic differential equation defining this process with parameters ρ and σ. This equation is known as the Langevin equation in the physical theory of Brownian motion (Wio, 2013). In general, the Langevin equation for a general Markov process is $d\alpha(t) = v_1 (\alpha, t)dt + v_2 (\alpha, t)dW (t)$, where v_1 and v_2 are arbitrary functions.

$$d\alpha(t) = -\rho\alpha(t)dt + \sigma dW (t) \tag{51}$$

ρ and σ play a role in the covariance of the process. Mathematically, the covariance and correlation of two random variables $\alpha(k)$ and $\alpha(s)$ in the set of the defined OU process are $Cov(\alpha(k), \alpha(s)) = \frac{\sigma^2 e^{-\rho|k-s|}}{2\rho}$, and $orr(\alpha(k), \alpha(s)) =$

$e^{-\rho|k-s|}$, respectively. $W\ (t)$ is the standard Brownian motion process. The initial value $\alpha(a)$ of this stochastic differential equation is given in equation (52).

$$\alpha(a) \sim N(0, \frac{\sigma^2}{2\rho}) \tag{52}$$

Similar to the discrete-time case, the conditional probability density function of the continuous-time auto-regressive logit model is presented in equation (53). It should be noted that the probability density is conditioned on the entire path of the OU process.

$$f(t|\alpha(t)\ \ :\ t\ \in [a,b]) = \frac{e^{V(t)+\alpha(t)}}{\int_a^b e^{V(z)+\alpha(z)}dz} \tag{53}$$

I operationalize this model by obtaining the unconditional probability density function *f (t)*. This unconditional density function is obtained through the use of the Feynman path integral (Wio, 2013). The Feynman path integral concept was introduced by Feynman (1942) to measure the probability of a quantum system moving from one point to another (Feynman, 1942). When a measure depends on all possible paths a quantum particle may take to move from x_i to x_j , path integrals over all possible paths are used to calculate the expected value of such measures. Feynman defines path integrals as the limit of a sequence of finite-dimensional integrals, similar to the definition of the Riemann integral, which is the limit of the sequence of finite sums (Linetsky, 1997). This definition helps justify the need for path integrals when comparing the continuous-time case and the discrete-time case. In the discrete-time case explained in the previous subsection, the time horizon was divided into K intervals, and the unconditional density function of the choice was obtained using K integrals. Therefore, the continuous-time case can be obtained by increasing the number of intervals, and allowing K $\rightarrow \infty$. As a result, the continuous-time case can be defined as the limit of a sequence of integrals when the number of intervals approaches infinity, which leads to a path integral in principle.

Intuitively, the conditional density function $f\ (t|(\alpha(t)\colon t\ \in\ [a,b]))$ must be integrated over all the possible paths that the OU process $(\alpha(t)\colon t\ \in\ [a,b])$ may follow in order to obtain the unconditional probability density function *f (t).* For this purpose, I start with the Chapman-Kolmogorov equation, which defines the transition density $g(\cdot)$ (i.e., conditional probability density, which is the conditional density of a normal distribution for the OU process, (Grigoriu, 2002, Øksendal, 2003)) between any two arbitrary time points in the process. I select the endpoints of the interval of interest $[a,b]$ such that a < m < b.

$$g(\alpha_b, b|\alpha_a, a) = \int_{-\infty}^{\infty} g(\alpha_b, b|\alpha_m, m) g(\alpha_m, m|\alpha_a, a) d\alpha_m \tag{54}$$

Following Wio (2013), I partition the interval [a, b] into K intervals of length Δ, where $\Delta = \frac{b-a}{K}$ (Wio, 2013). The endpoints of these intervals are as follows: $\delta_k\ =\ a\ +\ k\Delta$. Note that $\delta_0 =\ a$ and $\delta_K\ =\ b$. These partitions with this notation are applied to equation (54) as follows:

$$g\left(\alpha_{\delta_K}, \delta_K \middle| \alpha_a, a\right) =$$

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} g\left(\alpha_{\delta_k}, \delta_k \middle| \alpha_{\delta_{k-1}}, \delta_{k-1}\right) \dots g\left(\alpha_{\delta_1}, \delta_1 \middle| \alpha_{\delta_0}, \delta_0\right) d\alpha_{\delta_{k-1}} \dots d\alpha_{\delta_1} \tag{55}$$

Equation (55), with $K \rightarrow \infty$**,** may be interpreted as an integration over all possible paths that the OU process could follow from a given starting point $\alpha(a)$ (i.e.,

$\alpha(\delta_0))$ to a given terminal point $\alpha(b)$ (i.e., $\alpha(\delta_K)$). Furthermore, I define the following conditional probability density $f\ (t|\alpha(b), \alpha(a))$ as follows:

$$f(t|\alpha(b), \alpha(a)) =$$

$$\lim_{k\to\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{e^{V(t)+h_{\alpha(t)}}}{\int_a^b e^{V(z)+h_{\alpha(z)}} dz} g\left(\alpha_{\delta_k}, \delta_k \middle| \alpha_{\delta_{K-1}}, \delta_{K-1}\right) \dots g\left(\alpha_{\delta_1}, \delta_1 \middle| \alpha_{\delta_0}, \delta_0\right) d\alpha_{\delta_{k-1}} \dots d\alpha_{\delta_1}$$

(56)

where $h_\alpha(\cdot)$ is a function defined in equation (37), but for the current set of partitions and closed intervals. This limit of the number of partitions that tends to infinity K → ∞ (or Δ → 0) is a Feynman path integral, which is the limit of a sequence of finite-dimensional multiple integrals. The formalism of the Feynman path integral for a general Markov process is given in Wio (2013) (Wio, 2013), and for the Ornstein Uhlenbeck process in Wio, et al. (1989) (Wio, et al., 1989). For this problem, I write equation (56) following this formalism. I begin by writing the discretized version using notation of the Langevin equation found in equation (51) as follows:

$$\alpha_{\delta_k} - \alpha_{\delta_{k-1}} = \left\{-\gamma\rho\alpha_{\delta_k} - (1-\gamma)\rho\alpha_{\delta_{k-1}}\right\}\Delta + W_{\delta_k} - W_{\delta_{k-1}}, \tag{57}$$

where $W_{\delta_k} = W(\delta k)$, and also $W(\cdot)$ is a Brownian motion with parameter σ (or diffusion parameter σ^2). I assume this for $W(\cdot)$ from this point forward. In addition, γ is an arbitrary parameter defined as 0 ≤ γ ≤ 1. This parameter represents the discretization scheme used for the integration of equation (51). This is important, as it will be shown later that the Feynman path integral is dependent on the discretization scheme. Common discretization schemes are: Ito, which corresponds to γ = 0 (or pre-point discretization scheme), and Stratonovich, which corresponds to γ = 0.5 (or midpoint discretization scheme). In theory, the different discretization schemes are equivalent, but in practice, they have different numerical convergence properties. I will adopt the Ito scheme, and thus $\gamma = 0$. Readers are referred to Langouche, et al. (1980) for details (Langouche, et al., 1980).

I first need to write equation (55) for the standard Brownian Motion, and derive the corresponding equation for the OU process by transformation. For the standard Brownian Motion (Wiener process (Gardiner, 2009, Van Kampen, 1992)), Wio (2013) shows that (Wio, 2013):

$$P(W_2, \delta_2 | W_1, \delta_1) = \frac{1}{\sqrt{2\pi\sigma^2(\delta_2 - \delta_1)}} \exp[-\frac{1}{2\sigma^2(\delta_2 - \delta_1)}(W_2 - W_1)^2] \tag{58}$$

A measure in the path space known as the *Wiener* measure (Berry and Schulman, 1981, Wio, 1990) is obtained by substituting equation (58) into equation (55) (For $\delta_k - \delta_{k-1}$ assumed constant equal to Δ):

$$\prod_{k=1}^{K} \frac{dW_{\delta_k}}{\sqrt{2\pi\Delta\sigma^2}} e^{-\frac{1}{2\Delta\sigma^2}\Sigma_k \left(W_{\delta_k} - W_{\delta_{k-1}}\right)^2} \tag{59}$$

The wiener measure is the desired probability of following a given path. When $\Delta \to 0$ and $K \to \infty$, the exponential in equation (59) can be written as:

$$e^{-\frac{1}{2\sigma^2}\int_{\delta_0}^{\delta_K} dz \left(\frac{dW}{dz}\right)^2} \tag{60}$$

Following equation (55) for the Wiener process, if we integrate over all intermediate point, we recover equation (61) as the probability density of Weiner process. Therefore, the Wiener path integral can be expressed as (Berry and Schulman, 1981, Langouche, et al., 1982):

$$P(W_K, \delta_K | W_0. \delta_0) = \int e^{-\frac{1}{2\sigma^2}\int_{\delta_0}^{\delta_K} dz \left(\frac{dW}{dz}\right)^2} \mathcal{D}[W(\tau)] \tag{61}$$

The expression inside the integral is the continuous version of equation (55) over all possible paths (Notation $\mathcal{D}[W(\tau)]$ is the standard notation to show the integral is a path integral over all possible paths). In order to obtain the path integral for the more general process of this section, I rewrite equation (57) as in equation (62) and use transformation. I use the Wiener measure and the respective Jacobian in equation (63) to write the discretized form of equation (56) in the discretized version of the formalism of the Feynman path integral. This leads to equation (64).

$$W_{\delta_k} = \alpha_{\delta_k} - \alpha_{\delta_{k-1}} + \{\gamma\rho\alpha_{\delta_k} + (1-\gamma)\rho\alpha_{\delta_{k-1}}\}\Delta + W_{\delta_{k-1}} \tag{62}$$

$$J = det\left(\frac{\partial W_{\delta_k}}{\partial \alpha_{\delta_k}}\right) = \prod_{k=1}^{K}(1 + \Delta\gamma\rho) \tag{63}$$

The Jacobian J is approximately equal to $e^{\Delta\gamma \sum_{k=1}^{K} \rho}$ for $\Delta \approx 0$. This approximation is used in equation (64).

$$f(t|\alpha(b), \alpha(a)) =$$

$$\lim_{k\to\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \frac{e^{V(t)+h_{\alpha(t)}}}{\int_a^b e^{V(z)+h_{\alpha(z)}} dz} e^{\Delta\gamma K\rho} \prod_{k=1}^{K} \frac{d\alpha_{\delta_k}}{\sqrt{2\pi\Delta\sigma^2}} e^{-\frac{1}{2\Delta\sigma^2}\sum_{k=1}^{K}\left(\alpha_{\delta_k} - \alpha_{\delta_{k-1}} + \{\gamma\rho\alpha_{\delta_k} + (1-\gamma)\rho\alpha_{\delta_{k-1}}\}\Delta\right)^2}$$

(64)

Now, I take the continuous limit of the following terms as $K \to \infty$ (or $\Delta \to 0$).

$$J = e^{\Delta\gamma\,\sum_{k=1}^{K}\rho} \to e^{\gamma\rho\int_{\delta_0}^{\delta_K}dz} = e^{\gamma\rho\int_a^b dz} \tag{65}$$

Note that $\delta_0 = a$, and $\delta_K = b$, and $[\delta_0\,, \delta_K\,] = [a, b]$. I use a and b from this point forward to avoid repetition.

$$e^{\frac{\Delta}{2}\sum_{k=1}^{K}\left(\frac{\alpha_{\delta_k}-\alpha_{\delta_{k-1}}}{\Delta}\right)^2} \to e^{\frac{1}{2}\int_a^b dz\left(\frac{d\alpha}{dz}\right)^2} \tag{66}$$

$$e^{\frac{\Delta}{2}\sum_{k=1}^{K}\left(\gamma\rho\alpha_{\delta_k}+(1-\gamma)\rho\alpha_{\delta_{k-1}}\right)^2} \to e^{\frac{\rho^2}{2}\int_a^b \alpha^2(z)dz} \tag{67}$$

$$e^{\frac{\Delta}{2}\sum_{k=1}^{K}2\left(\frac{\alpha_{\delta_k}-\alpha_{\delta_{k-1}}}{\Delta}\right)\left(\gamma\rho\alpha_{\delta_k}+(1-\gamma)\rho\alpha_{\delta_{k-1}}\right)} \to e^{\rho\int_a^b \alpha(z)d\alpha} \tag{68}$$

$$h_{\alpha(t)} \to \alpha(t) \tag{69}$$

Hence, equation (54) can be written for $K \to \infty$ (or $\Delta \to 0$) as

$$f(t|\alpha(b), \alpha(a)) = \int_{\alpha(a)}^{\alpha(b)} \frac{e^{V(t)+\alpha(t)}}{\int_a^b e^{V(z)+\alpha(z)}dz}\mathfrak{D}[\alpha(t)]e^{-\frac{1}{\sigma^2}\int_a^b L\left[\alpha(z),\frac{d\alpha(z)}{dz}\right]dz} \tag{70}$$

Where $L\left[\alpha(z), \frac{d\alpha(z)}{dz}\right] = \frac{1}{2}\left(\frac{d_\alpha}{dz} + \ \rho\alpha(z)\right)^2 - \gamma\rho\sigma^2$ is known as the *Lagrangian* in the quantum literature (Feynman and Hibbs, 1965, Wio, 2013). In addition, let's define the $\mathcal{S}[\alpha(t)] = \int_a^b L[\alpha(z), \frac{d\alpha(z)}{dz}]dz$, which is known as the *action* in the aforementioned literature. Now equation (70) can be rewritten using the formalism of the Feynman path integral as:

$$f(t|\alpha(b), \alpha(a)) = \int_{\alpha(a)}^{\alpha(b)} \frac{e^{V(t)+\alpha(t)}}{\int_a^b e^{V(z)+\alpha(z)}dz}\mathfrak{D}[\alpha(t)]e^{-\frac{1}{\sigma^2}\mathcal{S}[\alpha(t)]} \tag{71}$$

Notice that the Lagrangian is dependent on the discretization scheme (γ), and I adopt the Ito scheme (γ = 0) as mentioned previously. Furthermore, $\mathfrak{D}[\alpha(t)]$ is the continuous representation of the path-dependent measure defined in equation (59), and similarly, equation (71) is the continuous representation of the limit of the multiple integrals in equation (64).

Equation (71), as mentioned previously, is an integration over all possible paths that the OU process could follow from a given starting point $\alpha(a)$ (i.e., $\alpha(\delta_0))$ to a given terminal point $\alpha(b)$ (i.e., $\alpha(\delta_K))$. However, both the initial point (i.e., $\alpha(a)$) with assumed density function in equation (51) and the terminal point are not deterministic in this case, and are assumed to be random. Therefore, the unconditional density f (t) is as follows:

$$f(t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{\alpha(a)}^{\alpha(b)} \frac{e^{V(t)+\alpha(t)}}{\int_a^b e^{V(z)+\alpha(z)}dz}\mathfrak{D}[\alpha(t)]e^{-\frac{1}{\sigma^2}S[\alpha(t)]}\right) g(\alpha_a, a)d\alpha_a d\alpha_b \tag{72}$$

The log-likelihood function for a random sample of size N for the continuous-time autoregressive continuous logit model, where n is an index for an individual in the sample, is given by

$$LL(\beta, \rho, \sigma) =$$

$$\sum_{n=1}^{N} ln(\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left(\int_{\alpha(a)}^{\alpha(b)} \frac{e^{V(t)+\alpha(t)}}{\int_a^b e^{V(z)+\alpha(z)}dz}\mathfrak{D}[\alpha(t)]e^{-\frac{1}{\sigma^2}S[\alpha(t)]}\right) g(\alpha_a, a)d\alpha_a d\alpha_b) \tag{73}$$

where β, ρ, and σ are parameters to be estimated. Theoretically, this log-likelihood function may be maximized to obtain the maximum likelihood estimates (Cramer, 1986) as well as the covariance matrix to obtain the standard errors of the parameters. In practice, the maximization of this log-likelihood is difficult due to the complex nature of the density function. I adopt the maximum simulated likelihood (Train, 2009) for the estimation. I also adopt the Euler-Maruyama method to simulate the Ornstein-Uhlenbeck process. The Euler-Maruyama method, essentially partitions the interval [a, b] into several small partitions of size Δ, thus equation (51) is written as follows using the previous notation for the partition of the interval [a, b]:

$$\alpha_{\delta_k} = (1 - \rho\Delta)\alpha_{\delta_{k-1}} + \sigma\sqrt{\Delta}\,\epsilon_{\delta_k} \tag{74}$$

where, ϵ_{δ_k} is a random variable following the standard normal distribution. The simulation of the OU process applying the Euler-Maruyama method is as follows:

1. Partition the interval [a, b] into K partitions, where $\Delta = \frac{b-a}{K}$. Set the number of desired simulations of the process (e.g., maximum_number_of_simulation = 100). Set simulation index to 1 (sim = 1).

2. Set k = 1.

3. Generate a realization for the initial value α(a) using the assumed probability density function of α(a).

4. Generate a realization for ϵ_{δ_k} . Apply equation (74).

5. Continue to the previous step until k = K. set sim=sim+1.

6. Check if sim = maximum_number_of_simulation; if yes then stop, if no then continue to step 2.

The simulation of the OU process allows sample paths (α(t) : t ∈ [a, b]) to be known, and the continuous-time autoregressive continuous logit density function can now be calculated numerically for a given set of values for the parameters of the density function. Therefore, multiple sample paths are generated and the calculations of this density function are averaged across them; this is the procedure of maximum simulated likelihood for one observation. Certainly, this must be repeated for all observations in the sample in order to calculate the log-likelihood function in equation (73) for a given set of parameters. Lastly, the logic of this numerical method is to return to equation (64) with a very small Δ.

4.2.2 RELATION TO THE DISCRETE-TIME AUTOREGRESSIVE CONTINUOUS LOGIT MODEL

The Euler-Maruyama method highlights the relation between both models. I restate equations (74) and (38),

$$\alpha_{\delta_k} = (1 - \rho\Delta)\alpha_{\delta_{k-1}} + \sigma\sqrt{\Delta}\,\epsilon_{\delta_k} \tag{75}$$

$$\alpha_k = \rho\alpha_{k-1} + \sigma\epsilon_k \tag{76}$$

Note that to avoid confusion of notation, $\alpha_{\delta_k} = \alpha_k$. Also, ϵ_k was redefined as following a standard normal distribution for comparison purposes. $\rho_{discrete} = (1 - \rho_{continuous}\Delta)$, and $\sigma_{discrete} = \sigma_{continuous}\sqrt{\Delta}$. It could be argued that the discrete-time autoregressive continuous logit model is the Euler-Maruyama numerical approximation to the continuous-time autoregressive continuous logit.

4.2.3 MONTE CARLO EXPERIMENT

I follow the same procedure for the Monte Carlo experiment presented in the discrete case, and I only mention the differences to avoid repetition. For the current implementation, I used pseudo-random draws, and also assumed α_a to be deterministic and zero. In addition, I fixed σ to be equal to 1.2 for identification, and only allowed β and ρ for estimation. The values of σ and ρ correspond to the selected values in the Monte Carlo experiment of the discrete case, following equations (75) and (76). The true parameters are presented in Table 6. For the synthetic data, the departure time choices of the synthetic travelers are generated using the acceptance-rejection method (Casella and Berger, 2002), as applied to the probability density function of the continuous-time autoregressive continuous logit model. Furthermore, the continuous-time autoregressive continuous logit model is computed numerically using the Euler-Maruyama method, as mentioned previously.

The results of the Monte Carlo experiment are presented in Table 7. In summary, the continuous-time autoregressive continuous logit model was able to recover the true parameters for this sample of size 200, but additional experiments are required to study the model. The T-Stat (null) values correspond to the t test with the hypothesis $Parameter = 0$. The T-Stat (true value) values correspond to the t-test with the hypothesis $Parameter = True\ parameter$. We can see that all null t-tests are rejected with $\alpha = 0.05$, and none of the true value t-tests can be rejected with $\alpha = 0.05$. The mathematical derivation of the continuous model and its relationship with the discrete case highlight why the results and their implications are similar to the discrete case.

Table 6. True parameters for generating the synthetic data - continuous-time autoregressive continuous logit model

Systematic Utility V(t)	
μ	1.00
βtt	-0.50
$\beta 1$	-4.73
$\beta 2$	3.70
$\beta 3$	0.00
$\beta 4$	2.43
$\beta 5$	0.00
$\beta 6$	2.46
$\beta 7$	0.00
$\beta 8$	0.55
ρ	1.2
σ_ϵ	1.2
Travel time TT(t)	
$\alpha 0$	0.0317
$\alpha dist$	0.0072
αdel	-1.4865
$\alpha 1$	7.0780
$\alpha 2$	-11.2967
$\alpha 3$	-5.6571
$\alpha 4$	12.5130
$\alpha 5$	1.8769
$\alpha 6$	-5.6582
$\alpha 7$	-0.2359
$\alpha 8$	0.9182
$\alpha 9$	2.3598
$\alpha 10$	-1.9606
$\alpha 11$	-0.7701
$\alpha 12$	1.7074
$\alpha 13$	0.1102
$\alpha 14$	-0.5433
$\alpha 15$	0.0018
$\alpha 16$	0.0630

Table 7. Estimates of the parameters using the synthetic data - continuous-time autoregressive continuous logit model

Parameters V(t)	True values	Estimates	Std. Error	T-Stat (null)	T-Stat (true value)
μ	1.00	-	-	-	-
βtt	-0.50	-0.58	0.068	-8.590	-1.240
$\beta 1$	-4.73	-4.62	0.140	-32.895	0.781
$\beta 2$	3.70	3.72	0.230	16.202	0.118
$\beta 3$	0.00	-	-	-	-
$\beta 4$	2.43	2.47	0.124	19.810	0.327
$\beta 5$	0.00	-	-	-	-
$\beta 6$	2.46	2.40	0.110	21.744	-0.515
$\beta 7$	0.00	-	-	-	-
$\beta 8$	0.55	0.65	0.105	6.180	0.950
ρ	1.20	1.16	0.132	8.794	-0.284
σ_ϵ	1.20	-	-	-	-
$LL(\hat{\beta}, \hat{\Sigma}_\alpha)$			-255.751		
Sample size			200		
Number of Partitions for the Euler-Maruyama			72		

I also illustrate the differences between the continuous logit and the Continuous-time autoregressive continuous logit model in Figure 7. The observations are similar to the discrete case.

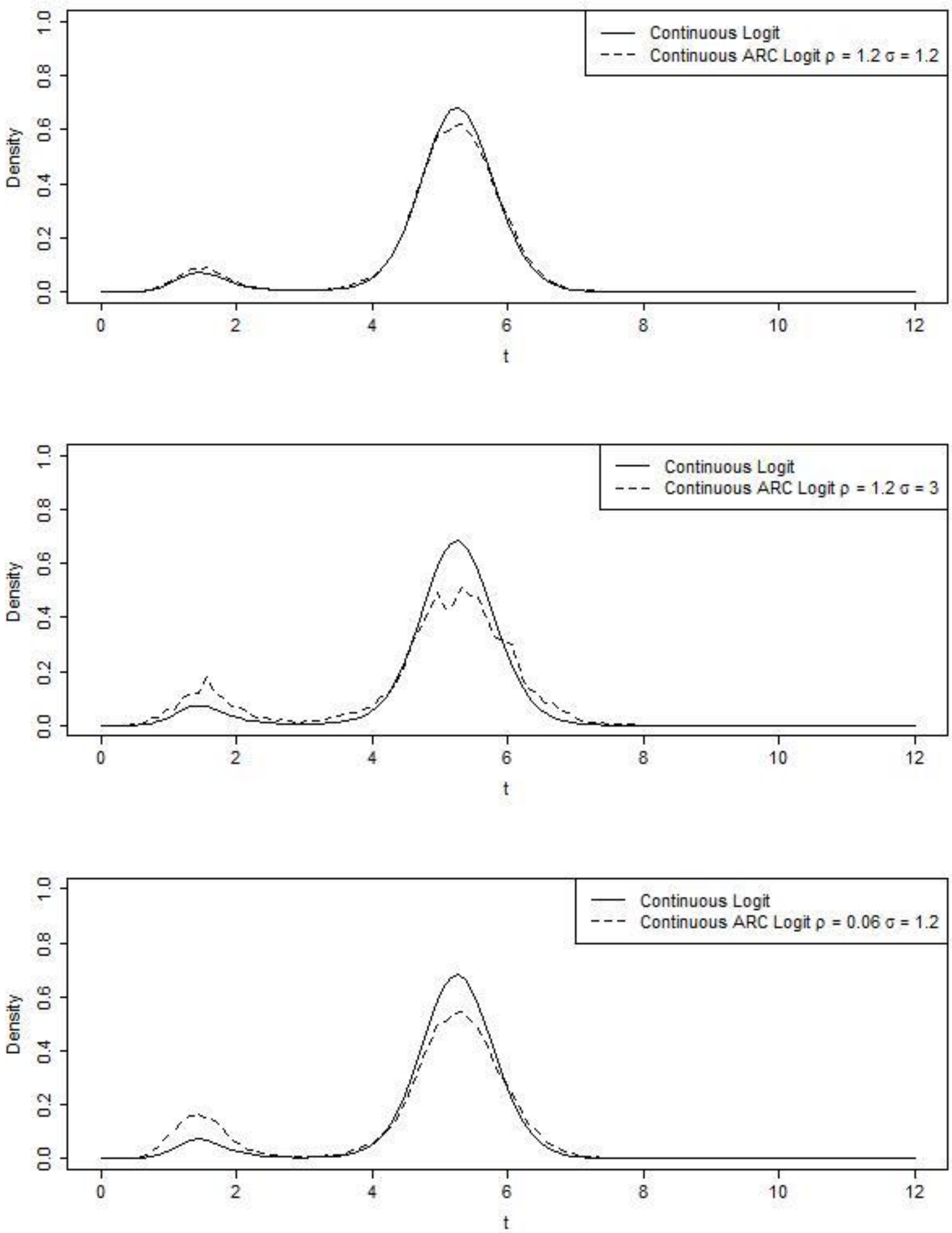


Figure 7. Continuous-time autoregressive continuous logit (Euler-Maruyama Partitions: 72) vs. continuous logit.

4.3 SUMMARY AND DISCUSSION

In this chapter, I formulated a continuous choice model, referred to as the autoregressive continuous logit model by combining the continuous logit with an autoregressive process following a discrete approach leading to the discrete-time autoregressive continuous logit and a continuous approach leading to the continuous-time autoregressive continuous logit. The autoregressive continuous logit model is the only continuous choice model, besides the continuous cross-nested logit (Lemp, et al., 2010), capable of explicitly representing correlations across alternatives in the continuous spectrum. In the discrete-time autoregressive continuous logit, the correlations are modeled for an a priori specified set of discrete periods in contrast to the continuous-time autoregressive continuous logit. In addition, it was shown that the discrete-time autoregressive continuous logit is the Euler-Maruyama approximation of the continuous-time autoregressive continuous logit. Consequently, the simpler discrete-case model can be used in practice if approximation is accepted. The approximation error can become arbitrarily small by increasing the number of intervals (K). Furthermore, I presented Monte Carlo experiments to study the models numerically, and showed they are capable of recovering the true parameters. This means the models are ready to be applied to real-world problems in which true parameters are not known.

The strength of the introduced models is their ability to predict demand in very small time intervals or to generate point prediction of travel demand. As previously explained, the fine time resolution for the demand is very useful for integration with

DTA and modeling time-dependent policies such as variable speed limit. The main limitation of the current models is their run-time. Due to the multiple integrals in the density function, the estimation of the model takes time. However, once the model is estimated, the required time for simulation and prediction is very short, similar to the continuous logit. The estimation time can be decreased by introducing approximation methods for calculating the integrals, as introduced in the third chapter. However, the effect of these approximations and their approximation errors need to be studied further. For future research, I plan to do more theoretical work and numerical experiments to study the identification of the proposed models. The findings of this chapter are published in Ghader, et al. (2019) (Ghader, et al., 2019).

5. COPULA-BASED CONTINUOUS CROSS-NESTED LOGIT

This chapter utilizes continuous cross-nested logit framework to model time-of-day choice, and employs a statistical method known as copula (Sklar, 1973, Trivedi and Zimmer, 2007) to capture the correlation between two dependent time-of-day choices. Copula facilitates the modeling of the correlation without knowing the actual bivariate distribution, and models the correlation by a closed form. This chapter presents a framework to model the arrival time to an activity choice and the departure time from an activity choice simultaneously in the continuous time setting. The framework considers both the correlation between two choices, and the correlation between alternatives of each choice; therefore, considers all the possible correlations (both on a tour, and in a day) in continuous time. Lemp, et al. (2012) captured these correlations in a discrete time setting using Multinomial Probit (Lemp, et al., 2012). Modeling the two choices together makes this framework compatible with the mainstream of activity-based models. The fully continuous, multi-dimensional, modeling framework presented in this chapter can act as a substitute for simpler time-of-day components of many activity-based models based on activity scheduling

The computation of the model is simplified using numerical integration rules discussed in the third chapter to approximate the continuous cross-nested logit's PDF and the approximation error is shown to be infinitesimal. The approximation makes the continuous cross-nested logit a practical tool to model time-of-day choice. The models' ability to find the true values is shown through Monte Carlo simulation. Moreover, in order to showcase the empirical applications, the models are estimated using real data.

INRIX data, which is a popular commercial travel time data, is used to regress travel time during the day. Regressed travel times were utilized in combination with observations from a household travel survey to jointly estimate choices of arrival to an activity and departure from the activity.

This chapter is organized as follows: formulation of the introduced framework; numerical analysis of the framework using Monte Carlo simulation; an empirical application example of the framework to showcase how it can be applied in the real world applications; and conclusions.

5.1 FORMULATION

5.1.1 COPULA

The choice of arrival time to an activity and departure time from the activity can be modeled together using their joint probability distribution; however, the true bivariate distribution between the two choices is not known. One way of obtaining the joint distribution is to assume a simple but flexible functional form for the joint distribution, known as copula, and estimate the function from the data, if data on the joint choices are available. Copula is a relatively simple function with specific characteristics that make it able to model the joint distribution (Trivedi and Zimmer, 2007). Copula can be better understood using the Sklar's theorem (Sklar, 1973):

$$F(y_1, \dots, y_m) = C(F_1(y_1), \dots, F_m(y_m);\ \theta) \tag{77}$$

Where F is the joint distribution function (CDF) of variables y_1 to y_m. F_i is the marginal distribution of variable i, C is the copula CDF Function, and θ is the parameter of copula known as the dependence parameter, which measures the dependence between marginal distributions. It can be seen that copula is a function defined on uniform (0,1) variables that produces the join distribution. Equation (77) shows how copula can be used to obtain the joint distribution from the marginal distributions. The main benefit of copula is that it separates the marginal distributions and the dependence parameter. The modeler needs to choose a copula function suitable for the data, define the marginal distributions separately, and use the data to estimate the dependence parameter. The dependence parameter can be estimated simultaneously with marginal distributions' parameters using MLE (Cramer, 1986). Another method is to estimate the marginal distributions and the copula parameters separately using Inference Function for Margins (IFM) (Joe and Xu, 2016). IFM is computationally less intensive, but the standard errors obtained from IFM need to be corrected using Bootstrapping. There are various types of copula functions introduced in the literature that differ in the form of the C function. Choice of the copula function depends on the form of the correlation between the variables. Readers can refer to Trivedi and Zimmer (2007) for a detailed explanation of various copulas (Trivedi and Zimmer, 2007). As a simple example, the form of the Frank copula, which is a widely used copula, can be seen in equation (78).

$$C(u_1, u_2, \theta) = -\theta^{-1} \log\left\{1 + \frac{(e^{-\theta u_1}-1)(e^{-\theta u_2}-1)}{e^{-\theta}-1}\right\} \quad (78)$$

In equation (78), $u_1$and u_2 represent marginal distributions, and θ is the dependence parameter. Copula has been utilized in a variety of transportation studies, such as studying the joint choice of vehicle type and miles traveled (Spissu, et al., 2009), studying residential self-selection (Bhat and Eluru, 2009), modeling spatial correlations (Bhat and Sener, 2009), modeling injury severity (Eluru, et al., 2010), modeling physical activity participation level (Sener, et al., 2010), and modeling the correlation of travel time and travel time reliability (Torkjazi, et al., 2018).

5.1.2 CONTINUOUS CROSS NESTED LOGIT MODEL

Marginal distributions need to be specified by the modeler before copula can be applied for the joint distribution. Each marginal distribution in my case is a time-of-day distribution. The model assumed for the time-of-day choice determines the form of the marginal distributions. One simple form of the choice models widely used in the literature for time-of-day choice modeling (Ben-Akiva and Abou-Zeid, 2013, Popuri, et al., 2008, Zeid, et al., 2006) is the multinomial logit (Ben-Akiva and Lerman, 1985). In the multinomial logit, the time horizon is divided into certain intervals and each interval represents an alternative. As previously discussed, the multinomial logit is unable to model correlations between alternatives and has the IIA property (Ben-Akiva and Lerman, 1985).

As discussed in the literature review chapter, researchers have relaxed the IIA property and considered the correlation between alternatives in time-of-day choice modeling with the Multivariate Extreme Value (MEV) models and mixture MEV models (Bhat, 1998, Börjesson, 2006, Chin, 1990, De Jong, et al., 2003, Hess, et al., 2007, Small, 1987). The nested logit is one of the MEV models, which deals with nested alternatives. If it is assumed that one alternative may belong to more than one nest, cross-nested logit is obtained, which can model very flexible correlation structures

(Ben-Akiva and Bierlaire, 1999, Small, 1987, Vovsha, 1997, Wen and Koppelman, 2001). The probability mass function for alternative k in the cross-nested logit is given by

$$P_k = \frac{\sum_{m=1}^{M}((\alpha_{km}y_k)^{\rho_m}[\sum_{j\in C_m}(\alpha_{jm}y_j)^{\rho_m}]^{\frac{1}{\rho_m}-1})}{\sum_{n=1}^{M}[\sum_{j\in C_n}(\alpha_{jn}y_j)^{\rho_n}]^{\frac{1}{\rho_n}}} \tag{79}$$

In equation (79), m indices the nest (assume M nests exist in total) ; α_{ij} is the allocation parameter describing the degree to which the alternative i belongs to the nest j; ρ_i is the inclusive value of the nest i, describing the level of correlation in the nest; and C_i is the subset of alternatives that belong to the nest i. $\rho_m \geq 1$ should be true in order to be consistent with the random utility theory. In addition, $\alpha_{ij} \geq 0$ and $\forall i;\ \sum_{m=1}^{M} \alpha_{im} = 1$ should be satisfied.

The continuous cross-nested logit can be seen as the continuous counterpart for the cross-nested logit. In chapter three, I showed how the multinomial logit can be seen as an approximation of the continuous logit. A similar relationship may exist between the cross-nested logit and the continuous cross-nested logit. In the continuous cross-nested logit, each time point t_i in the time horizon is an alternative. Each time point also represents the center of a nest, which contains alternatives from $t_i - h$ to $t_i + h$. h describes the minimum time interval between uncorrelated alternatives. Alternative t_i belongs to several nests (in fact an infinite number of nests). Similar to the cross-nested logit, the degree to which the alternative t_i belongs to the nest centered at t_j is described by the allocation parameter $\alpha(t_i, t_j)$. α is normalized in a way that it integrates to 1 over the range of possible nests for each t_i. The probability density function at time t_k in the continuous cross-nested logit is given by

$$P(t_k) = \frac{\int_{t_k+h}^{t_k-h} [\alpha(t_k,w) y(t_k)]^{\rho} \left(\int_{w-h}^{w+h} [\alpha(r,w) y(r)]^{\rho} dr\right)^{\frac{1}{\rho}-1} dw}{\int_0^{24} \left(\int_{q-h}^{q+h} [\alpha(r,q) y(r)]^{\rho} dr\right)^{\frac{1}{\rho}} dq} \tag{80}$$

In which $y_i = e^{v_i}$, and V_i is the systematic utility of alternative i. α and h were previously defined. Similar to the cross-nested logit, ρ is known as the inclusive value, which describes the amount of correlation in the nests.

The form of the α function, which may take various forms, needs to be specified. The triangular function, which simplifies the calculations, is widely used in the literature:

$$\alpha(t_i, t_j) \begin{cases} \frac{h-|t_i-t_j|}{h^2} & if\ |t_i - t_j| \le h \\ 0 & otherwise \end{cases} \tag{81}$$

According to equation (81), $h \geq 0.5$ should be satisfied, since each alternative should belong to at least one nest. Also, ρ is assumed to be equal in all nests for simplicity.

5.1.3 UTILITY FUNCTION SPECIFICATION

In order to use the continuous cross-nested logit model for the arrival time and the departure time choices, the systematic utility should be defined for each one of the choices separately. For the covariates, the continuous cross-nested logit model requires continuously time-varying covariates. These covariates may be smooth functions, continuous functions, and even step functions. Lemp and Kockelman (2010), Lemp (2009), Popuri, et al. (2008), and Zeid, et al. (2006) describe a methodology for obtaining travel time and travel time variability as time-varying functions through the use of regression models (Lemp and Kockelman, 2010, Lemp, 2009, Popuri, et al., 2008, Zeid, et al., 2006). These regression models are estimated using OLS. In addition, socio-economic covariates may also be included in the specification of the continuous cross-nested logit model.

For time-of-day choice modeling Ben-Akiva and Abou-Zeid (2013) introduced a specification for the $V(\cdot)$ function (systematic utility function) summarized as follows (Ben-Akiva and Abou-Zeid, 2013),

$$V(t) = \sum_{r=1}^{R} z_r s(t) + \beta_{tt} TT(t) + \cdots \tag{82}$$

$$s(t) = \sum_{k=1}^{K} \left[\beta_{2k} \sin\left(\frac{2k\pi t}{24}\right) + \beta_{2k-1} \cos\left(\frac{2k\pi t}{24}\right)\right] \tag{83}$$

where t can be either arrival time for the arrival choice to the activity or departure time for the departure time choice from the activity. $s(t)$ is a trigonometric function that captures the cyclicality of the time-of-day choices for the 24-hour day (The utility value should be equal in $t = 0$ and $t = 24$). It also makes the utility function flexible and enables it to show various trends over the day, as shown in the chapter three. The β is a vector of parameters to be estimated. K indicates the number of frequencies used in s, usually selected based on cross-validation after trying different values and comparing the goodness-of-fit results (Friedman, et al., 2001). Increasing K makes the function more flexible. Also, the $s(t)$ function is interacted with a vector of socioeconomics covariates (z) to capture the observed heterogeneity of the travelers' time-of-day preferences. R is the length of the socioeconomic covariates vector (including the intercept). Other variables may be added, such as the travel time function ($\beta_{tt}TT(t)$, where β_{tt} is a parameter to be estimated based on the methodology of the previously mentioned literature (Lemp and Kockelman, 2010, Lemp, 2009, Popuri, et al., 2008, Zeid, et al., 2006)).

In the framework of this chapter, the utility function for the arrival time and the departure time need to be specified for marginal distributions, and the copula is used to obtain the joint distribution. Equation (80) will be used for both marginal distributions. I also apply OLS regression similar to Lemp and Kockelman (2010), Lemp (2009), Popuri, et al. (2008), and Zeid, et al. (2006) using travel time data to obtain the travel time function (Lemp and Kockelman, 2010, Lemp, 2009, Popuri, et al., 2008, Zeid, et al., 2006). Details of the regression model are presented later in the chapter.

5.1.4 ESTIMATION

Nest inclusive value ρ, the minimum time interval between uncorrelated alternatives h, the copula parameter θ, and the parameter of the utility function form the set of parameters to be estimated. The Copula parameter can be estimated together with the continuous cross-nested logit parameters and the utility function parameters using the MLE method (Cramer, 1986). The form of the likelihood function can be seen in equation (85).

$$LL = \sum ln(f(t_a, t_d)) \qquad (84)$$

$$= \sum_{i=1}^{N} \sum_{j=a,d} ln\left(f_{ji}\left(t_{ji} \middle| X,\ \beta_j,\ \eta_j,\ Y_j\right)\right)$$

$$+ \textstyle\sum_{i=1}^{N} ln({}^{\partial^2 C \left(\left(F_d \middle| X,\ \beta_d,\ \eta_d,\ Y_d\right),\ \left(F_a \middle| X,\ \beta_a,\ \eta_a,\ Y_a\right);\theta\right)} \Big/ {}_{\partial F_a \partial F_d}) \quad (85)$$

Another method of estimation is the IFM method (Joe and Xu, 2016). In the IFM, the marginal distributions are estimated first and the copula parameter is estimated afterward. This method is consistent, but not efficient, and the standard errors need to be corrected using the Bootstrap algorithm (Efron and Tibshirani, 1994).

5.1.5 APPROXIMATION

One likelihood evaluation requires two density evaluations and one copula evaluation (equation (85)). One copula evaluation in a bivariate case requires evaluating two CDFs (equation (77)). The continuous-cross-nested logit density function contains two double integrals. Consequently, the likelihood evaluation requires too many integral calculations. These integrals were first computed numerically using an adaptive quadrature based on the 21-point Gauss-Kronrod quadrature. This quadrature is discussed by Piessens, et al. (2012) (Piessens, et al., 2012), and available in the R statistical package (RCore, 2012). The error tolerance of the quadrature was set to an order of magnitude equaling 10^{-10}. The computations were very time consuming and made the model estimation impractical. One way of tackling this computational problem is by using approximation tools such as numerical integration. Substituting the integrals by sums using the midpoint rule, which is a simple numerical integration rule discuss in the third chapter, decreased the CDF evaluation time by two orders of magnitude. For this chapter, the integral intervals were divided into 5-minutes sub-intervals, with each sub-integral approximated using the midpoint rule. The approximation error was in the order of 0.001 and negligible across all observations. Figure 8 shows approximated (red) versus true (black) PDF for a simulated observation. This approximation technique is used to add practicality to the framework.

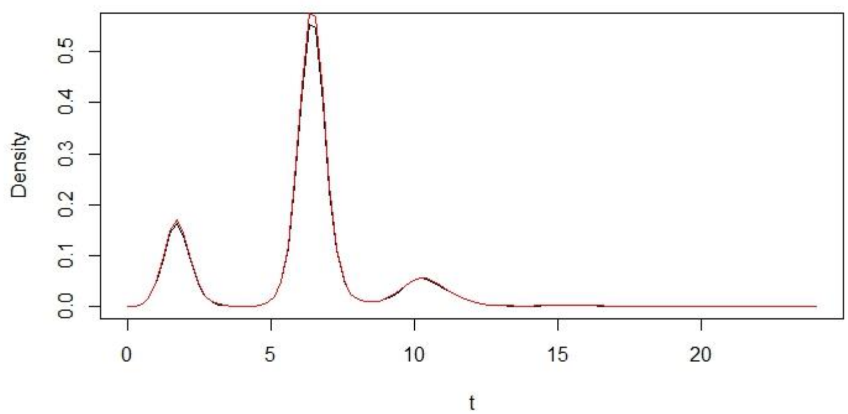


Figure 8. Approximation error in PDF evaluation using midpoint rule with 25 subintervals.

5.2 MONTE CARLO EXPERIMENT

In this section, I present the Monte Carlo experiment to test the proposed framework. Monte Carlo simulation can illustrate if the modeling framework is able to recover the assumed true parameters. I assume the following systematic utility function for the synthetic data set:

$$\mu V(t;\beta) = \mu(\,\beta_{tt}TT(t) + \textstyle\sum_{h=1}^{4}\left[\beta_{2h}\, sin\left(\frac{2h\pi}{24}t\right) + \beta_{2h-1}\, cos\left(\frac{2h\pi}{24}t\right)\right]) \tag{86}$$

This systematic utility function follows the specification of Ben-Akiva and Abou-Zeid (2013) for time-of-day choice modeling for a trip (Ben-Akiva and Abou-Zeid, 2013). The true parameters were chosen based on the estimates of the continuous logit model from table 5.2 in Lemp (2009), with the exception of the parameter of the travel time function $TT\ (t)$ (Lemp, 2009). This parameter was set to -0.50, because the 95% confidence interval of the travel time estimate in the model of Lemp (2009) contains 0. Also, the μ parameter is set to 1 for identification. In addition, the ρ parameter was set to a value of 2, and the h was set to 0.75 for this experiment based on the estimates of Lemp (2009). The value of the copula dependence parameter θ was also set to 6. The value of the true parameters are presented in Table 8.

The travel time function $TT\ (t)$ follows the travel speed regression model and uses the estimates from table 1 of Popuri, et al. (2008) (Popuri, et al., 2008). The dependent variable $m(t)$ of the regression model is $m(t) = ln(\frac{reported\ speed}{free\ flow\ speed})$. Therefore, the predictions for a given time t are the multiplier $m(t)$ for the free flow speed of the origin-destination pair of a traveler. In addition, the covariates are interacted with the delay, which is defined as $(1 - \frac{peak\ network\ speed}{free\ flow\ speed})$. The specification of the travel speed regression model used for the travel time function *TT (t)* is:

$$m(t) = \alpha_0 + \alpha_{dist}\, ln(distance) + \alpha_{delay} delay + \sum_{h=1}^{4} \left[\alpha_{2h} e^{h\, sin\left(\frac{2\pi}{24}t\right)} delay + \alpha_{2h-1} e^{h\, cos\left(\frac{2\pi}{24}t\right)} delay\right] + \sum_{h=1}^{4} \left[\alpha_{2h+8} e^{h\, sin\left(\frac{4\pi}{24}t\right)} delay + \alpha_{2h+7} e^{h\, cos\left(\frac{4\pi}{24}t\right)} delay\right]$$

(87)

The travel time function is defined as

$$TT(t) = \frac{distance}{free\ flow\ speed\ .e^{m(t)}} \tag{88}$$

The travel time function for arrival at time t gives the travel time experienced by the travelers when they arrive to the destination at time t. The travel time function for departure at time t gives the experienced travel time when travelers depart from the origin at time t. I assume that the travel distance (unit in miles) is distributed as a Gamma distribution with the shape parameter of two, and the scale parameter of four. The Gamma distribution was chosen because studies have shown that it is a good fit for travel distances (Ben-Akiva and Watanatada, 1981). Also, I assume that the free flow speed is distributed as a discrete uniform distribution with the following outcomes: 40, 50, 60, 70, 80, and 90. The unit used is miles per hour (MPH). For the peak network speed, I assume a discrete uniform distribution with the following outcomes: 15, 20, 25, 30, 35, and 40. The unit is miles per hour (MPH). The true parameters for the travel time function $TT\ (t)$ are presented in Table 8.

Table 8. True parameters for generating the synthetic data in the Monte Carlo experiment

Systematic Utility V (t)		
μ	1.00	
βtt	-0.50	
β1	-4.73	
β2	3.70	
β3	-0.51	
β4	2.43	
β5	2.40	
β6	2.46	
β7	1.34	
β8	0.55	
ρ	2	
h	0.75	
θ	6	
Travel time TT (t)		
	Arrival	Departure
α0	0.0317	0.0291
αdist	0.0072	0.0093
αdel	-1.4865	−17.3535
α1	7.0780	19.9279
α2	-11.2967	12.5456
α3	-5.6571	−15.3596
α4	12.5130	−9.4571
α5	1.8769	4.5522
α6	-5.6582	2.6285
α7	-0.2359	−0.4790
α8	0.9182	−0.2404
α9	2.3598	10.2721
α10	-1.9606	3.5345
α11	-0.7701	−8.0543
α12	1.7074	−3.2051
α13	0.1102	2.6979
α14	-0.5433	1.2742
α15	0.0018	−0.3379
α16	0.0630	−0.1915

I do not estimate the parameters of the travel time function $TT\ (t)$. They are assumed as constants and they are incorporated into the travel time function $TT\ (t)$, which is a covariate of the systematic utility function $V(t)$. This is the practice followed for time-of-day choice modeling in the research literature. In addition, I only simulate a cross-sectional data set. In other words, one observation is one synthetic traveler.

For this experiment, I generate a sample size of 400 observations for the preliminary analysis of the model. For each synthetic traveler, I simulate (i.e., draw from the respective distribution) the travel distance, peak network speed, and free flow speed. The arrival time and the departure time choices of the synthetic travelers are generated by using the acceptance-rejection method (Casella and Berger, 2002), as applied to the Frank copula density function with continuous cross-nested logit marginals, evaluated at the true parameters in Table 8. MLE was conducted afterward to estimate the parameters for comparison with their respected true values. Lastly, the Monte Carlo experiment was coded in R (RCore, 2012). Table 4 summarizes the findings of the Monte Carlo experiment. In summary, the estimates are not significantly different from their corresponding true values, suggesting that the model is able to recover the true parameters for this sample of 400 observations, but additional experiments are required to study the model. The T-Stat (true value) values correspond to the t-test with the hypothesis $Parameter = True\ parameter$. We can see that none of the true value t- tests can be rejected with $\alpha = 0.05$.

Table 9. Estimates of the parameters using the synthetic data – copula-based continuous cross-nested logit

Parameters V (t)	True values	Estimates	Std. Error	T-Stat(true value)
βtt	-0.50	-0.469	1.247	0.024
$\beta 1$	-4.73	-4.851	0.330	-0.366
$\beta 2$	3.70	3.672	0.281	-0.099
$\beta 3$	-0.51	-0.593	0.120	-0.691
$\beta 4$	2.43	2.409	0.358	-0.058
$\beta 5$	2.40	2.300	0.211	-0.473
$\beta 6$	2.46	2.412	0.260	-0.184
$\beta 7$	1.34	1.388	0.134	0.358
$\beta 8$	0.55	0.497	0.098	-0.540
ρ	2	1.855	4.814	-0.030
h	0.75	0.719	1.31	-0.023
θ	6	6.188	0.441	0.426
LL($\hat{\beta}$, $\hat{\Sigma}$ α)	-2284.546			
Sample size	400			

5.3 EMPIRICAL APPLICATION

In this section, the model estimation process using real data is presented. A continuous function of time for travel time, similar to Popuri, et al., 2008 is first estimated using INRIX data (Popuri, et al., 2008). This function was combined with observations from the 2007-2008 TPB-BMC household travel survey to estimate the copula-based continuous cross-nested logit model. Details of the used datasets in addition to the details of the travel time regression and the model estimation results are presented below.

5.3.1 DATA

5.3.1.1 INRIX Data

INRIX is a commercial company that provides data analytics based on GPS trajectories to its users. INRIX gathers data from millions of vehicles in more than 32 countries, and converts the raw data into easy-to-understand speed and travel time data. INRIX data can be used to obtain minute-by-minute travel time and speed information on the covered road segments. The spatial unit of this data is Traffic Message Channel (TMC). In my application, INIRX minute-by-minute travel time data was used to estimate travel time as a continuous function of time. The INRIX data can provide detailed information about travel time and travel time variability (Tang, et al., 2015).

5.3.1.2 2007-2008 TPB-BMC Household Travel Survey

This survey was conducted by the Transportation Planning Board (TPB) during 2007-2008 to gather information on the socio-economic, demographic, and trip-making characteristics of Washington, D.C. and Baltimore-area residents. This survey is very similar to the National Household Travel Survey (NHTS), but it is more focused on the D.C.-Baltimore area. About 31,000 persons from 14,000 households participated in this survey. This survey includes information about household, person, vehicle, and trip. It is geocoded at the Travel Analysis Zone (TAZ) level. Home-work-home tour observations made by car by individuals who have flexible work schedule were chosen for this example to estimate the copula-based continuous cross-nested logit model.

5.3.2 TRAVEL TIME REGRESSION

As previously mentioned, a continuous function of time should be used for the travel time covariate in the systematic utility function. Here, I use the methodology explained by Lemp and Kockelman (2010), Lemp (2009), Popuri, et al. (2008), and Zeid, et al. (2006) for obtaining travel time as a time-varying function through the use of regression models (Lemp and Kockelman, 2010, Lemp, 2009, Popuri, et al., 2008, Zeid, et al., 2006). A widely-used method for estimating travel time function is using a linear regression model as follows:

$$ln\left(\frac{V_i}{V_{free\,flow,i}}\right) = \beta_0 + +\beta_1[ln(distance)_i] + (delay)_i\,(\beta_2 + \sum_{k=1}^{M_1}\left[\beta_{2k+1}\sin\left(\frac{2k\pi t_i}{24}\right)\right) + \beta_{2k+2}\,cos\left(\frac{2k\pi t_i}{24}\right)]\,) + (\sum_{k=1}^{M_2}[\,\beta_{2k+2M_1+1}\sin(\frac{2k\pi t_i}{24}) + \beta_{2k+2M_1+2}cos(\frac{2k\pi t_i}{24})\,]) + \varepsilon_i \tag{89}$$

Here, i denotes the i'th observation. Instead of using travel time as the dependent variable, the speed is used, which incorporates the influence of distance on different Origin-Destination (OD) pairs. In this way, the model can be used for all OD pairs. Travel speeds (V_i) are collected from INRIX dataset.

Trip distance has an influence on travel speed and comes into the model in a logarithmic form. The variable delay is defined as:

$$delay = (1 - \frac{network\ peak\ speed}{network\ free\ flow\ speed}) \tag{90}$$

Free flow and peak speed for each OD pair were obtained using INRIX data. Peak speed is defined as the minimum of AM peak speed and PM peak speed. The delay variable works as a measure of congestion between each OD pair. A larger delay value for an OD pair shows a greater level of congestion on that OD pair.

Additionally, the delay variable was made to interact with a trigonometric function in order to project the effect of delay to any time-of-day and make the function more flexible. The cyclical and continuous shape of the trigonometric function shows the cyclical and continuous nature of time. The product of the delay variable and the trigonometric function is the contribution of delay to the dependent variable. In addition, another trigonometric function was included in the model without being interacted with the delay to capture the variation that the delay variable fails to capture. This means that time-of-day can affect travel time not only because of different levels of congestion, but also because of other reasons that influence the speed at different times of day.

In order to estimate the regression model, 40 OD pairs located in Maryland, Virginia, and the District of Columbia were selected. Since the seasonal effect is not included as an independent variable in this model, a random day is chosen for each OD pair to prevent any seasonal bias in the model. For each OD pair, INRIX travel times for departing and arriving at the beginning of each hour were used as the data points for departure and arrival travel time regressions respectively. Therefore, 960 data points in total were used in each of the regressions.

After organizing the data, the truncation points of the trigonometric functions M1 and M2 had to be determined. M1 and M2 were assumed to take values between 0 and 10, so 100 possible models were tested. In order to fit the best model, a 7-fold cross-validation was conducted (Friedman, et al., 2001). The model with $M1 = 1$ and $M2 = 4$, which had the minimum Mean Square Error (MSE), was selected as the final model.

Table 10 and Table 11 show the regression results for arrival and departure travel times. The results are consistent with expectations. Distance has a positive influence on the travel speed, indicating that a higher speed is expected for longer trips. As we can see, the delay has a negative coefficient and is statically significant, as expected; speed is low where the delay is long.

Table 10. Estimated parameters for arrival travel time regression

Parameter	Estimate	Std. Error
(Intercept)	0.035	0.009
ln(distance)	0.012	0.003
delay	-0.655	0.029
sin(2πt/24)*delay	0.157	0.041
cos(2πt/24)*delay	0.174	0.041
sin(2πt/24)	-0.007	0.011
cos(2πt/24)	0.016	0.011
sin(4πt/24)	0.011	0.005
cos(4πt/24)	0.016	0.005
sin(6πt/24)	-0.013	0.005
cos(6πt/24)	-0.015	0.005
sin(8πt/24)	-0.001	0.005
cos(8πt/24)	0.005	0.005
Number of observations		960
R squared		0.443
F(12,947)		62.855

Table 11. Estimated parameters for departure travel time regression

Parameter	Estimate	Std. Error
(Intercept)	0.027	0.009
ln(distance)	0.014	0.004
delay	-0.632	0.03
sin(2πt/24)*delay	0.138	0.042
cos(2πt/24)*delay	0.171	0.042
sin(2πt/24)	-0.007	0.011
cos(2πt/24)	0.017	0.011
sin(4πt/24)	0.009	0.005
cos(4πt/24)	0.015	0.005
sin(6πt/24)	-0.01	0.005
cos(6πt/24)	-0.014	0.005
sin(8πt/24)	-0.002	0.005
cos(8πt/24)	0.005	0.005
Number of observations		960
R squared		0.418
F(12,947)		56.584

Figure 9 shows the modeled travel time profile for both departure and arrival in one randomly selected OD pair. It demonstrates that the profile shows two peaks, and that the profile for arrival is similar to the departure profile, with the expected small shift to the right.

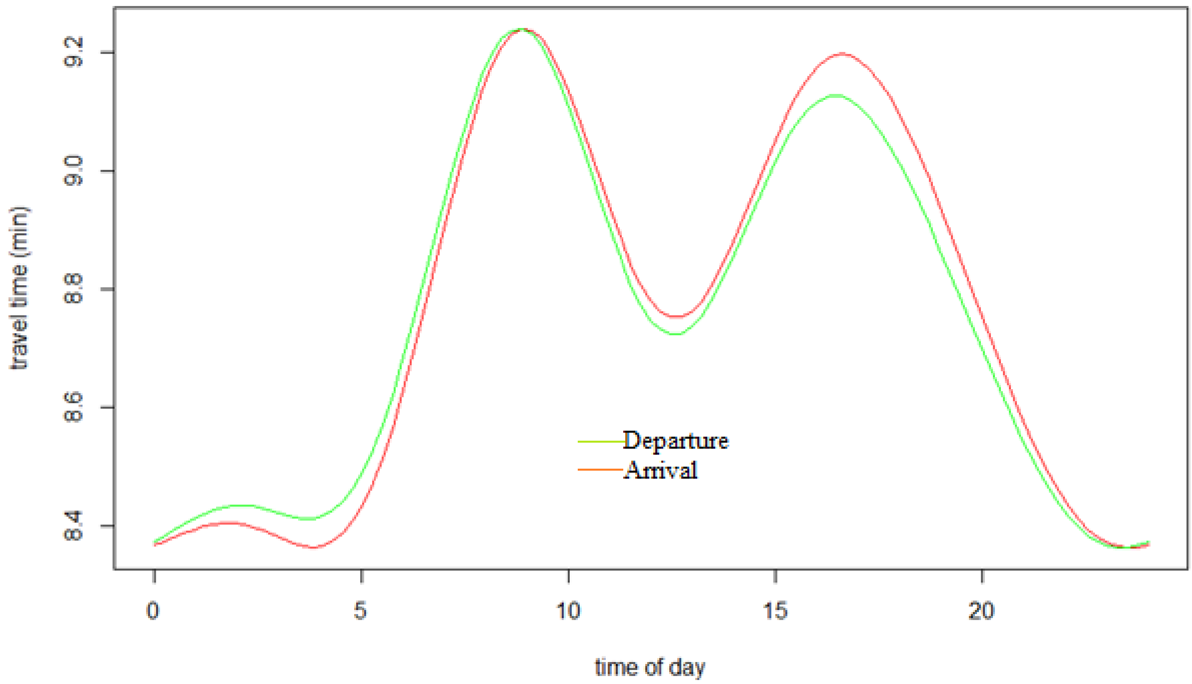


Figure 9. Travel time profiles for arrival and departure

5.3.2 UTILITY SPECIFICATION

I assumed the following utility function for the arrival to the activity choice, which is similar to the function assumed for the departure from the activity:

$$V_{arr}(t;\beta) = \beta_{tt}TT_{arr}(t) + \beta_{early}(PT_{arr} - t)\mathbb{I}_{PT_{arr}>t} + \beta_{late}(t - PT_{arr})\mathbb{I}_{t>PT_{ar}} + \sum_{h=1}^{2} age\left[\beta_{2h}\, sin\left(\frac{2h\pi}{24}t\right) + \beta_{2h-1}\, cos\left(\frac{2h\pi}{24}t\right)\right]) + \sum_{h=1}^{2} income\left[\beta_{2h+4}\, sin\left(\frac{2h\pi}{24}t\right) + \beta_{2h+3}\, cos\left(\frac{2h\pi}{24}t\right)\right]) \quad (91)$$

In which β is the vector of parameters to be estimated, TT_{arr} is the arrival travel time function, estimated before based on INIRX data, evaluated at time t. PT_{arr} is the preferred arrival time, which is assumed to be equal to the work start time. PT_{dep} in the utility function of the departure from the activity is the preferred departure time, which is assumed to be equal to the work end time. The second and third terms in the utility function, both of which contain PT are called earliness and lateness penalties respectively. They are added to the utility function to capture the scheduling preference of the travelers. $\mathbb{1}$ is the indicator function that prevents both earliness and lateness penalties to be added to the utility function simultaneously. Two socio-economic variables, person age and household income were interacted with the trigonometric function to capture some level of heterogeneity.

5.3.3 COPULA SELECTION

As previously mentioned, various copula functions are available in the literature. Some can only capture positive or negative correlations, some others are suitable for both. Some can capture strong tail correlations, some others fail to capture any tail correlations. I take an approach that is based on discussions in Armstrong (2003), Bhat and Eluru (2009), Nelsen (2007), and Trivedi and Zimmer (2007) to choose a copula suitable for my application, among candidate copula functions (Armstrong, 2003, Bhat and Eluru, 2009, Nelsen, 2007, Trivedi and Zimmer, 2007). The approach can be summarized in the following steps:

1. Estimate each marginal distribution separately
2. Calculate the marginal cumulative probabilities for each original data observation
3. transform the marginal cumulative probabilities to standard normal variables using the inverse of the normal CDF

4. Plot the transformed variables

5. Calculate the Kendell's τ rank correlation statistic (Kendell, 1955) for the transformed variables

6. Draw pairs of random uniform variables from the candidate copulas with the Kendall's τ value obtained in step 5

7. Repeat steps 3 and 4 for the pairs of step 6

8. Compare the plot of original data with the plots corresponding to the candidate copulas and choose the copula whose plot better match the observed data plot

Figure 10 shows these plots for my application. The top left figure shows the plot corresponding to the survey observations. The Kendell's τ rank correlation statistic for the data is equal to 0.21. The rest of the graphs correspond to the 5 candidate copulas tested: FGM, Clayton, Gaussian, Frank, and Gumbel. The main observations in the observed data plot helping the copula selection are the asymmetry and the stronger left-tail correlation in comparison with the right-tail correlation. These observations match better with the Clayton copula; therefore, Clayton copula was used for this empirical application. It is worthwhile to mention that in cases with a small value of rank correlation such as the case of this example, it may be difficult to differentiate between different copulas. The differences between the copulas become more obvious as the value of the rank correlation increases.

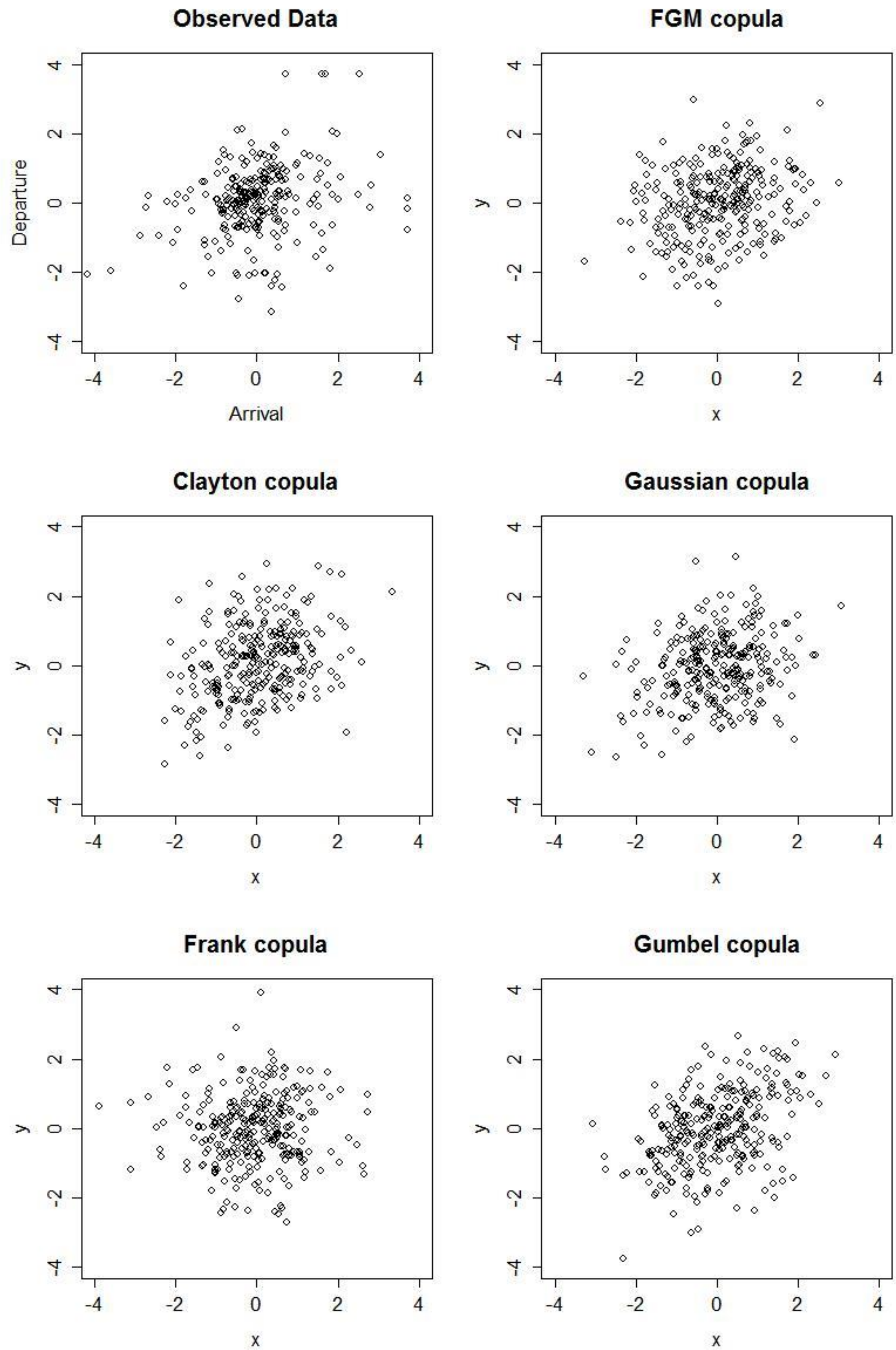


Figure 10. Comparing the data with the draws from different copulas

5.3.4 MODEL ESTIMATION RESULTS

In this part, observed data from the travel survey were used to estimate the copula-based continuous cross-nested logit model for home-work-home car tours by individuals with a flexible work schedule in the Washington, D.C.-Baltimore area. The form of the systematic utility function is assumed to be similar to equation (91). Continuous travel time functions, estimated from the INRIX data and presented before, were used for the travel time covariate. As trip purpose has a significant effect on departure time and arrival time choices, time-of-day models should differ by trip purpose. Here, the focus is on work trip purpose. Also, the correlation of the arrival and departure time is limited for those with a fixed work schedule, as their arrival and departure are usually dictated by their employer. Furthermore, the INRIX travel time data only includes driving, therefore, the focus of this example would be on car trips. As a result, the observed car tours of home to work and work to home purpose by individuals with a flexible work schedule were selected from the 2007-2008 TPB-BMC HHTS as the sample.

Table ***12*** shows the estimated parameters with their level of significance. The parameter h is estimated to be 2.37 and significant, which shows that alternatives that are less than 2.37 hours from each other are correlated. The parameter ρ is also estimated to be 3.14 and significant, showing that the amount of correlation within each nest is significant and positive, as expected. As previously shown, the correlation between the two choices is positive and small, but still statistically significant. The parameters for the travel time and scheduling penalties are all negative and significant as expected.

The β_{early} value is estimated to be smaller in value compared to the β_{leta}, which shows that the penalty for being late is higher than the penalty for being early.

Table 12. Estimation results for HBW car tours in Washington D.C.-Baltimore area

Parameter	Estimate	Standard Error	t-Stat
CCNL h	2.37291	0.05903	40.19837
CCNL ρ	3.14524	0.646072	4.86825
Copula θ	0.21264	0.105191	2.021466
β_t	-0.47202	0.022428	-21.046
β_{early}	-0.34261	0.054219	-6.319
β_{late}	-0.77547	0.134176	-5.7795
β_1	-0.14282	0.025317	-5.64127
β_2	0.031122	0.20505	0.151778
β_3	-0.04068	0.137083	-0.29675
β_4	0.053923	0.104276	0.517118
β_5	-0.11719	0.187788	-0.62405
β_6	0.057519	0.006978	8.242906
β_7	0.109079	0.087105	1.25227
β_8	0.039678	0.110464	0.359194
-Log-Likelihood	1003.927		
# of observations	285		
AIC	2035.854		
BIC			

5.3.5 COMPARISON WITH OTHER MODELS

It is worthwhile to compare the copula-based continuous cross-nested logit model estimation results with estimation results from some other available methods. In this subsection, I compare the copula-based continuous cross-nested logit with the following models: two independent multinomial logits, two independent continuous logits, two independent continuous cross-nested logits, and copula-based continuous logit. The specification of the utility for all the models is similar to equation (91). For the multinomial logit model, 24 one-hour alternatives were used. The value of the utility function was evaluated at the midpoint of alternatives. For the copula-based continuous logit, the Clayton copula was used, similar to the copula-based continuous cross-nested logit.

Table 13. Comparison of different models

Parameter	Independent MNLs	Independent CLs	Independent CCNLs	Copula-based CL	Copula-based CCNL
CCNL h	-	-	2.42508	-	2.37291
CCNL ρ	-	-	3.25442	-	3.14524
Copula θ	-	-	-	0.19620	0.21264
β_t	-0.45271	-0.46936	-0.51052	-0.32324	-0.47202
β_{early}	-0.42901	-0.59186	-0.49534	-0.67888	-0.34261
β_{late}	-0.76434	-0.66823	-0.67277	-0.54134	-0.77547
β_1	-0.0143	-0.0121	-0.0105	-0.04764	-0.14282
β_2	0.01869	0.00359	0.01456	-0.11912	0.031122
β_3	-0.011	-0.0051	-0.0083	-0.10992	-0.04068
β_4	0.00027	-0.0034	-0.0002	-0.16802	0.053923
β_5	-0.0636	-0.066	-0.0481	0.09656	-0.11719
β_6	-0.0824	-0.0802	-0.0912	-0.18465	0.057519
β_7	-0.0352	-0.0618	-0.0176	0.08218	0.109079
β_8	-0.0197	-0.014	-0.0111	-0.16767	0.039678
-Log-Likelihood	1130.096	1102.274	1026.705	1098.161	1003.927
# of observations	285	285	285	285	285
AIC	2228.192	2226.548	2079.41	2220.322	2035.854
BIC	2322.369	2266.725	2126.892	2264.151	2086.988

We can see that the continuous logit model is performing better than the multinomial logit in terms of all the tested goodness-of-fit measures. This shows that the benefits of modeling time with a continuous variable may also show themselves in terms of improvements in goodness-of-fit. The continuous cross-nested logit model is performing better than the continuous logit in terms of all the tested goodness-of-fit measures. This highlights that the benefits of modeling correlation among alternatives may also include an improvement in the goodness-of-fit. The copula-based continuous logit is performing better than the continuous logit, which shows that modeling the correlation among choices can lead to improvements in log-likelihood, even if the correlation is small. Finally, we can see that the copula-based continuous cross-nested logit model is the best among the tested models in terms of all the tested goodness-of-fit measures. This model, not only estimates behaviorally meaningful parameters such as h, ρ, and θ, it also leads to a better log-likelihood. However, there is a cost involved with the usage of this model in terms of the computation time. For some applications, the simpler models on the left side of Table 13 might suffice. There is a trade-off between the computation time and the benefits gained by using more complex models. Considering the significant ongoing advancements in computing, and the growing demand for a detailed time-of-day modeling of individual level activates during the day with fine time resolution, the more complex models on the right side of Table 13 can soon go from the sate-of-the-art to the state-of the-practice and be utilized in travel demand models.

5.4 SUMMARY AND DISCUSSION

This chapter utilized copula to model choices of arrival to the activity and departure from the activity simultaneously in continuous time. The continuous cross-nested logit is able to model the correlation between different times of day and copula is able to model the correlation between two choices. Consequently, all the possible correlations are considered in continuous time by the introduced framework. Numerical approximation methods were used to decrease the computational burden and make this framework a practical tool to be used for real applications. The approximation error was shown to be very small. Numerical analysis of the model using Monte Carlo simulation illustrated that the assumed true values can be recovered by the model, suggesting that the model is ready to be applied to real data. The application of this modeling framework with real data was showcased using travel time data from INRIX and observations from a household travel survey. The results showed that the proposed framework performs better than its simpler counterparts in terms of the goodness-of-fit. The continuous nature of the framework makes it suitable for applications requiring travel demand in fine time resolution, such as DTA integration; real-time applications for connected and automated vehicles or real-time traffic control; and evaluating time-dependent policies such as variable speed limit, dynamic tolling, etc. The multi-dimensionality of the introduced framework makes it compatible with the majority of activity-based and tour-based models. The time-of-day modeling framework introduced in this chapter can replace the time-of-day component of many activity-based models based on activity scheduling in order to make these models more powerful tools that can capture all types of correlation in continuous time. Some of the findings in this chapter are published in Ghader, et al. (2017) (Ghader, et al., 2017).

6. CONCLUDING REMARKS

Any travel demand model, whether trip-based, tour-based, or activity-based, tries to replicate the real-world travel environment. Time is an essential component of any travel; therefore, a sufficient travel demand model should have a time component. Many travel demand models suffer from a lack of time component or a simplified time component dividing the demand among aggregate time intervals. Without an adequate time component, the model would not be sensitive to temporal changes of demand, while many demand management policies are targeting the temporal distribution of demand. There are various benefits associated with modeling travel demand in continuous time. The first benefit is in terms of application. A model with a continuous time component would be able to predict travel demand in any small time interval, or even give a point prediction of travel demand. Such models would be sensitive to any time-dependent policy, capturing even small shifts in the temporal distribution of demand. With a continues time component, the model would be a suitable tool for evaluating the effect of variable speed limit, dynamic message sign, dynamic tolling, or any other time-dependent policy on travel demand, while a model with aggregate intervals would not be sensitive to changes within its intervals. The second benefit is in integration with traffic assignment. Travel demand with fine time resolution can improve traffic assignment in DTA models. The current activity based model and DTA model integrations usually use demand in 1-hour or 30-minutes intervals. The integration can benefit from travel demand with finer time resolution provided by the demand models with continuous time. The third benefit is in terms of data requirement. A discrete choice model requires alternative-specific data, and the data size increases by increasing the number of alternatives. For modeling time using a discrete variable with small-time intervals, one requires a significant amount of data. The fourth benefit is that the modeler does not need to choose suitable discrete alternatives for continuous time-of-day modeling. There is no theoretical foundation helping the modeler to choose the discrete alternatives for the time variable which is continuous in nature.

In this dissertation, I studied some of the current practices for time-of-day modeling in more depth and introduced new models for continuous time-of-day modeling. I showed that some of the current models in the literature, such as the multinomial logit model, are an approximated form of the continuous logit. This highlights that in many cases, the use of discrete choice models is in-fact using an approximation of their continuous choice model counterpart. I studied the approximation error. I also demonstrated how new continuous and discrete choice models can be obtained from continuous logit by approximation techniques, and compared their approximation error. Through numerical and analytical results, I showed how small changes in the approximation technique may lead to better models in terms of approximation error. I also highlighted the benefits of using approximations in the form of continuous choice models in comparison with the approximations in the form of discrete choice models. The continuous choice models not only keep the time variable in its correct form, but they also have significantly smaller approximation errors.

Due to the IIA property of the continuous logit, I introduced and formulated the autoregressive continuous logit, and showed how it can be derived from a combination of continuous logit and a stochastic process. I derived two versions of the autoregressive continuous logit, one in continuous form, and one in discrete form, and showed that the discrete-time autoregressive continuous logit is the Euler-Maruyama approximation of the continuous-time autoregressive continuous logit. Through Monte Carlo experiments, I showed that the introduced autoregressive continuous logit model is able to recover the true parameters, therefore, it is ready to be applied to real world-data where the true parameters are unknown.

One of the main applications of the time-of-day models is in activity-based travel demand models. In the activity-based demand models, activities are modeled with their temporal and spatial dimensions. Each activity has an arrival time and a departure time; therefore, time-of-day choice for activity participation includes more than one choice situation. I extended the continuous choice models that were limited to one-dimensional choice into a multi-dimensional framework. I introduced the copula-based continuous cross-nested logit as a framework that can capture all possible sorts of correlation required for an activity-based model. I used copula to capture the correlation between dependent choices of arrival to the activity and departure from the activity. I studied the copula-based continuous cross-nested logit model numerically using Monte Carlo experiments to show that it is able to recover the true parameters. I also showcased the use of the copula-based continuous cross-nested logit framework in a real-world situation using real-world data. Through the empirical application example, I compared the copula-based continuous cross-nested logit with the multinomial logit, the continuous logit, the continuous cross-nested logit, and the copula-based continuous logit, and showed that it outperforms them in terms of the goodness-of-fit.

The work of this dissertation can be followed in the future in several directions. The first is studying the identification of the continuous choice models with mixing distributions. In the fourth chapter, I discussed that the autoregressive continuous choice model may have identification issues with both parameters unconstrained and mentioned that the identification of such continuous models is not studied in-depth in the literature. The second direction is hybrid choice modeling in continuous time for incorporating scheduling preferences. Information about preferred schedules is usually missing in the surveys. These preferences can be modeled using latent variables and hybrid choice modeling. I previously studied hybrid-choice time-of-day modeling in discrete time (Ghader, et al., 2018), and I plan to extend that work to continuous time. The third direction is toward numerical methods for simplifying the calculations required for the likelihood evaluation, or heuristics and metaheuristics for MLE optimization of the continuous time-of-day models. The biggest current limitation in the introduced models is their model estimation computation time. Introducing simplifications and approximations or fine-tuning the optimization algorithms to find the optimum parameters in fewer iterations may address the main limitation of the introduced models.

REFERENCES

Abkowitz, M. D. (1981), 'An Analysis of the Commuter Departure Time Decision', *Transportation* Vol. 10, No. 3, pp. 283-297.

Armstrong, M. (2003), 'Copula Catalogue. Part 1: Bivariate Archimedean Copulas', *Unpublished paper available at http://www.cerna.ensmp.fr*.

Ben-Akiva, M. and M. Abou-Zeid. (2013), 'Methodological Issues in Modelling Time-of-Travel Preferences', *Transportmetrica A: Transport Science* Vol. 9, No. 9, pp. 846-859.

Ben-Akiva, M. and M. Bierlaire. (1999), 'Discrete Choice Methods and Their Applications to Short Term Travel Decisions', *Handbook of Transportation Science*, Springer.

Ben-Akiva, M., N. Litinas and K. Tsunokawa. (1985), 'Continuous Spatial Choice: The Continuous Logit Model and Distributions of Trips and Urban Densities', *Transportation Research Part A: General* Vol. 19, No. 2, pp. 119-154.

Ben-Akiva, M. and T. Watanatada. (1981), 'Application of a Continuous Spatial Choice Logit Model', *Structural analysis of discrete data with econometric applications*, pp. 320-343.

Ben-Akiva, M. E. and S. R. Lerman. (1985), *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT press.

Berry, M. and L. Schulman. (1981), 'Techniques and Applications of Path Integration'.

Bhat, C. R. (1996), 'A Hazard-Based Duration Model of Shopping Activity with Nonparametric Baseline Specification and Nonparametric Control for

Unobserved Heterogeneity', *Transportation Research Part B: Methodological* Vol. 30, No. 3, pp. 189-207.

———. (1998), 'Accommodating Flexible Substitution Patterns in Multi-Dimensional Choice Modeling: Formulation and Application to Travel Mode and Departure Time Choice', *Transportation Research Part B: Methodological* Vol. 32, No. 7, pp. 455-466.

———. (1998), 'Analysis of Travel Mode and Departure Time Choice for Urban Shopping Trips', *Transportation Research Part B: Methodological* Vol. 32, No. 6, pp. 361-371.

Bhat, C. R. and N. Eluru. (2009), 'A Copula-Based Approach to Accommodate Residential Self-Selection Effects in Travel Behavior Modeling', *Transportation Research Part B: Methodological* Vol. 43, No. 7, pp. 749-765.

Bhat, C. R. and I. N. Sener. (2009), 'A Copula-Based Closed-Form Binary Logit Choice Model for Accommodating Spatial Correlation across Observational Units', *Journal of Geographical Systems* Vol. 11, No. 3, pp. 243-272.

Bhat, C. R. and J. L. Steed. (2002), 'A Continuous-Time Model of Departure Time Choice for Urban Shopping Trips', *Transportation Research Part B: Methodological* Vol. 36, No. 3, pp. 207-224.

Binmore, K. G. (1982), *Mathematical Analysis: A Straightforward Approach*, Cambridge University Press.

Bolduc, D. and M. Ben-Akiva. (1991), 'A Multinomial Probit Formulation for Large Choice Sets', (Ed.)^(Eds.), *LES METHODES D'ANALYSE DES*

COMPORTEMENTS DE DEPLACEMENTS POUR LES ANNEES 1990-6E CONFERENCE INTERNATIONALE SUR LES COMPORTEMENTS DE DEPLACEMENTS, CHATEAU BONNE ENTENTE, QUEBEC, 22, 23, 24 MAI 1991, *VOLUME 2*.

Burden, R. L. and J. D. Faires. (2001), 'Numerical Analysis', (Ed.)^(Eds.), Brooks/cole Pacific Grove, CA.

Börjesson, M. (2006), 'Departure Time Modelling: Applicability and Travel Time Uncertainty', *Transportation Research Part B: Methodological*.

———. (2008), 'Joint Rp–Sp Data in a Mixed Logit Analysis of Trip Timing Decisions', *Transportation Research Part E: Logistics and Transportation Review* Vol. 44, No. 6, pp. 1025-1038.

Cardell, N. S. and F. C. Dunbar. (1980), 'Measuring the Societal Impacts of Automobile Downsizing', *Transportation Research Part A: General* Vol. 14, No. 5, pp. 423-434.

Casella, G. and R. L. Berger. (2002), *Statistical Inference*, Duxbury Pacific Grove, CA.

Chin, A. T. (1990), 'Influences on Commuter Trip Departure Time Decisions in Singapore', *Transportation Research Part A: General* Vol. 24, No. 5, pp. 321-333.

Chiou, L. and J. L. Walker. (2007), 'Masking Identification of Discrete Choice Models under Simulation Methods', *Journal of Econometrics* Vol. 141, No. 2, pp. 683-703.

Costinett, P., M. Jain, R. Moeckel, E. Wardell and T. Weidner. (2009), 'Maryland Statewide Transportation Model User's Guide: Draft Report', (Ed.)^(Eds.), Tech rep, Maryland DOT State Highway Administration.

Cramer, J. (1986), 'Econometric Applications of Maximum Likelihood Methods Cambridge University Press', *New York, New York*.

De Jong, G., A. Daly, M. Pieters, C. Vellay, M. Bradley and F. Hofman. (2003), 'A Model for Time of Day and Mode Choice Using Error Components Logit', *Transportation Research Part E: Logistics and Transportation Review* Vol. 39, No. 3, pp. 245-268.

Efron, B. and R. J. Tibshirani. (1994), *An Introduction to the Bootstrap*, CRC press.

Eluru, N., R. Paleti, R. M. Pendyala and C. R. Bhat. (2010), 'Modeling Injury Severity of Multiple Occupants of Vehicles: Copula-Based Multivariate Approach', *Transportation Research Record* Vol. 2165, No. 1, pp. 1-11.

Enders, W. (2004), 'Applied Econometric Time Series, by Walter', *Technometrics* Vol. 46, No. 2, pp. 264.

Feynman, R. P. (1942), 'The Principle of Least Action in Quantum Mechanics', Princeton University Princeton, New Jersey.

Feynman, R. P. and A. Hibbs. (1965), *Quantum Mechanics and Path Integrals [by] Rp Feynman [and] Ar Hibbs*, McGraw-Hill.

Friedman, J., T. Hastie and R. Tibshirani. (2001), *The Elements of Statistical Learning*, Springer series in statistics New York.

Gadda, S., K. Kockelman and P. Damien. (2009), 'Continuous Departure Time Models: A Bayesian Approach', *Transportation Research Record: Journal of the Transportation Research Board*, No. 2132, pp. 13-24.

Gander, W. and W. Gautschi. (2000), 'Adaptive Quadrature—Revisited', *BIT Numerical Mathematics* Vol. 40, No. 1, pp. 84-101.

Gardiner, C. (2009), 'Stochastic Methods', (Ed.)^(Eds.), Springer Berlin.

Ghader, S., C. Carrion, L. Tang and L. Zhang. (2016), 'Revisiting the Continuous Logit: Numerical Analysis and Application to Time of Day Choice Modeling', (Ed.)^(Eds.), *Transportation Research Board 95th annual Meeting*.

Ghader, S., C. Carrion and L. Zhang. (2019), 'Autoregressive Continuous Logit: Formulation and Application to Time-of-Day Choice Modeling', *Transportation Research Part B: Methodological* Vol. 123, pp. 240-257.

Ghader, S., A. Darzi and L. Zhang. (2018), 'Hybrid Choice Framework for Time-of-Day Choice in Trip-Base Models', (Ed.)^(Eds.).

Ghader, S., L. Tang, C. Carrion, A. Asadabadi and L. Zhang. (2017), 'A Copula-Based Continuous Cross-Nested Logit Model for Tour Scheduling in Activity-Based Travel Demand Models', (Ed.)^(Eds.), *Transportation Research Board 96th Annual Meeting*.

Grigoriu, M. (2002), 'Stochastic Calculus-Applications in Science and Engineering, Birkhuser', (Ed.)^(Eds.), Springer-Verlag, New York, NY.

Hammersley, J. (2013), *Monte Carlo Methods*, Springer Science & Business Media.

Hendrickson, C. and E. Plank. (1984), 'The Flexibility of Departure Times for Work Trips', *Transportation Research Part A: General* Vol. 18, No. 1, pp. 25-36.

Hess, S., A. Daly, C. Rohr and G. Hyman. (2007), 'On the Development of Time Period and Mode Choice Models for Use in Large Scale Modelling Forecasting Systems', *Transportation Research Part A: Policy and Practice* Vol. 41, No. 9, pp. 802-826.

Hess, S., J. W. Polak, A. Daly and G. Hyman. (2007), 'Flexible Substitution Patterns in Models of Mode and Time of Day Choice: New Evidence from the Uk and the Netherlands', *Transportation* Vol. 34, No. 2, pp. 213-238.

Hyman, G. (1997), 'The Development of Operational Models for Time Period Choice', *Department of the Environment, Transport and the Regions, HETA Division, London*.

Joe, H. and J. J. Xu. (2016), 'The Estimation Method of Inference Functions for Margins for Multivariate Models'.

Johnston, J. and J. DiNardo. (1972), 'Econometric Methods', *New York* Vol. 19, No. 7, pp. 22.

Kendell, M. (1955), 'Rank Correlation Methods', *Charles Griffin and Company: London*.

Khuri, A. I. (2004), 'Applications of Dirac's Delta Function in Statistics', *International Journal of Mathematical Education in Science and Technology* Vol. 35, No. 2, pp. 185-195.

Komma, A. and S. Srinivasan. (2008), 'Modeling Home-to-Work Commute-Timing Decisions of Workers with Flexible Work Schedules', (Ed.)^(Eds.), *Transportation Research Board 87th Annual Meeting*.

Langouche, F., D. Roekaerts and E. Tirapegui. (1980), 'Short Derivation of Feynman Lagrangian for General Diffusion Processes', *Journal of Physics A: Mathematical and General* Vol. 13, No. 2, pp. 449.

———. (1982), 'Functional Integration And', *Semiclassical Expansions*.

Lemp, J. and K. Kockelman. (2010), 'Empirical Investigation of Continuous Logit for Departure Time Choice with Bayesian Methods', *Transportation Research Record: Journal of the Transportation Research Board*, No. 2165, pp. 59-68.

Lemp, J. D. (2009), 'Capturing Random Utility Maximization Behavior in Continuous Choice Data: Application to Work Tour Scheduling'.

Lemp, J. D., K. M. Kockelman and P. Damien. (2010), 'The Continuous Cross-Nested Logit Model: Formulation and Application for Departure Time Choice', *Transportation Research Part B: Methodological* Vol. 44, No. 5, pp. 646-661.

———. (2012), 'A Bivariate Multinomial Probit Model for Trip Scheduling: Bayesian Analysis of the Work Tour', *Transportation Science* Vol. 46, No. 3, pp. 405-424.

Linetsky, V. (1997), 'The Path Integral Approach to Financial Modeling and Options Pricing', *Computational Economics* Vol. 11, No. 1, pp. 129-163.

Liu, Y.-H. and H. Mahmassani. (1998), 'Dynamic Aspects of Commuter Decisions under Advanced Traveler Information Systems: Modeling Framework and Experimental Results', *Transportation Research Record: Journal of the Transportation Research Board*, No. 1645, pp. 111-119.

Mathews, J. H. and K. D. Fink. (2004), *Numerical Methods Using Matlab*, Pearson London, UK:.

McFadden, D. (1976), 'The Mathematical Theory of Demand Models', *Behavioral Travel Demand Models*, pp. 305-314.

———. (1978), *Modelling the Choice of Residential Location*, Institute of Transportation Studies, University of California.

McFadden, D. and K. Train. (2000), 'Mixed Mnl Models for Discrete Response', *Journal of applied Econometrics* Vol. 15, No. 5, pp. 447-470.

Morokoff, W. J. and R. E. Caflisch. (1994), 'Quasi-Random Sequences and Their Discrepancies', *SIAM Journal on Scientific Computing* Vol. 15, No. 6, pp. 1251-1279.

———. (1995), 'Quasi-Monte Carlo Integration', *Journal of computational physics* Vol. 122, No. 2, pp. 218-230.

Nelsen, R. B. (2007), *An Introduction to Copulas*, Springer Science & Business Media.

Pawitan, Y. (2001), *In All Likelihood: Statistical Modelling and Inference Using Likelihood*, Oxford University Press.

Piessens, R., E. de Doncker-Kapenga, C. W. Überhuber and D. K. Kahaner. (2012), *Quadpack: A Subroutine Package for Automatic Integration*, Springer Science & Business Media.

Popuri, Y., M. Ben-Akiva and K. Proussaloglou. (2008), 'Time-of-Day Modeling in a Tour-Based Context: Tel Aviv Experience', *Transportation Research Record: Journal of the Transportation Research Board*, No. 2076, pp. 88-96.

RCore, T. (2012), 'R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria', (Ed.)^(Eds.), ISBN 3-900051-07-0, URL http://www.R-project.org.

Sener, I. N., N. Eluru and C. R. Bhat. (2010), 'On Jointly Analyzing the Physical Activity Participation Levels of Individuals in a Family Unit Using a Multivariate Copula Framework', *Journal of Choice Modelling* Vol. 3, No. 3, pp. 1-38.

Shreve, S. E. (2004), *Stochastic Calculus for Finance Ii: Continuous-Time Models*, Springer Science & Business Media.

Sklar, A. (1973), 'Random Variables, Joint Distribution Functions, and Copulas', *Kybernetika* Vol. 9, No. 6, pp. (449)-460.

Small, K. A. (1982), 'The Scheduling of Consumer Activities: Work Trips', *The American Economic Review*, pp. 467-479.

———. (1987), 'A Discrete Choice Model for Ordered Alternatives', *Econometrica: Journal of the Econometric Society*, pp. 409-424.

Spissu, E., A. R. Pinjari, R. M. Pendyala and C. R. Bhat. (2009), 'A Copula-Based Joint Multinomial Discrete–Continuous Model of Vehicle Type Choice and Miles of Travel', *Transportation* Vol. 36, No. 4, pp. 403-422.

Süli, E. and D. F. Mayers. (2003), *An Introduction to Numerical Analysis*, Cambridge university press.

Tang, L., S. Ghader, S. Mishra and L. Zhang. (2015), 'Impact of Travel Reliability on Mode Choice: Empirical Analysis with Observed Travel Time Variation Data in Washington, Dc', (Ed.)^(Eds.).

Torkjazi, M., P. S. Mirjafari and H. Poorzahedy. (2018), 'Reliability-Based Network Flow Estimation with Day-to-Day Variation: A Model Validation on Real Large-Scale Urban Networks', *Journal of Intelligent Transportation Systems* Vol. 22, No. 2, pp. 121-143.

Train, K. E. (2009), *Discrete Choice Methods with Simulation*, Cambridge university press.

Trivedi, P. K. and D. M. Zimmer. (2007), *Copula Modeling: An Introduction for Practitioners*, Now Publishers Inc.

Van Kampen, N. G. (1992), *Stochastic Processes in Physics and Chemistry*, Elsevier.

van Vuren, T., S. Carmichael, J. Polak, G. Hyman and S. Cross. (1999), 'Modelling Peak Spreading in Continuous Time', (Ed.)^(Eds.), *TRANSPORTATION PLANNING METHODS. PROCEEDINGS OF SEMINAR F, EUROPEAN TRANSPORT CONFERENCE, 27-29 SEPTEMBER 1999, CAMBRIDGE, UK.*

Vickrey, W. S. (1969), 'Congestion Theory and Transport Investment', *The American Economic Review*, pp. 251-260.

Vovsha, P. (1997), 'Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area', *Transportation Research Record: Journal of the Transportation Research Board*, No. 1607, pp. 6-15.

Walker, J. and M. Ben-Akiva. (2002), 'Generalized Random Utility Model', *Mathematical social sciences* Vol. 43, No. 3, pp. 303-343.

Walker, J. L. (2001), 'Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables', Massachusetts Institute of Technology.

Walker, J. L., M. Ben-Akiva and D. Bolduc. (2007), 'Identification of Parameters in Normal Error Component Logit-Mixture (Neclm) Models', *Journal of Applied Econometrics* Vol. 22, No. 6, pp. 1095-1125.

Wang, J. J. (1996), 'Timing Utility of Daily Activities and Its Impact on Travel', *Transportation Research Part A: Policy and Practice* Vol. 30, No. 3, pp. 189-206.

Wen, C.-H. and F. S. Koppelman. (2001), 'The Generalized Nested Logit Model', *Transportation Research Part B: Methodological* Vol. 35, No. 7, pp. 627-641.

Wio, H. S. (1990), *Introducción a Las Integrales De Camino: Curso De Tercer Ciclo*, Universitat de les Illes Balears, Secretariat de Publicacions i Intercanvi Científic.

———. (2013), *Path Integrals for Stochastic Processes: An Introduction*, World Scientific.

Wio, H. S., P. Colet, M. San Miguel, L. Pesquera and M. Rodriguez. (1989), 'Path-Integral Formulation for Stochastic Processes Driven by Colored Noise', *Physical Review A* Vol. 40, No. 12, pp. 7312.

Xiong, C. and L. Zhang. (2013), 'Positive Model of Departure Time Choice under Road Pricing and Uncertainty', *Transportation Research Record: Journal of the Transportation Research Board*, No. 2345, pp. 117-125.

Zeid, M., T. Rossi and B. Gardner. (2006), 'Modeling Time-of-Day Choice in Context of Tour-and Activity-Based Models', *Transportation Research Record: Journal of the Transportation Research Board*, No. 1981, pp. 42-49.

Zhang, L. (2007), 'Developing a Positive Approach to Travel Demand Forecasting: Theory and Model', *Traffic and Transportation Theory* Vol. 17, pp. 136-150.

Zhang, L., D. Yang, S. Ghader, C. Carrion, C. Xiong, T. F. Rossi, M. Milkovits, S. Mahapatra and C. Barber. (2018), 'An Integrated, Validated, and Applied Activity-Based Dynamic Traffic Assignment Model for the Baltimore-Washington Region', *Transportation Research Record*, pp. 0361198118796397.

Øksendal, B. (2003), 'Stochastic Differential Equations', *Stochastic Differential Equations*, Springer.