

Copyright

by

Jason David Lemp

2009

**The Dissertation Committee for Jason David Lemp certifies that this is the
approved version of the following dissertation:**

**CAPTURING RANDOM UTILITY MAXIMIZATION BEHAVIOR IN
CONTINUOUS CHOICE DATA: APPLICATION TO WORK TOUR
SCHEDULING**

Committee:

Kara M. Kockelman, Supervisor

Randy B. Machemehl

Paul Damien

Stephen Donald

S. Travis Waller

**CAPTURING RANDOM UTILITY MAXIMIZATION BEHAVIOR IN
CONTINUOUS CHOICE DATA: APPLICATION TO WORK TOUR
SCHEDULING**

by

Jason David Lemp, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December, 2009

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to Dr. Kara Kockelman for being an excellent mentor. Without her expert guidance and support over the past four years, this dissertation would not have been possible. I would also like to acknowledge preliminary and final dissertation committee members: Drs. Randy Machemehl, Paul Damien, Stephen Donald, Travis Waller, and Michael Walton. Many thanks are due to each for their participation and advice. I am especially indebted to Dr. Damien for his expert statistical guidance.

Special thanks are due to Ms. Annette Perrone for her encouragement and assistance with many logistical and administrative matters. I also wish to thank all members, past and present, of the University of Texas Transportation Program for their support over the past several years. I would like to acknowledge the financial support of the National Cooperative Highway Research Program (NCHRP), Project HR 08-57, and the Strategic Highway Research Program (SHRP) II, Project CO4, for allowing me to concentrate on research and academics during my PhD studies.

Lastly, I must acknowledge the support offered by my family, especially my wife, Sally. Her compassion and understanding, particularly over the several months leading to the completion of this dissertation, has been immeasurable. Thank you for always being there with your love and encouragement.

CAPTURING RANDOM UTILITY MAXIMIZATION BEHAVIOR IN CONTINUOUS CHOICE DATA: APPLICATION TO WORK TOUR SCHEDULING

Jason David Lemp, Ph.D.

The University of Texas at Austin, 2009

Supervisor: Kara M. Kockelman

Recent advances in travel demand modeling have concentrated on adding behavioral realism by focusing on an individual's activity participation. And, to account for trip-chaining, tour-based methods are largely replacing trip-based methods. Alongside these advances and innovations in dynamic traffic assignment (DTA) techniques, however, time-of-day (TOD) modeling remains an Achilles' heel. As congestion worsens and operators turn to variable road pricing, sensors are added to networks, cell phones are GPS-enabled, and DTA techniques become practical, accurate time-of-day forecasts become critical. In addition, most models highlight tradeoffs between travel time and cost, while neglecting variations in travel time. Research into stated and revealed choices suggests that travel time variability can be highly consequential.

This dissertation introduces a method for imputing travel time variability information as a continuous function of time-of-day, while utilizing an existing method for imputing

average travel times (by TOD). The methods employ ordinary least squares (OLS) regression techniques, and rely on reported travel time information from survey data (typically available to researchers), as well as travel time and distance estimates by origin-destination (OD) pair for free-flow and peak-period conditions from network data.

This dissertation also develops two models of activity timing that recognize the imputed average travel times and travel time variability. Both models are based in random utility theory and both recognize potential correlations across time-of-day alternatives. In addition, both models are estimated in a Bayesian framework using Gibbs sampling and Metropolis-Hastings (MH) algorithms, and model estimation relies on San Francisco Bay Area data collected in 2000.

The first model is the continuous cross-nested logit (CCNL) and represents tour outbound departure time choice in a continuous context (rather than discretizing time) over an entire day. The model is formulated as a generalization of the discrete cross-nested logit (CNL) for continuous choice and represents the first random utility maximization model to incorporate the ability to capture correlations across alternatives in a continuous choice context. The model is then compared to the continuous logit, which represents a generalization of the multinomial logit (MNL) for continuous choice. Empirical results suggest that the CCNL out-performs the continuous logit in terms of predictive accuracy and reasonableness of predictions for three tolling policy simulations. Moreover, while this dissertation focuses on time-of-day modeling, the CCNL could be used in a number of other continuous choice contexts (e.g., location/destination, vehicle usage, trip durations, and profit-maximizing production).

The second model is a bivariate multinomial probit (BVMNP) model. While the model relies on discretization of time (into 30-minute intervals), it captures both key dimensions of a tour's timing (rather than just one, as in this dissertation's application of the CCNL model), which is important for tour- and activity-based models of travel demand. The

BVMNP's ability to capture correlations across scheduling alternatives is something no existing two-dimensional choice models of tour timing can claim.

Both models represent substantial contributions for continuous choice modeling in transportation, business, biology, and various other fields. In addition, the empirical results of the models evaluated here enhance our understanding of individuals' time-of-day decisions. For instance, average travel time and its variance are estimated to have a negative effect on workers' utilities, as expected, but are not found to be that practically relevant here, probably because most workers are rather constrained in their activity scheduling and/or work hours. However, correlations are found to be rather strong in both models, particularly for home-to-work journeys, suggesting that if models fail to accommodate such correlations, biased application results may emerge.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION	1
1.1 Overview and Motivation.....	1
1.2 Existing Methods of Activity Scheduling	3
1.2.1 Discrete Methods.....	3
1.2.2 Continuous Methods.....	4
1.3 Limitations of Existing Methods	5
1.4 Advantages of Bayesian Techniques in TOD Modeling.....	7
1.5 Study Objectives	9
1.6 Organization	11
1.7 Chapter Summary	12
CHAPTER 2: LITERATURE REVIEW.....	13
2.1 Travel Timing Models.....	13
2.1.1 Discrete Choice Models.....	13
2.1.2 Continuous Models.....	18
2.1.3 Time-of-Day Modeling Summary	20
2.2 Reliability Measures	21
2.2.1 Travel Time Distribution	21
2.2.2 Schedule Delay Methods	23
2.3 Bayesian Statistics	24
2.3.1 Bayesian Theory.....	24
2.3.2 Prior Choice	25
2.3.3 MCMC Simulation	26
2.3.4 Convergence Assessment.....	32
2.4 Chapter Summary	34
CHAPTER 3: TIME-OF-DAY MODELING METHODS	35
3.1 Continuous Logit	36
3.1.1 Continuous Logit Specification.....	36

3.1.2 Continuous Logit Parameter Estimation via MCMC Simulation	39
3.2 Continuous Cross-Nested Logit.....	41
3.2.1 (Discrete) Cross-Nested Logit.....	41
3.2.2 Continuous Cross-Nested Logit	42
3.2.3 CCNL Model Behavior and Properties.....	46
3.2.4 CCNL Parameter Estimation via MCMC Simulation	56
3.3 Bivariate Multinomial Probit.....	58
3.3.1 Random Utility Framework and Model Specification	59
3.3.2 Error Correlation Structure	62
3.3.3 BVMNP Parameter Estimation via MCMC Simulation.....	69
3.4 Chapter Summary	73
CHAPTER 4: IMPUTING TIME-VARYING NETWORK VARIABLES	75
4.1 Methodology.....	75
4.1.1 Automobile Travel Times	76
4.1.2 Automobile Travel Time Variability	78
4.1.3 Transit Level-of-Service Attributes.....	80
4.2 Data Description	82
4.3 Empirical Results	88
4.4 Chapter Summary	97
CHAPTER 5: EMPIRICAL RESULTS OF CONTINUOUS TOD MODELS	98
5.1 Work-Tour Departure Time Data	98
5.2 Model Estimation Details	101
5.3 Empirical Findings.....	102
5.3.1 Out-of-Sample Predictions.....	110
5.3.2 Economic Welfare Demonstration	112
5.4 Chapter Summary	122
CHAPTER 6: EMPIRICAL RESULTS OF DISCRETE TOD MODELS	123
6.1 Work-Tour Scheduling Data	123
6.2 Model Estimation Details	125

6.3 Empirical Findings.....	131
6.3.1 Individual-Specific Covariate Effects.....	136
6.3.2 Out-of-Sample Predictive Performance.....	142
6.3.3 Economic Welfare Demonstration	146
6.4 Chapter Summary	155
CHAPTER 7: CONCLUSION	157
7.1 Summary	157
7.2 Opportunities for Future Research.....	162
7.3 Concluding Remarks	165
APPENDIX A: R CODE FOR CCNL MODEL ESTIMATION	167
APPENDIX B: R CODE FOR BVMNP MODEL ESTIMATION.....	176
REFERENCES.....	186
VITA.....	198

LIST OF TABLES

Table 3.1: Correlation Coefficients for Varying ρ and Distance Variables	49
Table 4.1: Descriptive Statistics for Variables Used in Automobile Mode Regressions.....	85
Table 4.2: Descriptive Statistics for Variables Used in Drive-to-Transit Mode Regression	86
Table 4.3: Descriptive Statistics for Variables Used in Walk-to-Transit Mode Regression	87
Table 4.4: Automobile Mode Speed Regression Model Estimates for Departure Time- and Arrival Time-Based Models	89
Table 4.5: Automobile Mode Travel Time Variance Regression Model Estimates for Departure Time- and Arrival Time-Based Models.....	93
Table 4.6: Transit Mode Travel Time Variance Regression Model Estimates	96
Table 5.1: Descriptive Statistics of Explanatory Variables Used in Continuous TOD Models.....	100
Table 5.2: Continuous Logit and CCNL Model Estimation Results.....	105
Table 5.3: Continuous Logit and CCNL Model Estimation Results (Cont'd)	106
Table 5.4: Predicted Departure Time Proportions for Five TOD Periods and Four Simulations.....	120
Table 6.1: BVMNP Parameter Estimation Results for Arrival Time Variables.....	133
Table 6.2: BVMNP Parameter Estimation Results for Return Time Variables	134
Table 6.3: BVMNP Parameter Estimation Results for Covariance Parameters	135
Table 6.4: MSEs of MNL and BVMNP Model Predictions.....	145
Table 6.5: Predicted Arrival Time Proportions for Three TOD Periods and Four Simulations under CAR, AR1, and MNL Model Specifications	150
Table 6.6: Predicted Return Time Proportions for Three TOD Periods and Four Simulations under CAR, AR1, and MNL Model Specifications	153

LIST OF FIGURES

Figure 3.1: Allocation Parameter Illustration	45
Figure 3.2: Utility Profiles and 90% Confidence Intervals for Examples 1 and 2	50
Figure 3.3: Predictive Densities with h of 0.5 and Varying Values of ρ for Example 1	52
Figure 3.4: Predictive Densities with h of 1.0 and Varying Values of ρ for Example 1	53
Figure 3.5: Predictive Densities with h of 2.0 and Varying Values of ρ for Example 1	54
Figure 3.6: Predictive Densities with h of 2.0 and Varying Values of ρ for Example 2	56
Figure 4.1: Distributions of Dependent Variables for (a) Automobile Speeds, (b) Drive-to-Transit Variances, and (c) Walk-to-Transit Variances	88
Figure 4.2: Delay Coefficient Variation in Automobile Mode Speed Regressions across Departure and Arrival Times	90
Figure 4.3: Automobile Mode Free-Flow Speed Coefficient Variation for Different Delay Variable Values	91
Figure 4.4: Delay Coefficient Variation in Automobile Mode Travel Time Variance Regressions across Departure and Arrival Times.....	94
Figure 4.5: Automobile Mode Travel Time Variance Variation for Different Delay Variable Values and Mean Time-Invariant Covariates	95
Figure 5.1: Estimated Density of Outbound Departure Time.....	101
Figure 5.2: Gender, Age, Worker Status, and Income Effects on Average Individuals' Predictive Densities for Continuous Logit and CCNL.....	107
Figure 5.3: Household Size, Number of Tours, Travel Distance, and CBD Effects on Average Individuals' Predictive Densities for Continuous Logit and CCNL.....	109
Figure 5.4: Out-of-Sample Likelihood Predictions for Continuous Logit and CCNL...	111

Figure 5.5: Distribution of Consumer Surplus Change for Continuous Logit and CCNL under Three Tolling Policy Simulations.....	115
Figure 5.6: Distribution of Consumer Surplus Change for Three Tolling Policy Simulations under Continuous Logit and CCNL Specifications.....	116
Figure 5.7: Distribution of Travelers' Departure Time Choices for Four Simulations under Continuous Logit and CCNL.....	117
Figure 5.8: Distribution of Travelers' Departure Time Choices for Continuous Logit and CCNL by Simulation Exercise.....	118
Figure 5.9: AM Peak Period Arrival Time Shifts to Shoulder and Off-Peak Periods for Three Tolling Policy Simulations under Continuous Logit and CCNL Models.....	121
Figure 6.1: Densities of Home-to-Work Arrival Times and Work-to-Home Departure Times	124
Figure 6.2: Trace Plot of Parameter Draws versus Iteration Number for Four Selected CAR Model Parameters (Every 200 th Draw from Last 100,000).....	129
Figure 6.3: Trace Plot of Parameter Draws versus Iteration Number for Four Selected AR1 Model Parameters (Every 200 th Draw from Last 100,000)	130
Figure 6.4: Gender, Age, Work Status, and Income Effects on Arrival Time Profiles for BVMNP Models.....	137
Figure 6.5: Household Size, Other Tours, Travel Distance, and CBD Effects on Arrival Time Profiles for BVMNP Models	138
Figure 6.6: Gender, Age, Work Status, and Income Effects on Return Time Profiles for BVMNP Models.....	140
Figure 6.7: Household Size, Other Tours, Travel Distance, and CBD Effects on Return Time Profiles for BVMNP Models	141
Figure 6.8: Out-of-Sample Predictions for MNL versus BVMNP Models.....	144
Figure 6.9: Consumer Surplus Change Distribution for CAR and AR1 Specifications under Three Tolling Policy Simulations	147

Figure 6.10: Consumer Surplus Change Distribution for Three Tolling Policy
Simulations under CAR, AR1, and MNL Model Specifications 149

Figure 6.11: AM Peak Period Arrival Time Shifts to Shoulder and Off-Peak Periods
for Three Tolling Policy Simulations under CAR, AR1, and MNL Models 152

Figure 6.12: PM Peak Period Return Time Shifts to Shoulder and Off-Peak Periods
for Three Tolling Policy Simulations under CAR, AR1, and MNL Models 155

NOTATION

The following summarizes the notation used in this dissertation, grouped by their chapter of first appearance.

Chapter 2:

$p(\theta|Y, X)$: The posterior distribution of parameters

$p(Y|\theta, X)$: The likelihood function

$p(\theta)$: The prior distribution of parameters

P : A Markov chain transition matrix

$q(\theta^n|\theta^{n-1})$: Proposal density for Metropolis-Hastings algorithm

X : A matrix of explanatory variables

Y : A vector of dependent variables

z_G : Geweke's (1992) convergence diagnostic

$s(a, b)$: Metropolis-Hastings proposal acceptance probability of transitioning from state a to state b

θ : A vector of parameters

π_0 : Initial distribution of θ in Markov chain

π : Unique stationary distribution of Markov chain

Chapter 3:

b_1 and b_2 : Choice bounds for continuous choice model

c_{pq} : BVMNP's degree of closeness measure between implied duration of arrival alternative p and return alternative q , relative to a baseline/preferred duration

\mathcal{D}_m : Subset of alternatives in nest m

$F(\varepsilon_k)$: Cumulative density function for random error term ε_k

$F(\varepsilon_k, \varepsilon_j)$: Joint cumulative density function for random error terms ε_k and ε_j

$g_{ip}(t)$: Time-varying level-of-service variable p for individual i evaluated at time t

$G(\cdot)$: GEV model generating function
 h : CCNL's nest size parameter
 I_j : Identity matrix of dimension J
 J : Number of discrete alternatives for choice model
 \mathcal{J} : BVMNP's set of arrival/return time alternatives
 λ : BVMNP's covariance deflation factor for drawing utilities in model estimation
 K : Number of individual-specific variables in utility functions
 M : Number of nests in discrete cross-nested logit
 p_{t_k} : Choice density at time t_k
 P_k : Probability of choice k in discrete choice model
 \mathcal{P} : Number of time-varying level-of-service variables
 Q : Number of cyclical interaction terms for utility functions
 s : Distance between discrete alternatives (where alternatives are ordered intervals)
 t : Continuous time variable
 U_j : Random (latent) utility of alternative j
 $U_i(t)$: Time-varying (over time t) random (latent) utility function for individual i
 U_{aj} : BVMNP's random (latent) arrival time-specific utility for arrival time alternative j
 U_{rl} : BVMNP's random (latent) return time-specific utility for return time alternative l
 V_j : Systematic utility of alternative j
 $V_i(t)$: Time-varying (over time t) systematic utility function for individual i
 V_{aj} : BVMNP's systematic arrival time-specific utility for arrival time alternative j
 V_{rl} : BVMNP's systematic return time-specific utility for return time alternative l
 w_{jq} : BVMNP's degree of closeness measure between arrival or return specific
alternatives j and q
 \mathcal{X}_{ij} : BVMNP's row vector of individual-specific attributes interacted with cyclical
functions of arrival time j for individual i
 \mathcal{X}_{il} : BVMNP's row vector of individual-specific attributes interacted with cyclical
functions of return time l for individual i

x_i : BVMNP's matrix of explanatory variables for individual i
 $y_j = e^{V_j}$: Exponent of alternative j 's systematic utility
 α_{jm} : CNL's allocation parameter for alternative j in nest m
 $\alpha(r, q)$: CCNL's allocation parameter for time choice r in time q 's nest
 β : Vector of parameters in utility functions
 β_a : BVMNP's vector of utility function parameters for arrival time variables
 β_r : BVMNP's vector of utility function parameters for return time variables
 $\bar{\beta}$: Prior mean for BVMNP's β
 $\gamma(t)$: Collection of cyclical functions evaluated at time t , interacted with time-invariant covariates for continuous choice models
 ε_j : Random error component for alternative j
 $\varepsilon_i(t)$: Time-varying (over time t) random error components for individual i
 ε_{aj} : BVMNP's random error arrival-specific error term for arrival time alternative j
 ε_{rl} : BVMNP's random error return-specific error term for return time alternative l
 ζ : Euler's constant
 η_p : Utility function parameter on time-varying level-of-service variable p
 η_a : BVMNP's vector of utility function parameters for time-varying arrival time variables
 η_r : BVMNP's vector of utility function parameters for time-varying return time variables
 κ_{aj} : BVMNP's random noise term for arrival alternative j
 κ_{rl} : BVMNP's random noise term for return alternative l
 $\lambda_{CAR,a}$: BVMNP's arrival-specific correlation parameter for CAR covariance specification
 $\lambda_{CAR,r}$: BVMNP's return-specific correlation parameter for CAR covariance specification
 $\lambda_{CAR,d}$: BVMNP's duration-specific correlation parameter for CAR covariance specification

- $\lambda_{AR1,\alpha}$: BVMNP's arrival-specific correlation parameter for AR1 covariance specification
- $\lambda_{AR1,r}$: BVMNP's return-specific correlation parameter for AR1 covariance specification
- $\lambda_{AR1,d}$: BVMNP's duration-specific correlation parameter for AR1 covariance specification
- Λ : Vector parameter for posterior distribution of BVMNP's β
- $\mu_{CAR,1}$: Baseline duration parameter 1 for CAR covariance specification
- $\mu_{CAR,2}$: Baseline duration parameter 2 for CAR covariance specification
- $\mu_{AR1,1}$: Baseline duration parameter 1 for AR1 covariance specification
- $\mu_{AR1,2}$: Baseline duration parameter 2 for AR1 covariance specification
- ρ_m : CNL's inclusive value parameter for nest m
- ρ : CCNL's inclusive value parameter
- σ_a^2 : BVMNP's variance for arrival alternative utilities
- σ_r^2 : BVMNP's variance for return alternative utilities
- Σ_{CAR} : BVMNP's utility covariance matrix for CAR covariance specification
- Σ_{AR1} : BVMNP's utility covariance matrix for AR1 covariance specification
- Σ_{full} : BVMNP's utility covariance matrix for full-time workers making no additional tours
- Σ_{part} : BVMNP's utility covariance matrix for part-time workers and/or those making additional tours
- Σ_β : Prior covariance matrix for BVMNP's β
- Ω : Covariance matrix parameter for posterior distribution of BVMNP's β

Chapter 4:

- $(Delay)_q$: Delay variable for origin-destination pair q
- $(IVTT_{skim})_q$: Transit in-vehicle travel time for origin-destination pair q
- $(TTT_{skim})_q$: Transit total travel time for origin-destination pair q

X_{iq} : Row vector of explanatory variables in speed and variance regressions for trip i made between origin-destination pair q

ϵ_i : Random error term for trip i in regression equations

ι_1 and ι_2 : Parameters related to delay variable in speed and variance regressions

ι_3 and ι_4 : Parameters related to travel times in transit variance regressions

v_1 and v_2 : Vectors of parameters related to trip-specific explanatory variables in speed and variance regressions

v_3 and v_4 : Vectors of parameters related to trip-specific explanatory variables in transit variance regressions

$\psi_j(t)$: The j^{th} cyclical function interacted with delay variable

ω_{1jl} and ω_{2jl} : Parameters related to cyclical function j taken to the power of l and interacted with the delay variable in speed and variance regressions

Chapter 6:

size_p : Size (in minutes) of choice alternative p

size_{def} : Size (in minutes) of the default choice alternatives

$\sigma_{CAR,a,\text{first}}^2$: Variance parameter for first arrival time (boundary) alternative for CAR specification

$\sigma_{CAR,a,\text{last}}^2$: Variance parameter for last arrival time (boundary) alternative for CAR specification

$\sigma_{CAR,r,\text{first}}^2$: Variance parameter for first return time (boundary) alternative for CAR specification

$\sigma_{CAR,r,\text{last}}^2$: Variance parameter for last return time (boundary) alternative for CAR specification

$\sigma_{AR1,a,\text{first}}^2$: Variance parameter for first arrival time (boundary) alternative for AR1 specification

$\sigma_{AR1,a,\text{last}}^2$: Variance parameter for last arrival time (boundary) alternative for AR1 specification

$\sigma_{AR1,r,first}^2$: Variance parameter for first return time (boundary) alternative for AR1 specification

$\sigma_{AR1,r,last}^2$: Variance parameter for last return time (boundary) alternative for AR1 specification

$\tau_{CAR,ar}$: BVMNP's size parameter relating to arrival and return correlations for CAR specification

$\tau_{CAR,d}$: BVMNP's size parameter relating to duration correlations for CAR specification

$\tau_{AR1,ar}$: BVMNP's size parameter relating to arrival and return correlations for AR1 specification

$\tau_{AR1,d}$: BVMNP's size parameter relating to duration correlations for AR1 specification

CHAPTER 1: INTRODUCTION

1.1 Overview and Motivation

Travel demand forecasting has been widely used over the past 40 years. The earliest models of travel demand sought to inform long-term infrastructure investments, rendering models of trip timing less important than trip generation, mode, and destination choices. But the focus of decision-making has shifted from long-term capital investments to shorter-term policies, such as congestion management, promotion of alternative transport modes, and demand management (Bhat and Koppelman 1999). This is due in part to the onset of environmental and transportation legislation, like the 1990 Clean Air Act Amendment and the Intermodal Surface Transportation Efficiency Act (ISTEA), but also due to rising financial costs of such investments, as practical space limits are reached in many urban areas.

In addition, our understanding of travel behavior and the tools available for computing have progressed tremendously since the earliest models. As policy decisions and tools have changed, so too have the questions posed to our travel demand forecasting models, and behavioral travel theory has become an increasingly important component. With improvements in dynamic traffic assignment (DTA) techniques and applications (see, e.g., Hobeika 2005, Lin et al. 2008, and DYNASMART-P 2009), better models of travel timing are needed, compatible with the relatively fine time resolution used in DTA methods. Even with recent advances in travel demand theory and DTA techniques, models of activity scheduling, a fundamental aspect of any trip or tour, are typically greatly simplified in model specifications and applications, particularly when compared to the treatment of mode and destination choices. In fact, Vovsha et al. (2005) and TRB (2007) point to time-of-day (TOD) modeling as a major weakness of most travel demand model systems.

In addition to the need for better models of activity timing, the importance of travel time variability and its affect on travel behavior is becoming more prominent. Most models emphasize travel time and travel cost, while ignoring travel time variability, though it can be highly consequential (Vovsha et al. 2005). In some instances, estimates of the value of travel time reliability (measured as the difference between the 80th or 90th and 50th percentile travel times) exceed the value-of-travel-time (see, e.g., Lam and Small 2001, Small et al. 2005, and Bhat and Sardesai 2006), suggesting that travel time reliability considerations are very important for forecasting travel, along with average travel times. DTA techniques offer a natural method for capturing travel time variability in models.

These issues are of great consequence when examining road pricing policies, which have become more common in recent years (Schofer 2005). Vovsha et al. (2005) suggests that accounting for reliability and dealing with individuals' TOD choices are among the most important modeling challenges in road pricing policy evaluation (as well as accounting for heterogeneity among users' values-of-travel-time). For peak period pricing policies, TOD variations are of particular importance, since such policies are generally intended to shift travelers' TOD choices to off-peak or shoulder periods. Given the current modeling weaknesses in these areas, it is not surprising that toll road traffic and revenue forecasts suffer from great amounts of uncertainty, and are often biased high¹ (see, e.g., Bain and Wilkins 2002, George et al. 2003, Bain and Plantagie 2003 and 2004, Bain and Polakovic 2005, Lemp and Kockelman 2009). In order to accommodate these uncertainties in traffic and revenue forecasts, credit agencies often take a conservative approach, reducing growth rate expectations and carefully examining future toll schedule increases (Bain et al. 2006). While understanding and accommodating uncertainty in forecasts is certainly necessary, improving the quality of forecasts is also important.

¹ For instance, the average ratio of actual-to-forecast traffic volumes for the first year of toll road operation is about 0.77 (Bain and Polakovic 2005). In other words, forecasts tend to over-predict traffic volumes by about 30%, on average.

Of course, modeling activity timing has other important applications as well. For instance, activity timing models could help transit agencies to optimally schedule transit vehicles. Understanding the times in which people are most likely to choose to travel as well as the effect of transit quality of service (e.g., access, egress, wait, and travel times, costs, and reliability) with respect to competing transport modes (like the automobile) has important consequences in constructing transit vehicle schedules.

This dissertation develops models of activity timing that allow travel time variability (as well as other level-of-service variables) to be incorporated into the model specification, and each model is developed in random utility theory. The first two models (the continuous logit and continuous cross-nested logit [CCNL]) consider time-of-day choice in a continuous choice context, while the third model (the bivariate multinomial probit [BVMNP]) considers discrete time intervals. However, the BVMNP considers the two-dimensional activity scheduling context of a travel tour, rather than the one-dimensional context of the continuous logit and CCNL. Model specifications are discussed in great detail in Chapter 3. In the following section, existing methods for modeling activity scheduling behavior are introduced.

1.2 Existing Methods of Activity Scheduling

Existing methods for modeling activity scheduling behavior can essentially be broken into two subclasses: those that rely on discretization of the TOD element and those that treat it as a continuous response variable. Here, a brief summary of these two methods is offered. Chapter 2 details the methods in greater detail.

1.2.1 Discrete Methods

Supporting applications of random utility maximization theory, the GEV class of models (McFadden 1978) has become a mainstay in travel behavior analysis of discrete choice behaviors. McFadden's (1973) multinomial logit (MNL) model represents the most familiar and straightforward of these models. However, the MNL suffers from the

independence of irrelevant alternatives (IIA) property, which results in equivalent cross-elasticities across each pair of choice alternatives².

The nested logit model (Williams 1977, McFadden 1978, and Daly and Zachary 1979) relaxes this assumption, allowing correlations to emerge across similar alternatives. However, choice alternatives in common nests still retain the IIA property. Vovsha (1997) introduced the cross-nested logit (CNL) model (later generalized by Ben-Akiva and Bierlaire [1999], Wen and Koppelman [2001] and Papola [2004]³), which allows choice alternatives to appear in multiple nests, thereby, offering more flexible correlation patterns than the nested logit. Small's (1987) ordered GEV (OGEV) model represents a special case of the CNL, where alternatives are ordered in nature (e.g., departure time choice). Thus, each nest in the OGEV contains consecutive alternatives in a sequence.

While the models described above all have important applications in travel behavior research, a number of travel-related decisions are inherently continuous in nature, including TOD choice⁴. Of course, one can discretize these choices and employ discrete choice methods. In fact, this is often done in activity scheduling research (see, e.g., Abkowitz 1981, Small 1982 & 1987, Hendrickson and Plank 1984, Chin 1990, Vovsha 1997, Bhat 1998a & b, Steed and Bhat 2000, Vovsha and Bradley 2004, Abou Zeid et al. 2006, Popuri et al. 2008, among others).

1.2.2 Continuous Methods

Other research has treated departure time as a continuous response variable. The usual approach involves a hazard function (see, e.g., Wooldridge 2002). The hazard function defines the probability that an agent will leave its current state at some particular time.

² In other words, if one alternative's attributes improve (e.g., travel time under the alternative decreases), the probability of that alternative draws probability away from each other alternative equally (in a relative sense).

³ These generalizations reformulated the model to allow each nest to have its own nesting (or inclusive value) parameter.

⁴ Other examples include residential location choice, travel destination choice, and vehicle usage.

There are a number of hazard function forms, including the proportional hazard specification, non-parametric hazard forms, and accelerated hazard functions (see, e.g., Bhat and Pinjari 2008).

Since continuous methods do not require discretization of the time continuum, they may be preferable in certain applications. For instance, DTA may require departure time data at a very fine temporal resolution, something that continuous methods can provide. Most of the existing activity scheduling applications of continuous time methods have focused on activity duration (e.g., Ettema et al. 1995, Niemeier and Morita 1996, Bhat 1996, Yee and Niemeier 2000, Srinivasan and Guo 2003, Lee and Timmermans 2007, among others), though there has been limited application of such models in a departure time analysis setting (e.g., Wang 1996, Bhat and Steed 2002, Komma and Srinivasan 2008, and Gadda et al. 2009).

1.3 Limitations of Existing Methods

While there are many advantages to the discrete and continuous response models described above, all suffer from some limitations. Bhat and Steed (2002) indicate a number of weaknesses of discrete choice methods. First, discrete choice models require that interval boundaries be set, which is usually done in an arbitrary manner. If interval boundaries are changed, different model results emerge⁵. Second, discretization always results in points that are very close lying on either side of an interval boundary. Such points may be viewed as very similar options from the decision-maker's perspective, but cannot be treated adequately with discrete choice methods. Allowing for correlation across alternatives in discrete models can alleviate this issue to some extent, but continuous treatment of departure time seems preferable. Discrete treatment of time also results in a loss of temporal resolution, and, as DTA methods become more common,

⁵ In a departure time setting, for instance, one could build an AM peak interval (from 6am to 9am). If instead, this interval was split into two separate intervals (e.g., one from 4:30am to 7:30am and another from 7:30am to 10:30am), one could not presume application results (and certainly not calibration results) would be the same.

value of such models is diminished. Smaller intervals can alleviate this difficulty to some extent, but issues will remain unless continuous models are pursued.

In addition, to the inherent weaknesses of using discrete choice methods for continuous response variables, there are four other key issues to recognize. First, some models have taken the approach of modeling TOD choice for only a small portion of the day (e.g., the AM peak period). There are certainly advantages to taking this approach, but such models may not be applicable in large-scale travel demand model systems. Second, models that employ relatively large TOD choice alternatives over the entire day choice context (e.g., early morning, AM peak, midday, PM peak, and evening) may not be so valuable either. Third, as activity-based modeling has taken center stage over the last several years, there is a need for models that offer a two-dimensional choice construct (since travel tours have [at least] two timing dimensions). And finally, there is a need for models that allow correlations to emerge across similar alternatives. Clearly, one would expect correlations to be present across adjacent discrete time intervals. Currently there are no such applications in which all four of these ideas are integrated cohesively. While there have been applications to consider each of the first three of these considerations (e.g., 30-minute or 1-hour time intervals over the entire day for two time choice dimensions [see, e.g., Vovsha and Bradley 2004, Abou Zeid et al. 2006, and Popuri et al. 2008]), there have been none to integrate the correlations and two timing dimensions. And, those that have incorporated correlations have only done so over a limited temporal context or have used broad TOD choice alternatives (see, e.g., Small 1987, Chin 1990, Bhat 1998a, and de Jong et al. 2003).

Although discretization of continuous response variables may be inappropriate, one of the main advantages of the GEV class of models is that they are based in random utility theory. Random utility theory is a mainstay of travel choice modeling and provides a solid foundation for estimation of economic welfare, which can be used for project evaluation and policy analysis. In addition, it can provide meaningful relationships

between activity scheduling behavior and other traveler choices (e.g., destination, mode, and route).

Existing continuous response models are not based in random utility theory, and thus, offer none of the advantages of such models. In addition, there do not appear to be any applications of such models incorporating the two timing dimensions of a travel tour. However, continuous models do not suffer from the interval boundary issues of discrete choice models, and offer no loss in temporal resolution. In the next section, some strengths of Bayesian statistical techniques are explored.

1.4 Advantages of Bayesian Techniques in TOD Modeling

Since Bayesian statistics are used to estimate the TOD models developed in this dissertation, this section highlights comparisons between these methods relative to classical techniques.

Bayesian statistics offer more straightforward parameter estimation and interpretation as compared to classical (or frequentist) statistics (Wagenmakers et al. 2008). For instance, Bayesian methods avoid reliance on sample size asymptotics (needed with typical maximum likelihood estimation) and offer richer distributional inference of estimated parameters (e.g., the multivariate distributions of all parameters are estimated rather than point estimates of means and covariance). In addition, if the likelihood function is multimodal, classical maximum likelihood methods face issues relating to local maxima convergence. Since Bayesian methods do not search for parameter values that attain the maximum likelihood (draws from the posterior distribution are taken instead), this is not an issue (Huber and Train 2001). Of course, assessing convergence with Bayesian methods can be problematic, since parameter draws converge to a distribution rather than a point (Train 2009).

Bayesian methods are typically more flexible than classical techniques since they rely on Markov chain Monte Carlo (MCMC) simulation with conditional distributions for model estimation. Reliance on these conditional distributions allows models to be decomposed into smaller sub-problems that are often simpler to manage. Moreover, Bayesian methods can estimate non-trivial variables of interest, such as latent utilities. Classical methods generally must deal with the whole model directly, and such latent variables must be integrated out to obtain unconditional likelihood distributions. Often such integration is analytically and/or numerically complex.

Bayesian methods also offer the opportunity for the analyst to specify prior beliefs regarding model parameters. In practice, however, it is often useful to assume little prior knowledge about the parameters, but in some circumstances, the specification of the prior can be critical. In this dissertation, very little prior knowledge will generally be assumed.

Also, hierarchical modeling in Bayesian statistics offers great advantages over classical techniques. While hierarchical methods are not examined in this dissertation, it is useful to understand their value in TOD modeling. Vovsha et al. (2005) notes the importance of understanding the heterogeneity in users' values-of-travel-time. Bayesian methods can handle such heterogeneity with ease through hierarchical model specification of relevant parameters (i.e., those related to cost, travel time, and/or reliability), while classical techniques generally must rely on maximum simulated likelihood techniques. Of course, to achieve simplicity in such hierarchical models with Bayesian methods, the analyst is often limited to a narrower subset of distributional assumptions than with classical methods (Huber and Train 2001). Nonetheless, other distributional assumptions can be used with Bayesian methods; they may simply complicate the analysis.

Finally, Bayesian estimation offers clear opportunities for examining uncertainty in predictions. This is because Bayesian estimation results in obtaining the actual distribution of relevant parameters, rather than point estimates. Uncertainty in traffic

forecasts (as pointed out in Section 1.1) is of great importance, especially for toll road projects where traffic directly leads to revenue. While one may be able to make informed guesses about the distribution of parameters using classical methods (e.g., via confidence intervals), such approximations would not be ideal, particularly when parameters exhibit high degrees of correlation among one another. In the following section, the main objectives of this dissertation are outlined.

1.5 Study Objectives

There are five key objectives for the models of activity scheduling developed in this dissertation. First, and foremost, the model should be developed within the random utility theory construct. While this is rather limiting in terms of the types of models one may conceive (e.g., no continuous models of TOD choice have been developed in random utility theory), it offers a number of advantages. For instance, it forms a basis for which consumer surplus and economic welfare calculations can be performed. In addition, it offers opportunities to relate TOD choice to other travel choices within a large-scale travel demand model system. The models developed in this dissertation each have a basis in random utility theory.

The second objective is to develop models that recognize the continuous nature of activity timing adequately and across the entire 24-hour day period. While this does not necessitate the treatment of activity timing in a continuous context, time intervals should be sufficiently small to capture relatively small changes in time choice behavior. This is particularly important from a practical standpoint. As the policies being investigated have become more focused on demand and congestion management (particularly peak-period tolling policies), TOD choices occurring on relatively small scales have become increasingly important to recognize in model systems. Moreover, the advancement of DTA techniques relies (at least in some part) on the ability to capture relatively small shifts in demand over time.

The third objective of this dissertation is to develop a model that recognizes correlations across very similar TOD choice alternatives. Certainly one would expect adjacent 15-minute time interval alternatives to exhibit some degree of correlation, but in practice, correlations are often ignored to simplify the model and streamline estimation procedures. Realistically, such specifications assume that all of the dependence between alternatives is captured within alternative-specific utility specifications, which is unreasonable.

The fourth objective is to develop a model of tour scheduling behavior. A tour has (at least) two dimensions of travel timing (i.e., one for the outbound journey of travel and one for the inbound journey). Thus, the model should accommodate simultaneous choice of two alternatives. Often in practice, joint choices such as these are handled by describing the model in a single dimension and representing each model alternative as a pair of actual alternatives. Of course, in a two-dimensional activity timing model with small time intervals across the entire day, the number of alternatives to consider increases dramatically in two dimensions. For instance, if 30-minute intervals are considered, there are 48 alternatives in a single dimension and 1,176 in two dimensions (restricting joint alternatives to ones where inbound time is identical to or after outbound time). With certain model specifications, handling large numbers of alternatives is not necessarily so problematic (e.g., the multinomial logit), but with more complex models, large numbers of alternatives may be prohibitive. Thus, special techniques will be needed here.

Finally, the model should include transportation level-of-service variables. Since data on such variables is generally only available through network skims relevant to broad TOD periods (such as peak and off-peak), special methods are needed to thoughtfully impute these variables as they vary continuously over time. And, since data on reliability is generally not available at all, a new method for imputing travel time (un)reliability is introduced (which also varies continuously over time). For the automobile mode, simple regressions (a la Cambridge Systematics 2005, Abou Zeid et al. 2006, Popuri et al. 2008,

and Komma and Srinivasan 2008) are estimated to obtain time-varying travel times and reliability measures. For the transit mode, similar regressions are estimated to obtain reliability measures, though these measures do not vary continuously over time for the transit mode. Each of the activity timing models presented in this dissertation utilizes these imputed level-of-service (LOS) measures.

Three models of travel timing are introduced in this dissertation, though the first is only new in the context of activity scheduling and represents a special case of the second. The first two models are formed in random utility maximization theory, in a continuous choice context, and employ the imputed network LOS variables. In addition, the second model allows for correlations across similar alternatives. However, neither model is formulated in the context of tour scheduling decisions. The final model meets objectives one, three, four, and five, but represents time as 30-minute choice alternatives, which is not ideal, but sufficient in many contexts.

1.6 Organization

The remainder of this dissertation is organized as follows. Chapter 2 offers a detailed examination of the relevant literature, including that related to TOD modeling, reliability measures, and Bayesian statistics. Chapter 3 develops the three models investigated in this dissertation: the continuous logit, the continuous cross-nested logit (CCNL), and the bivariate multinomial probit (BVMNP). This chapter also details the Bayesian estimation procedures adopted for each model. In Chapter 4, the methods and estimation results for imputing time-varying LOS variables used in the TOD choice models are presented. Chapter 5 details the empirical results of the first two continuous choice models. In Chapter 6, empirical results of the BVMNP model for tour scheduling are discussed. And finally, Chapter 7 offers some concluding remarks and opportunities for future research.

1.7 Chapter Summary

As the questions posed to our travel demand models have changed over the past several decades, many model enhancements have emerged, such as activity- and tour-based methods and dynamic traffic assignment techniques. However, time-of-day modeling components of these models remain a major weakness. This dissertation contributes to the growing body of research on time-of-day modeling, by developing and applying three new models.

CHAPTER 2: LITERATURE REVIEW

As discussed in Chapter 1, there are several elements to the research presented in this dissertation. This chapter explores the related literature, emphasizing key methodological techniques, including models of travel timing. Over the past two decades, much research, and a variety of models have been devoted to uncovering the key determinants of travelers' trip timing behavior. Understanding the effect of travel time reliability represents a core element of such work. This chapter compares reliability measures identified in the literature, and concludes with a discussion of Bayesian statistical techniques, which are used for this dissertation's empirical analysis.

2.1 Travel Timing Models

Travel timing models developed in the literature can generally be broken into two broad categories: discrete choice methods and continuous techniques. Discrete choice methods require that the time dimension be discretized to provide a set of discrete alternatives. For instance, the 24-hour day may be discretized into 24 1-hour interval alternatives, or the 2-hour morning peak period may be broken into twelve 10-minute intervals. Such methods usually take a random utility maximization (RUM) model form in which travelers are assumed to choose the alternative that offers the greatest benefit (or utility). Continuous models generally do not take the form of a RUM model, but do not require discretization of the time dimension. The following sections discuss these two approaches, highlighting the advantages and weaknesses of each.

2.1.1 Discrete Choice Models

Discrete choice models offer a convenient form for time-of-day (TOD) modeling, since destination, mode, and other travel dimensions typically take discrete choice forms and such models are based on the behavioral premise that individuals seek to maximize their own utility when choosing alternatives. Moreover, if models of the generalized extreme value (GEV) family (such as the multinomial and nested logits) can be developed for

these purposes, relatively straightforward calculations for traveler welfare can emerge (see, e.g., de Jong et al. 2007, Zhao et al. 2008, and Kockelman and Lemp 2009). Traveler welfare is important for a variety of reasons, including environmental justice considerations, cost-benefit analyses, and project ranking and evaluation. Of course, such models require that the time dimension be broken into discrete intervals, for which interval boundaries are often set rather arbitrarily. Nonetheless, the convenient structural form of discrete choice models (GEV models in particular) offers many advantages in both model estimation and application, and they have been used a great deal in the literature to identify the determinants of departure time choice and activity durations, as described below.

Some of the earliest models of travel timing examined departure time choices, often focusing on commuter trips. The multinomial logit (MNL) model (McFadden 1973 and 1978), in particular, was widely used in the early years of trip timing research. Abkowitz (1981) and Small (1982) modeled departure times in this way, both using twelve 5-minute departure time intervals (spanning one hour during the AM peak period). These studies illuminated key demographic (e.g., age, income, and occupation) and network level-of-service (LOS) effects on departure time choice. Using broader TOD periods (e.g., peak, before peak, and after peak), McCafferty and Hall (1982) also modeled commuter departure time choice in a MNL framework. Similar TOD alternatives were considered by Saleh and Farrell (2005), finding that individuals with flexible work schedules are much more likely to shift the timing of their travel, which has important implications for many transportation policies, including congestion pricing.

All of the models discussed above use a MNL structure, which has well known limitations, including the assumption of independence across irrelevant alternatives (IIA). Such limitations have led to the use of more flexible model structures for departure time choice. For instance, Chin (1990) compared the MNL with the nested logit model (which relaxes the MNL's independence of irrelevant alternatives [IIA] assumption). The nested

logit model contained three nests for very early morning, early morning, and morning slots, thus allowing for correlations across alternatives in common nests. However, their estimation results yielded inclusive value parameter estimates that were inconsistent with random utility theory. This is probably due to the fact that the nested logit does not adequately describe the types of correlations one would expect in the context of departure time choice. For example, an alternative at the very end of the early morning period and one at the very beginning of the morning period would likely be correlated in some way, since they are adjacent over the time spectrum. However, the nested logit does not allow such correlations, and it presumes that the first alternative in the morning nest is as correlated with the last alternative in the morning nest as it is with the second alternative in the nest. Again, this simply does not seem an adequate representation of the correlation structure one would expect in a departure time choice context. Small (1987) introduced a more flexible model specification, the ordered generalized extreme value (OGEV) model. The OGEV model can be viewed as a generalization of the nested logit for ordered alternatives, allowing correlations to emerge between alternatives in order to reflect the ordered nature of those alternatives. However, Small's (1987) empirical investigation with twelve 5-minute time slots suggested the model was no better than the conventional MNL specification.

While the above models all focused on work-related travel, several research efforts have examined non-work travel. For instance, Okola (2003) investigated the determinants of TOD choice for elderly persons' recreational activities using six broad times of day. Estimation results suggested a propensity for choice of the morning peak and midday time periods. Using similarly broad periods, Steed and Bhat (2000) investigated MNL and OGEV model structures to reveal determinants of social, recreational, and shopping trip timing. It is questionable whether a great deal of correlation would exist between broad TOD intervals (such as the ones used here), and results of the study confirm these suspicions. The OGEV model did not perform any better than the MNL.

Other research has studied departure time choice in a joint setting with mode choice. In a joint choice model of mode and departure time, with four mode and seven 10-min departure time alternatives, Hendrickson and Plank (1984) found that commuters are less likely to shift modes than departure times. Contrary to these findings, Tringides et al. (2004) used bivariate probit models of mode and departure time, finding that workers are more constrained in their trip timing choice than mode, while non-workers are more likely to shift times than modes. However, with the broad time periods used here (peak vs. off-peak), such results are not so surprising. In another joint choice application, Bhat (1998a) developed a MNL-OGEV model, essentially nesting the ordered departure time choice dimension within an upper-level mode choice dimension. Empirical results indicated that the model outperforms both the nested logit and MNL formulations. Bhat (1998b) compared mixed multinomial logit (MMNL) models with the MNL model in a joint choice setting (of mode and departure time). Results indicated that the mixed models offer smaller value-of-travel-time estimates than the MNL, and if random utility components are introduced along both choice dimensions (i.e., mode and departure time), one obtains the best model fit. However, the analytical intractability of the MMNL model creates challenges for estimation and application.

As tour- and activity-based travel demand models have become more common in practice, models of tour timing (as opposed to trip timing) have taken center stage. Of course, tours have two timing dimensions (departure time and duration).⁶ The problem becomes more complex with a second timing dimension, especially since such models must account for individuals' scheduling constraints when more than one tour is undertaken during the day (e.g., tours cannot overlap in time). Vovsha and Bradley (2004) modeled tour timing in this way using a joint MNL framework. In this case, the time dimension was split into 1-hour intervals (over the entire day). The authors used a

⁶ Note that tours may potentially have many timing dimensions, one for each trip made. Usually this is resolved by selecting a "primary" activity from the collection of tour activities, with an activity having a start and end time. The timing of the primary activity is then modeled rather than the timing of the entire collection of activities. The other activities' timings may then be handled via other methods.

rather innovative approach in specifying the utility in order to recognize the ordered nature of the alternatives. In particular, four utility components were present: one each related to departure, return, and duration time, and one related to the maximum expected mode utility⁷. The first three of these components were computed using shift variables, where demographic characteristics of the individual were interacted with departure time, return time, and duration (each treated as continuous variables). Thus, the utility function is allowed to vary continuously over time, even though tour timing is not modeled in a continuous way. Abou Zeid et al. (2006) and Popuri et al. (2008) used similar methods of analysis with 30-minute intervals. In addition, both studies developed a unique utility function form, where demographic and other time-invariant covariates were interacted with cyclical functions, allowing for highly flexible utility shapes to emerge as a function of time. Practically, however, the model structures are identical to that of Vovsha and Bradley (2004). While this method offers advantages in terms of computational tractability and incorporation of multiple timing dimensions, it does not recognize likely correlations between adjacent (and possibly non-adjacent) time intervals and durations, as other methods do. The method is, however, useful for integrating these time choice dimensions rather seamlessly for activity-based model systems.

Using stated preference data, de Jong et al. (2003) estimated MMNL models for joint choice of mode and departure and return times. Because only four alternatives were included in the stated preference design, utility functions were constructed carefully. The findings suggested that arriving early or late by one minute was valued higher by commuters than one minute of travel time. In addition, simulation results implied that a travel time or cost increase during peak periods would indeed shift travelers to off-peak periods, a factor important in policy design and decision-making. However, the study relied on four choice alternatives presented in the stated preference experiment, where tradeoffs could be evaluated in terms of a specific tour that respondents actually made.

⁷ This was computed using logsums over mode choice alternatives specific to departure and arrival times in the activity scheduling model. It essentially serves as a method for nesting mode choice under TOD choice.

Generalizing these results for broader tour scheduling (e.g., choice sets over an entire day) would be challenging. And, while not entirely prohibitive, the computational intractability of the MMNL model is undesirable, though the MMNL does allow for a highly flexible error structure, and can be used to approximate any GEV-type model (McFadden and Train 2000).

2.1.2 Continuous Models

The second broad type of travel timing models treat time as a continuous response variable. While such models have not taken the form of a RUM model, they do not require special treatment of the time dimension (i.e., time is not discretized as it is with discrete choice methods). These models (usually) utilize what is called a hazard function. The hazard function defines the probability that an agent leaves its current state in a time interval with boundaries t and $t + h$, given that the agent has remained in its current state until time t (Wooldridge 2002). More specifically, the hazard function is the limit of this probability as h goes to zero. The hazard function can take a number of different forms, including non- and semi-parametric forms. This section highlights some key methodologies and results found in the literature.

The simplest form of the hazard model is the Cox proportional hazards model (Cox 1972). Such models assume that covariates affect the hazard function in a proportional way, and have been used by a number of researchers to investigate activity durations. For instance, Niemeier and Morita (1996) observed differences in the duration of trip-making activities for men and women, Yee and Niemeier (2000) examined how activity durations changed for households over a period of years using panel data, Srinivasan and Guo (2003) studied activity durations for shopping trips, and Zhong and Hunt (2005) explored weekend activity durations. While proportional hazard models are convenient, the restriction that covariates affect the hazard in a proportional manner is limiting. To allow for a more flexible structure, Ettema et al. (1995) suggested accelerated hazards, which allow covariates to have time-varying effects on the hazard function. In other

words, the accelerated hazard specification captures an individual's propensity to remain in his/her current state for more or less time depending on the time already allocated to the individual's current state. Using a similar approach, Popkowski Leszczyc and Timmermans (2002) found that the choice and timing of activities depends in some ways on the type and durations of previously conducted activities, which is not surprising.

Another neglected attribute of these models is that of unobserved heterogeneity across durations. Heckman and Singer (1984) established that failure to control for such unobserved heterogeneity may produce large biases in model estimation results. In the context of activity durations, Bhat (1996) was among the first to incorporate unobserved heterogeneity directly into model specification. In that study, non parametric proportional hazards were used (with heterogeneity controlled non-parametrically) in an examination of shopping activity durations. Results indicated significant effects of unobserved heterogeneity and that using parametric forms for the baseline hazard and unobserved heterogeneity can lead to estimation bias. Lee and Timmermans (2007) used latent class specifications with an accelerated hazard model to capture heterogeneity in activity durations for five activity types, finding a strong relationship between heterogeneity and demographics.

While the models discussed above emphasize activity durations, there is also a large amount of literature related to departure time modeling in a continuous time setting. For instance, Wang (1996) used a parametric proportional hazard rate specification to model activity start times. Somewhat differently, Bhat and Steed (2002) developed a non-parametric hazard model using Gamma mixing distributions to account for unobserved heterogeneity in shopping trip departure times. Travel time and travel cost were introduced in the model, though results suggested these had little effect on departure times, likely due to travel times and costs varying only between peak and off-peak periods (rather than continuously). Komma and Srinivasan (2008) used a similar model specification for commute trip departure time, but included continuously varying travel

times. Their findings suggested that travel time had an appropriate effect on departure time choice. Thus, it seems likely that travel time plays a role in TOD choice; one simply needs to appropriately characterize it.

In contrast to these earlier applications, Gadda et al. (2009) used Bayesian estimation techniques to model departure times for four trip types. This represents one of the few attempts to utilize Bayesian techniques in models of TOD choice. Gadda et al. (2009) used accelerated hazard functions with and without unobserved heterogeneity to model departure times of home-based work and non-home-based trips. However, the absence of time-varying travel time and cost data limited the models' abilities to predict how transportation LOS changes affect departure time choice.

2.1.3 Time-of-Day Modeling Summary

The continuous and discrete approaches for TOD modeling offer respective advantages and weaknesses. For instance, the main advantage of continuous time models is that they avoid arbitrary time interval selection required by discrete choice methods. Time interval selection creates a number of issues that are described in greater detail below. However, network LOS data are often unavailable over continuous time, and typically only estimated for broad TOD periods (e.g., peak and off-peak periods). If such variables are of interest for the continuous model, special techniques must be used to impute these continuously over time. In addition, continuous time models are not based in random utility theory. Random utility theory is an integral part of many travel demand model systems and offers a meaningful technique for estimating user benefits. Moreover, integrating such continuous time models (like those described here) with other travel choice dimensions in an econometrically meaningful way presents many challenges. Thus, hazard-based models of trip/tour timing have gone largely unused in large-scale transportation model systems.

On the other hand, discrete choice models (usually) are based in random utility theory, and offer all the benefits derived therein. However, they require discretization of the time dimension, which produces a number of issues as discussed by Bhat and Steed (2002). First, interval boundaries are often set in an arbitrary manner, and if boundaries are changed, different model results emerge. Second, two points very close in time may be classified in two different intervals with discrete models, when, in reality, these points may be perceived as very similar options. Allowing for correlations in discrete models can help alleviate this issue to some extent, but continuous treatment of departure time can alleviate this issue fully. Third, discrete treatment of time results in loss of temporal resolution in the model, and as dynamic traffic assignment (DTA) methods (which rely on input demand from origin-destination pairs across very small time increments) become more popular, this issue will become more critical. And last, discrete intervals set in model calibration must also be used in model application, which can limit the policies that can be evaluated (for instance, variable congestion charging and other traffic control measures).

In addition to model selection, this dissertation seeks a better understanding of the relationship between travel time reliability and trip timing. In the following section, reliability measures are explored.

2.2 Reliability Measures

One key element of the proposed work is the incorporation of travel time (un)reliability measures. This is particularly challenging since there is no consensus regarding how reliability should be measured. Two main approaches currently exist: travel time distributions and schedule delay methods.

2.2.1 Travel Time Distribution

Travel time distribution approaches assume that travel times have some underlying distribution (on a single route at the same time-of-day for the same origin-destination pair

across days of the week/month), which can be measured. Based on that distribution, it is possible to obtain standard deviations, variances, and 80th, 90th, or 95th percentile travel times, each of which can be used to measure travel time variability. In the case of percentiles, usually the difference is taken between some upper percentile travel time and the mean or median travel time as the measure. Many studies have used these types of measures. For instance, Small et al. (1999) used the standard deviation of travel times in a stated preference survey, and estimated binary route choice models. Their findings suggested that a minute of added standard deviation in travel times is valued more than an added minute of average travel time. However, critics argue that standard deviation is not such a useful measure because it does not relate well to everyday commuting experiences and it is not well understood by non-technical audiences (TTI and CS 2006).

Hensher (2001), Small et al. (2005), Brownstone and Small (2005), and Bhat and Sardesai (2006) all measured travel time reliability as the difference between some upper travel time percentile and median or mean travel time. While application settings and results vary over these studies, all found values of travel time reliability to be sizable in their choice contexts. The measurement of reliability using percentiles has several advantages over standard deviations. Not only does it relate explicitly to a travel time threshold for which there is some specific probability of higher travel times, but it is much easier for non-technical audiences to understand (TTI and CS 2006). In addition, van Lint et al. (2008) have shown that the distribution of travel times may in fact exhibit some skew, but standard deviations simply cannot capture that dimension.

Distributional measures of reliability, however, may not be the most appealing for models of travel timing. These measures will typically enter the utility function just as travel cost or travel time might, which presumes that the possibility of additional delay is what causes the disutility. In reality, the disutility associated with unreliable travel may be in arriving late (or early) to an activity location. In other words, some of the planned activity participation is lost, and distributional measures of reliability can only proxy for

the disutility of this lost activity participation. Schedule delay approaches take a step toward capturing this idea of lost activity participation.

2.2.2 Schedule Delay Methods

First introduced by Small (1982), schedule delay approaches capture the effect of travel time variability by entering a penalty term in the utility function for arriving earlier or later than the *preferred* arrival time. Only one of these costs will come into play at a time, since an individual cannot arrive early and late simultaneously. The penalties for late and early arrivals can take any number of analytical forms. A number of studies have employed these measures in a variety of contexts for both stated and revealed preference (see, e.g., Noland and Small 1995, Small et al. 1999, Lam and Small 2001, Bates et al. 2001, de Jong et al. 2003, and Tseng and Verhoef 2008).

The main advantage of this approach is that it provides valuable behavioral insight, in the sense that disutility associated with unreliable travel times is quantified by penalties for arrival at non-preferred times. In comparison to distributional methods, it more appropriately captures the real source of disutility. Of course, the schedule delay approach has its limitations as well. It requires knowledge of individuals' preferred arrival times. For model application, preferred arrival times could be simulated or otherwise specified; but, for model calibration, stated preference data are probably needed in most cases. In this regard, distributional measures may have an advantage since they are based on data supplied by the transportation network, though obtaining these distributional measures (even with network data in hand) is no simple task.

The previous two sections of this chapter have detailed literature specifically related to the TOD modeling context of this dissertation. In the next section, methods of Bayesian inference (which are used in the empirical work of this dissertation) are introduced.

2.3 Bayesian Statistics

As compared to many applications of classical (or frequentist) statistics to complex model specifications, parameter estimation and interpretation can be much more straightforward with Bayesian inference (Wagenmakers et al. 2008). While classical methods rely on asymptotics to suggest that parameter estimates converge properly as sample size becomes very large, Bayesian methods yield estimates of parameter distributions for any sample size. These distributions can suggest intervals on which parameters have a high probability of being bounded. Classical methods offer point estimates and confidence intervals, which must be strictly interpreted as long run frequencies (Gelman et al. 2004).

In practice, Bayesian statistics are generally more flexible since they rely on model estimation via Markov chain Monte Carlo simulation with conditional distributions (Wagenmakers et al. 2008). Such conditional distributions allow models to be decomposed into smaller sub-problems, which are generally easier to handle. Classical methods, however, must deal directly with the entire problem.

Finally, the specification of prior distributions offers analysts the opportunity to include any established intuition regarding parameters directly in the model. In fact, Bayesian methods can be viewed as a way of optimally combining such intuition or prior knowledge with observed data. Of course, in practice, it is often useful to assume as little as possible about the parameters (i.e., the idea of letting the data speak for itself). And, in this dissertation, this latter approach will typically be used. However, it is useful to note that such opportunities exist when using Bayesian statistics. In the following subsections, many fundamental techniques for applications of Bayesian methods are discussed.

2.3.1 Bayesian Theory

Bayesian inference is based on Bayes' theorem, which states that the joint distribution of two random quantities, A and B , is equal to the conditional distribution of A given B

multiplied by the marginal distribution of B . Suppose that our data is composed of Y and X , where Y represents a vector (or matrix in a multivariate setting) of dependent variables, and X represents a matrix of explanatory variables. Let θ be a vector of unknown parameters in the model. Bayes' theorem then states that the joint distribution of Y and θ given X can be written as follows:

$$p(Y, \theta|X) = p(\theta|Y, X)p(Y|X) = p(Y|\theta, X)p(\theta|X) \quad (2.1)$$

Here, $p(\theta|X) = p(\theta)$ (since θ is generally taken to be independent of X) is the density of the *prior distribution*, which represents the analyst's beliefs about the distribution of parameters before examining the data; $p(Y|\theta, X)$ is the *likelihood function*, representing the probability of obtaining the actual sample of dependent variables, Y , given X and choice of θ ; $p(\theta|Y, X)$ is the density of the parameters' *posterior distribution*, representing the analyst's updated beliefs about the distribution of θ , given the occurrence of Y and X ; and, $p(Y|X)$ is the marginal density of the data, which does not depend on θ , and, thus, is simply a constant. The posterior distribution can be rewritten as follows:

$$p(\theta|Y, X) \propto p(Y|\theta, X)p(\theta) \quad (2.2)$$

Thus, Bayesian statistics can be viewed as a mechanism for combining one's beliefs and observed data. The following subsections detail some important techniques in applications of Bayesian statistics that will be used in this dissertations empirical work.

2.3.2 Prior Choice

One key piece of Bayesian analysis is the prior distribution. In theory, the prior can be chosen as anything one likes. And, as sample size grows, the effect of the prior on the posterior becomes smaller, regardless of the choice of prior (Gelman et al. 2004). However, there are some practical considerations here. If the prior density is zero for

some values of θ , the posterior density for such values must also be zero, regardless of sample size and likelihood function.

When an analyst has little or no intuition regarding the parameters before examining the data, it is common to employ non-informative priors. Generally, such priors come from improper distributions. For instance, a uniform prior on $[-\infty, \infty]$ or a normal prior with $\sigma^2 = \infty$ represent two non-informative prior choices. In neither case does the density function integrate to 1, but in both cases one can treat the density values as constant. Thus, the posterior distribution becomes proportional to the likelihood alone. However, this can sometimes lead to improper posterior distributions (i.e., functions that are not actually density functions), which is problematic (Gelman et al. 2004). Alternatively, one could represent one's lack of prior knowledge with vague proper priors (e.g., a normal distribution with very large variance). In this dissertation, vague proper priors are used regularly.

When the prior and posterior distributions come from the same family of distributions, the prior is said to be conjugate to the posterior. For example, if the prior was chosen to be distributed normal, and the posterior was also normal (with different parameters values, of course), the prior is conjugate. Fully conjugate priors are often not achievable with more complex models, but conditional conjugacy is often very helpful. This refers to a prior that is conjugate to the conditional posterior distribution of some parameter subset, given all other parameters. As will become evident in the next subsection, such conditional posteriors are often very important in Bayesian statistics applications.

2.3.3 MCMC Simulation

Modern Bayesian inference relies heavily on Markov chain Monte Carlo (MCMC) simulation, including Gibbs sampling and the Metropolis-Hastings (MH) algorithm (see, e.g., Gelfand and Smith 1990, Smith and Roberts 1993, Gelman et al. 2004, and

Gamerman and Lopes 2006). The objective is to generate random draws from a Markov chain whose stationary distribution is the posterior.

Suppose one is interested in a random process that evolves over iterations. Further, say that the state of the process is θ^n at iteration n . Here, it will be assumed that the state space is discrete, since the results are the same for continuous distributions. A Markov chain is a sequence of random variables (e.g., $\theta^0, \theta^1, \dots$) where the distribution of θ^n given all previous values of θ ($\theta^0, \theta^1, \dots, \theta^{n-1}$) only depends on θ^{n-1} . That is, given the state of the process at iteration $n-1$, the distribution of θ^n is independent of all other previous states. Suppose the state space is discrete with states given by $\{1, 2, \dots, Q\}$, and the transition probabilities are written as $P(\theta^n = b | \theta^{n-1} = a) = P_{ab} = P_{a \rightarrow b}$. Thus, P_{ab} is the probability of transitioning from state a to state b , and $\sum_b P_{ab} = 1 \forall a$. The transition matrix is given by the following:

$$P = \begin{bmatrix} P_{11} & \cdots & P_{1Q} \\ \vdots & \ddots & \vdots \\ P_{Q1} & \cdots & P_{QQ} \end{bmatrix} \quad (2.3)$$

Now suppose that $\theta^0 \sim \pi_0$, where π_0 is a vector of probabilities (that sum to 1). By definition, $\theta^n \sim \pi_0 P^n$. If it is assumed that $P_{ab} > 0 \forall a, b$, and given that there are a finite number of states, then there is a unique stationary distribution, π , such that $\pi P = \pi$, and $E(f(\theta)) = \lim_{n \rightarrow \infty} (1/n) \sum_{i=1}^n f(\theta^i)$. In fact, these conditions are more than sufficient to show these results. One need only show the chain to be irreducible, positive recurrent, and aperiodic, in which case the chain is said to be ergodic (Gamerman and Lopes 2006).

Irreducibility and positive recurrence are somewhat related. A Markov chain is said to be irreducible if every state b , can be transitioned to from any other state a , for all state pairs a and b , though the transition need not be direct (Gamerman and Lopes 2006). A Markov

chain is positive recurrent if all its states are positive recurrent. The probability of returning to a recurrent state, starting at any other state is 1; and for a positive recurrent state, the expected number of iterations required before returning to it is finite (Gamerman and Lopes 2006). Finally, while a periodic chain may have a unique stationary distribution, aperiodicity establishes the convergence of transition probabilities for a large number of iterations (Gamerman and Lopes 2006).

Consider again a Markov chain transition matrix P (as in equation 2.3), with some fixed number of states Q , and that each element is strictly positive. Suppose, however, that instead of drawing θ^n conditional on all previous draws, θ^n is drawn conditional on all future draws. Using Bayes' theorem, the probability of θ^n taking state b is defined as follows:

$$\begin{aligned}
P(\theta^n = b | \theta^{n+1} = a_1, \theta^{n+2} = a_2, \dots, \theta^{n+k} = a_k) &= \frac{P(\theta^n = b, \theta^{n+1} = a_1, \dots, \theta^{n+k} = a_k)}{P(\theta^{n+1} = a_1, \dots, \theta^{n+k} = a_k)} \quad (2.4) \\
&= \frac{P(\theta^n = b)P(\theta^{n+1} = a_1 | \theta^n = b)P(\theta^{n+2} = a_2, \dots, \theta^{n+k} = a_k | \theta^{n+1} = a_1)}{P(\theta^{n+1} = a_1)P(\theta^{n+2} = a_2, \dots, \theta^{n+k} = a_k | \theta^{n+1} = a_1)} \\
&= \frac{P(\theta^n = b)P_{ba_1}}{P(\theta^{n+1} = a_1)} \\
&\equiv \frac{\pi_b P_{ba_1}}{\pi_{a_1}} \quad (2.5)
\end{aligned}$$

Thus, the reverse process is a Markov chain, where $P'_{ab} = \frac{\pi_b P_{ba}}{\pi_a}$. Here, π is the stationary distribution of the forward Markov chain process, P is the transition matrix of the forward Markov chain process, and P' is the transition matrix of the reverse Markov chain process. Moreover, a Markov chain is defined to be time reversible if the following conditions holds (Gamerman and Lopes 2006):

$$\pi_a P_{ab} = \pi_b P_{ba} \quad (2.6)$$

In words, a Markov chain is time reversible if the probability of obtaining an a to b transition is the same as the probability of obtaining a b to a transition. This is not equivalent to the condition where the one-step probability of transitioning from a to b is the same as the probability of transitioning from b to a , which would imply that P_{ab} must equal P_{ba} . Time reversibility is a property needed to construct the MH algorithm, discussed later.

Although the discussion in this section focused on discrete random variables, the methods and results hold in a continuous setting as well. In the next subsection, MCMC results are extended to illustrate their applicability in Bayesian statistics and to Gibbs sampling, in particular.

2.3.3.1 Gibbs Sampling

The Gibbs sampler results from defining a Markov chain in the parameter space so that the posterior density is the stationary distribution of the Markov chain. Suppose the full conditional posterior density of θ_i is given by $p(\theta_i|\theta_{i-}, Y, X) \forall i$, where θ_{i-} denotes the full set of parameters with θ_i (which may be a vector or scalar) removed. Suppose also that initial values are given for each parameter so that $\theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_k^0)$. The Gibbs sampler is defined as a Markov chain where, at each iteration, each parameter is drawn from its full conditional posterior distribution.⁸

Step 1: Draw θ_1^n from $p(\theta_1|\theta_2^{n-1}, \theta_3^{n-1}, \dots, \theta_k^{n-1}, Y, X)$.

Step 2: Draw θ_2^n from $p(\theta_2|\theta_1^n, \theta_3^{n-1}, \dots, \theta_k^{n-1}, Y, X)$.

⋮

Step k: Draw θ_k^n from $p(\theta_k|\theta_1^n, \theta_2^n, \dots, \theta_{k-1}^n, Y, X)$.

⁸ Note that this does not mean that each parameter is from the correct joint posterior distribution, at least not for the first number of iterations. A “burn-in” period is required for parameters to converge to the correct joint posterior.

Since θ_i^n comes from the correct conditional posterior for all i , θ^n must be from the correct joint posterior (for large enough n). In fact, only mild regularity conditions are need to show that draws of θ converge to the correct joint posterior distribution (see, e.g., Roberts and Smith 1994 and Tierney 1994).

This is an important result because often (particularly for complex models) the full joint posterior distribution cannot be recognized as coming from a familiar family of distributions and cannot be drawn from directly. The Gibbs sampler allows one to draw from a series of (simpler) conditional distributions instead, and these usually are constructed to be “easy” to draw from. This is typically accomplished through specification of conditionally conjugate priors. Such priors are conjugate to the conditional posterior distribution (rather than the full joint posterior). Of course, this is not always the case. The following subsection details the Metropolis-Hastings algorithm for drawing from distributions with unfamiliar and/or difficult forms.

2.3.3.2 *Metropolis-Hastings*

The Gibbs sampler presented above allows an analyst to draw parameters using a series of conditional distributions. Usually it is convenient to build a model in which each set of parameters is conditionally conjugate, to facilitate drawing from the conditional posteriors. However, in many cases, it may not be possible to construct the model in this way, and certain groups of parameters may have non-standard distributions (that are typically more difficult to draw from directly). The Metropolis-Hastings (MH) algorithm provides a rather simple method for generating such draws. Metropolis et al. (1953) first proposed the algorithm, which was later extended by Hastings (1970).

As with the Gibbs sampler, the goal is to construct a Markov chain with stationary distribution equal to the posterior distribution. However, suppose that draws of θ^n cannot be easily obtained (e.g., $p(\theta^n|\theta^{n-1})$ is some unknown density), but draws from the density $q(\theta^n|\theta^{n-1})$ are easy to generate. Here, q could be any density one likes.

Clearly, since choice of q is somewhat arbitrary, the Markov chain defined simply with transition density q will generally not converge to the posterior distribution of interest. Gamerman and Lopes (2006) discuss how, if the Markov chain is constructed to be time reversible, with q acting as a “proposal density,” then the chain will converge to the correct posterior distribution. The MH algorithm works as follows:

Step 1: Draw $\theta^n = b$ from $q(\theta^n | \theta^{n-1} = a)$.

Step 2: Compute $s(a, b) = \min \left\{ 1, \frac{\pi_b q(a|b)}{\pi_a q(b|a)} \right\}$.

Step 3: With probability $s(a, b)$, $\theta^n = b$.

With probability $1 - s(a, b)$, $\theta^n = a$.

Thus, each proposed draw of θ can either be accepted or rejected, where the probability of acceptance is given by $s(a, b)$. Here, π_a and π_b represent the posterior densities (or conditional posterior densities, as the case may be) of a and b , respectively. One need only be able to compute the ratio of $\frac{\pi_b}{\pi_a}$, not necessarily the posterior density itself, which can be important in some applications. The transition density of this Markov chain is given by the following:

$$P_{ab} = q(b|a)s(a, b) \forall a \neq b \tag{2.7}$$

$$\begin{aligned} \pi_a P_{ab} &= \pi_a q(b|a)s(a, b) = \pi_a q(b|a) \min \left\{ 1, \frac{\pi_b q(a|b)}{\pi_a q(b|a)} \right\} \\ &= \min \{ \pi_a q(b|a), \pi_b q(a|b) \} = \pi_b P_{ba} \end{aligned}$$

Thus, the transition density is time reversible with stationary distribution π . Roberts and Smith (1994) show that under mild regularity conditions for s and q , the Markov chain’s stationary distribution is equivalent to the correct posterior distribution.

One simplifying technique for choosing transition densities is to construct a symmetric chain (Gamerman and Lopes 2006). A chain is symmetric if for every pair of states, the

density is symmetric in its arguments (i.e., $q(a|b) = q(b|a)$). Thus, the transition densities cancel and so do not matter in computing the acceptance probability, $s(a, b)$. The acceptance criterion only depends on whether the proposal is more or less likely under the posterior distribution than the current set of parameters.

Random-walk chains represent another reasonable choice here. The proposal for a random walk chain takes the following form:

$$\theta^n = \theta^{n-1} + \varepsilon^n \tag{2.8}$$

Here, θ^n is the proposed parameter draw at iteration n , θ^{n-1} is the current draw of θ , and ε^n is a random disturbance term with density f_ε , generally taken to be independent and identically distributed (IID) across iterations (Gamerman and Lopes 2006). Though f_ε need not be symmetric, it should allow for both negative and positive values. If f_ε is symmetric about 0, then the chain is symmetric. Symmetric random walk chains are most common for practical applications, usually assuming f_ε is a normal density (Gamerman and Lopes 2006). One reason for using a random walk MH algorithm (rather than an independence chain, where the proposal is independent of previous draws) is that it can be simpler to calibrate in practice. Most MH algorithms require some degree of calibration to work properly, since it is not always clear what types of parameter values are most likely under the posterior. In this dissertation, the normal random walk proposal is used in all cases, and proposal variances are tuned to facilitate speedy convergence.

2.3.4 Convergence Assessment

One key element of most Bayesian applications is assessing convergence of generated draws. While one would typically like to obtain independent draws, neither the Gibbs sampler nor the MH algorithm will generate independent draws⁹. One simple approach

⁹ By definition, the Markov chains are defined to dependent on previous draws. Thus, each draw will show some signs of dependence.

for reducing dependence is to drop draws (e.g., keep every 5th or 10th draw). However, even with independent draws, there is generally no simple or conclusive way to identify when a Markov chain has converged (Gamerman and Lopes 2006), though a couple of methods are described here.

Gelfand and Smith (1990) suggest graphical techniques as informal checks of convergence. For instance, if an analyst obtains 1,000 draws he/she thinks are from the posterior density, estimated parameter densities using the first 200 could be plotted against estimated parameter densities for the last 500. If these parameter densities are indistinguishable, convergence is accepted. Alternatively, one might plot parameter draws as a function of the draw number (see, e.g., Gamerman 1997). If clear trending is evident in any parameters, it can be used as a reasonable indication that convergence has not been reached (Gamerman and Lopes 2006).

Others have attempted to provide more formal convergence diagnostics. One rather simple method was proposed by Geweke (1992). Suppose the analyst has n draws for which convergence will be tested. If A represents the set of draws from 1 to n_A , and B represents the set of draws from $n-n_B+1$ to n , with no overlap between sets A and B , Geweke's (1992) diagnostic compares the ergodic averages of the two sets, which should be very similar if convergence has been reached. Formally, the diagnostic is computed as follows:

$$z_G = \frac{\bar{\theta}_A - \bar{\theta}_B}{[\widehat{Var}(\theta_A) + \widehat{Var}(\theta_B)]^{1/2}} \xrightarrow{d} N(0,1) \quad (2.9)$$

Here, $\bar{\theta}_A$ and $\bar{\theta}_B$ are the means over sets A and B , respectively, and $\widehat{Var}(\theta_A)$ and $\widehat{Var}(\theta_B)$ are estimated variances over the two sets. As n_A and n_B grow in size, the distribution of z_G approaches standard normal if the chain has converged. Geweke (1992) recommends $n_A = 0.1n$ and $n_B = 0.5n$.

Other diagnostics have also been proposed. Gelman and Rubin (1992) used multiple chains starting at different points to formalize convergence criteria, and Zellner and Min (1995) proposed a method based on conditional distributions. Of course many others exist as well (see, e.g., Gamerman and Lopes 2006). However, no method can definitively prove convergence, and different problems may have different convergence criteria acceptance considerations (Gamerman and Lopes 2006). In this dissertation, informal graphical techniques and Geweke's (1992) convergence diagnostic (e.g., equation 2.9) are used, but more as informal checks of convergence rather than definitive measures.

2.4 Chapter Summary

This chapter reviewed a variety of literature related to this dissertation work. Section 2.1 offered a detailed examination of the methods used in TOD modeling research, discussing important considerations needed for such work as well as highlighting some key strengths and weaknesses of existing continuous- and discrete-response methods. Section 2.2 investigated different measures used to quantify travel time (un)reliability. Finally, Section 2.3 presented the fundamentals of Bayesian inference, with particular attention given to those methods used in this dissertation. In the next chapter, the statistical methods used in this dissertation's empirical work are presented.

CHAPTER 3: TIME-OF-DAY MODELING METHODS

In this chapter, one existing and two new time-of-day (TOD) model specifications are developed in detail. All of the methods presented here derive from principles of random utility, where each individual is assumed to make decisions in order to maximize his/her utility. A key reason for developing such models is that random utility theory offers a strong basis for economic welfare calculations, which can be important for project evaluation and cost-benefit analyses. In addition, all models consider the choice context of an entire 24-hour period. Three other desirable features also were sought. First, the models should allow for continuous choice (since the timing of travel is inherently continuous in nature). Second, the models should allow correlations across similar choice alternatives (i.e., those close on the temporal continuum). Finally, the models should allow for both outbound and inbound TOD choice of an activity or tour. While none of the models developed here is able to offer all three components, all are based on random utility maximization (RUM) and represent advances in behavioral modeling.

In Section 3.1, the continuous logit model is developed for departure time choice. The continuous logit offers a continuous choice setting, though it does not allow for correlations or two dimensions of travel timing¹⁰. While the continuous logit model is nothing new, the literature offers no applications of the model in a departure time setting, and its derivation here is important for derivation of the continuous cross-nested logit (CCNL) model. Section 3.2 presents the CCNL model formulation, also for departure time choice only. The CCNL offers a continuous choice setting and allows correlations to emerge across timing choices, but like the continuous logit, the model is developed for a single timing dimension¹¹. In Section 3.3, a bivariate multinomial probit (BVMNP)

¹⁰ While the continuous logit is presented here for a single timing dimension, adding a second timing dimension would not be too difficult (as done by Ben-Akiva and Watanatada [1981] and Ben-Akiva et al. [1985] in two-dimensional spatial applications) and would be computationally feasible to estimate.

¹¹ Like the continuous logit, adding a second timing dimension to the CCNL is theoretically possible, but estimation is computationally prohibitive (at this age of computing).

model is developed. Unlike the logit models, the BVMNP offers two-dimensional tour timing choice. In addition, the model allows correlations across alternatives to emerge. However, the model is not continuous in its choice contexts.

3.1 Continuous Logit

McFadden (1978) developed the GEV class of models for discrete choice applications to make use of random utility theory. Every GEV model is derived from a function $G(y_1, \dots, y_J)$, where $j = 1, \dots, J$ indexes the set of alternatives (with G satisfying some regularity conditions [see, e.g., McFadden 1978 and Small 1987]). If random utility for any alternative j is defined as a systematic component plus a random error component (where the joint density of all error components is distributed according to the extreme value distribution), as shown in equation 3.1, then the probability that alternative k is chosen (i.e., alternative k offers the maximum utility) is given by equation 3.2.

$$U_j = V_j + \varepsilon_j \quad (3.1)$$

$$P_k = \frac{y_k G_k(y_1, \dots, y_J)}{G(y_1, \dots, y_J)} \quad (3.2)$$

Here, $G_k(y_1, \dots, y_J)$ denotes the partial derivative of G with respect to y_k , and $y_j = e^{V_j}$, for all j . For the MNL model, G is given by equation 3.3 and the choice probability of alternative k takes equation 3.4's familiar form.

$$G(y_1, \dots, y_J) = \sum_{j=1}^J y_j \quad (3.3)$$

$$P_k = \frac{\exp(V_k)}{\sum_{j=1}^J \exp(V_j)} \quad (3.4)$$

3.1.1 Continuous Logit Specification

The continuous logit model represents a generalization of the MNL model for a continuous response variable (McFadden 1976, Ben-Akiva and Watanatada 1981, and

Ben-Akiva et al. 1985). The model can be derived directly from the random utility assumption. For instance, consider again the utility equation 3.1, but assume that utility varies continuously over departure time choice, t . The utility expression can thus be written as follows (for observation i):

$$U_i(t) = V_i(t) + \varepsilon_i(t) \quad (3.5)$$

Suppose the continuous response variable of interest, t , is bounded by b_1 and b_2 , discretize t such that t_j denotes the j^{th} discrete alternative (where $j = 1, \dots, J$), and let $t_1 = b_1$ and $t_J = b_2$. Also, suppose that ε_j (the error component for the j^{th} discrete alternative) is independent and identically distributed over all J alternatives. Now, suppose J (the number of discrete alternatives) is computed as $J = 1 + \frac{b_2 - b_1}{s}$, where s denotes the distance between each discrete alternative. Since t has been discretized, the model can now be written as a MNL, with generating function given by equation 3.3 and choice probabilities given by equation 3.4. As s decreases in size, the number of discrete alternatives grows, but the generating function and choice probabilities remain of the same form. However, in the limit as $s \rightarrow 0$, one obtains the continuous logit generating function and choice density function shown in equations 3.6 and 3.7, respectively.

$$G(y_1, \dots, y_J) = \int_{b_1}^{b_2} \exp(V(t)) dt \quad (3.6)$$

$$p_{t_k} = \frac{\exp(V(t_k))}{\int_{b_1}^{b_2} \exp(V(t)) dt} \quad (3.7)$$

Here, b_1 and b_2 define the bounds of the choice space. The choice density shown in McFadden (1976), Ben-Akiva and Watanatada (1981), and Ben-Akiva et al. (1985) appears slightly different than that shown in equation 3.7. Their derivations include an additional additive component in the utility equation, called the opportunity density (formulated as the natural logarithm of an attractiveness function), which describes the

density of choice alternatives at a particular point in space. The authors use this function in the context of location choice, since some locations have more opportunities than others (e.g., a body of water will have no household locating opportunities while a densely populated urban area may have many). Essentially this is done to reflect another dimension and ensure no aggregation biases. In the context of departure time choice, one may view this opportunity density function as a zero-one indicator. If an individual has already scheduled an activity during some particular time interval, the opportunity density will be zero over that interval, and the opportunity density will be one elsewhere. Here, the opportunity density will be ignored (i.e., it is assumed equal to one everywhere).

As developed above, the continuous logit model represents a generalization of the MNL model. Such a model of tour timing was described by Ettema and Timmermans (2003), except the denominator of the density function was discretized, and the model was estimated as a MNL. Moreover, since time-varying travel times and costs were not available to Ettema and Timmermans, such time-dependent covariates were not included in their tour scheduling model. Time-dependent covariates are developed for the models in this dissertation, as described in Chapter 4.

Since utility is arguably (smoothly) continuous in departure time choice, it deserves special attention here. Similar to the typical MNL model, covariates that do not vary over time alternatives (e.g., an individual's gender or age) cannot be introduced in the normal way. One could imagine any number of continuous utility forms to use in this context. For instance, Abou Zeid et al. (2006) and Popuri et al. (2008) specified systematic utility in a continuous framework¹², both using cyclical functions of departure time choice interacted with covariates. A similar systematic utility specification is adopted here, as shown in equation 3.8.

¹² Note that the models in these applications were discrete choice models, not continuous ones.

$$V_i(t) = X_i\beta\gamma(t) + \sum_{p=1}^{\mathcal{P}}\eta_p g_{ip}(t) \quad (3.8)$$

$$\gamma(t) = \left[\sin\left(\frac{2\pi t}{24}\right), \sin\left(\frac{4\pi t}{24}\right), \dots, \sin\left(\frac{2Q\pi t}{24}\right), \cos\left(\frac{2\pi t}{24}\right), \cos\left(\frac{4\pi t}{24}\right), \dots, \cos\left(\frac{2Q\pi t}{24}\right) \right]' \quad (3.9)$$

Here, X is a row vector of K individual specific variables; β is a matrix of parameters to be estimated with size $K \times 2Q$; and $\gamma(t)$ is a $2Q \times 1$ column vector consisting of cyclical functions of the departure time t . Note that some covariates may be interacted with fewer than $2Q$ cyclical functions by restricting the applicable elements of β to be zero. Time-varying variables (such as travel time, travel time variability, and cost) are represented by $g_{ip}(t)$, and there are \mathcal{P} of these variables, with parameters given by η_p . Since these variables vary over time, they need not enter the systematic utility in any special form. There are a couple of reasons for selecting this utility form. First, it allows utility to take on a rich assortment of shapes, including multimodal ones. In addition (and as pointed out by Abou Zeid et al. [2006] and Popuri et al. [2008]), 24 hours is a multiple of the period of each cyclical function, which offers day-to-day consistency in the utility function (e.g., utility at 0 and 24 hours is identical).

3.1.2 Continuous Logit Parameter Estimation via MCMC Simulation

In order to estimate the continuous logit's parameters, Bayesian methods are employed. In this work, the β prior is chosen to be independent (i.e., each parameter in β is independent of each other parameter in β under the prior) and normally distributed with vague prior information (i.e., large variance). Since one expects network variables to affect the utility function in a negative way, the η prior is chosen to be independent and distributed according to a log-normal multiplied by minus one, again with vague prior information. Given the large dataset ($n = 997$), the priors should play little role in the posterior distribution.

Draws from the posterior distribution of parameters are obtained by constructing a Markov Chain with stationary distribution equal to the posterior distribution, as described

in Chapter 2. Here, a one-step Metropolis-Hastings (MH) algorithm can be employed to draw all parameters simultaneously.¹³ A MH algorithm is needed here since the posterior distribution of parameters cannot be drawn from any standard distributions. The proposal density is chosen to be multivariate normal with mean given by the previous parameter draw (i.e., a normal random walk proposal). The η parameter deserves special attention here. Instead of drawing η directly in the proposal density, a transformation variable, $\eta' = \ln(-\eta)$, is drawn. By transforming η in this way, η' can take on negative or positive values and a normal proposal distribution is appropriate.

Initially, the covariance matrix is set to zero on all off-diagonal elements, and very small values on all diagonal elements. This ensures a high rate of acceptance during initial stages of the process. After 2,000 draws were obtained using the MH algorithm (and for every 20th draw thereafter), the proposal density's covariance matrix was updated (i.e., set to equal the estimated covariance matrix from previous draws, not including the current draw). Note that the covariance matrix used to generate proposals is actually deflated to increase the probability of proposal acceptance. Gelman et al. (2004) suggests a deflation factor of $5.8 / d$, where d is the number of parameters drawn. Here, the deflation factor was initially set to 0.2, and adjusted accordingly every 1,000 draws depending on the rate of proposal acceptance over the previous 1,000 draws. Finally, after 5,000 draws were obtained, and thereafter, the covariance matrix was estimated from the previous 5,000 draws (not including the current draw). Since the adaptive MH algorithm does not include information from the current value of the parameters (other than the mean of the proposal density), the algorithm will converge to the proper posterior distribution (Holden et al. 2009).

After it appeared the draws have converged to the posterior distribution, additional draws were obtained from the posterior, with only every 50th draw being retained for inference.

¹³ Gibbs sampling is not needed here, since all parameters can be drawn simultaneously from a single multivariate proposal distribution.

With an MH process such as this, it is inevitable for consecutive draws to be correlated. By dropping draws, this autocorrelation is reduced. Convergence was diagnosed using informal graphical techniques¹⁴ as well as Geweke's (1992) diagnostic. Results are presented in Chapter 5.

3.2 Continuous Cross-Nested Logit

The continuous cross-nested logit (CCNL) model represents a generalization of the cross-nested logit (CNL)¹⁵ for a continuous response variable; much like the continuous logit represents a generalization of the MNL. In formulating the CCNL here, it is convenient to first discuss the CNL for discrete choice.

3.2.1 (Discrete) Cross-Nested Logit

The CNL model for discrete alternatives has been well documented in the literature (see, e.g., Small 1987, Vovsha 1997, Ben-Akiva and Bierlaire 1999, Wen and Koppelman 2001, and Papola 2004, among others) and offers a rather flexible correlation structure for discrete alternatives. The CNL is formulated here to aid in the notation and formulation of the continuous cross-nested logit (CCNL) detailed in Section 3.2.2. The CNL's generating function and choice probabilities are shown in equations 3.10 and 3.11, respectively.

$$G(y_1, \dots, y_J) = \sum_{m=1}^M \left[\sum_{j \in \mathcal{D}_m} (\alpha_{jm} y_j)^{\rho_m} \right]^{\frac{1}{\rho_m}} \quad (3.10)$$

$$P_k = \frac{\sum_{m=1}^M \left((\alpha_{km} y_k)^{\rho_m} \left[\sum_{j \in \mathcal{D}_m} (\alpha_{jm} y_j)^{\rho_m} \right]^{\frac{1}{\rho_m} - 1} \right)}{\sum_{n=1}^M \left[\sum_{j \in \mathcal{D}_n} (\alpha_{jn} y_j)^{\rho_n} \right]^{\frac{1}{\rho_n}}} \quad (3.11)$$

¹⁴ In other words, the draws were plotted and investigated for patterns/trends. The absence of such trends is an indicator of convergence (see, e.g., Gamerman and Lopes 2006).

¹⁵ When not explicitly stated otherwise, the term cross-nested logit (or CNL) will refer to discrete choice contexts here.

Here, $m = 1, \dots, M$ indexes the set of nests, α_{jm} is an allocation parameter defining the degree to which alternative j is a member of nest m , ρ_m denotes the inclusive value parameter for nest m , and \mathcal{D}_m is the subset of alternatives in nest m . To be consistent with random utility theory, $\rho_m \geq 1, \forall m$. Furthermore, α_{jm} should satisfy the conditions $\alpha_{jm} \geq 0, \forall j, m$ and $\sum_{m=1}^M \alpha_{jm} = 1, \forall j$ (Wen and Koppelman 2001, Bierlaire 2006, Abbe et al. 2007, and Marzano and Papola 2008). As a practical matter, each subset of alternatives, \mathcal{D}_m , should be distinct (ie., $\mathcal{D}_m \neq \mathcal{D}_n, \forall m \neq n$) for model identification purposes.

3.2.2 Continuous Cross-Nested Logit

Similar to the way in which the MNL is generalized for continuous response, the CNL can be generalized, though some additional support is needed, as discussed here. Of great importance is nest composition. It makes good sense to think of these nests as small, contiguous intervals of the continuous spectrum of alternatives. And, since the response variable (departure time) is continuous, it seems reasonable to restrict attention to the case of ordered alternatives. Thus, each nest should be constructed so that it contains a set of sequential elemental alternatives.

The set of nests could be structured in a couple of different ways. For instance, one could construct a finite number of nests, similar to the discrete CNL model. However, a more general approach is to consider the set of nests in the same manner as the set of alternatives, effectively infinite. Such treatment requires parameterization of the inclusive value and allocation parameters, as discussed in more detail below.

As before, suppose the continuous response variable of interest, t , is bounded by b_1 and b_2 (e.g., 0 and 24 hours for trip departure time); t is discretized so that t_j denotes the j^{th} discrete alternative; and J is the total number of choice alternatives, computed as $J = 1 + \frac{b_2 - b_1}{s}$. In addition, suppose the number of nests equals the number of alternatives

(i.e., $J = M$), and the nest interval is given as $2h$ (i.e., each nest, m , is composed of elemental alternatives ranging from alternative $t_m - h$ to alternative $t_m + h$). Let $\alpha(t_j, t_m)$ denote the allocation parameter for alternative j in nest m and let $y(t_j) = e^{V(t_j)}$, the exponent of systematic utility for alternative j . Like the parallel between MNL and continuous logit, it is now possible to write the generating function and choice probabilities for this discretized model as they appear in equations 3.10 and 3.11. As before, taking the limit, $s \rightarrow 0$ (and $J = M \rightarrow \infty$), results in the generating function and choice density function for the CCNL model, as shown in equations 3.12 and 3.13, respectively.

$$G(y_1, \dots, y_J) = \int_a^b \left(\int_{q-h}^{q+h} [\alpha(r, q)y(r)]^\rho dr \right)^{\frac{1}{\rho}} dq \quad (3.12)$$

$$p_{t_k} = \frac{\int_{t_k-h}^{t_k+h} [\alpha(t_k, w)y(t_k)]^\rho \left(\int_{w-h}^{w+h} [\alpha(r, w)y(r)]^\rho dr \right)^{\frac{1}{\rho}-1} dw}{\int_{b_1}^{b_2} \left(\int_{q-h}^{q+h} [\alpha(r, q)y(r)]^\rho dr \right)^{\frac{1}{\rho}} dq} \quad (3.13)$$

Here, a single inclusive value parameter, ρ , is considered instead of allowing each nest to have a different parameter (which would require further parameterization). In the context of continuous response, this seems reasonable since one may expect similar amounts of correlation across every pair of alternatives separated by a common distance (e.g., alternatives 5 minutes apart in the AM share the same amount of latent information as alternatives 5 minutes apart in the PM). Of course, such correlations may differ at different points along the continuous spectrum and relaxation of this assumption may offer added model flexibility. As with the discrete CNL model, ρ is bounded below by 1, to be consistent with random utility theory. Unlike ρ , the allocation parameters, $\alpha(r, q)$, are parameterized here, though they could be taken as constant over the nest interval.

Suppose that for any nest m and alternative j , the allocation parameter, $\alpha(t_j, t_m)$, is parameterized such that, if $m = j$, then $\alpha(t_j, t_m)$ takes on its greatest value. And define h

so that, if $|t_j - t_m| \geq h$, then $\alpha(t_j, t_m) = 0$; otherwise, $\alpha(t_j, t_m) > 0$. In other words, the allocation parameter is zero if alternatives m and j are more than h units apart (e.g., one hour apart), and strictly positive otherwise. (Remember that the number of nests equals the number of alternatives.) Finally, $\alpha(t_j, t_m)$ must be normalized as shown in equation 3.14, just as it was for the CNL (for unbiased results [Abbe et al. 2007]).

$$\int_{b_1}^{b_2} \alpha(t_j, t_m) dt_m = 1, \forall j \quad (3.14)$$

With this constraint it can be shown that equations 3.12 and 3.13 reduce to equations 3.6 and 3.7 for $\rho = 1$ (i.e., the CCNL collapses to the continuous logit), similar to the manner in which the CNL collapses to the MNL for inclusive value parameters of 1. Even with these restrictions, the analyst retains a great deal of flexibility in the parameterization of the allocation parameters. Here, a simple triangular formulation is proposed, as shown in equation 3.15 (and illustrated graphically in Figure 3.1).

$$\alpha(t_j, t_m) = \begin{cases} \frac{h - |t_j - t_m|}{h^2} & \text{if } |t_j - t_m| \leq h \\ 0 & \text{otherwise} \end{cases} \quad (3.15)$$

Here, it is assumed that time rolls over at midnight, meaning that 0 hours and 24 hours are identical. This ensures that condition 3.14 holds for t_j “close to” (i.e., within h units of) either limit, b_1 or b_2 . In addition, it allows for correlations to emerge across these times, which seems reasonable given the cyclical nature of a day period. (Essentially, those reporting trips in the sample view an 11:59 pm departure time virtually the same as 12:01 am. Most travel surveys start and end trip reporting at 3 am anyhow.) By formulating α in this way, correlations across alternatives separated by a common distance are identical. Moreover, it allows a single parameter, h , to control behavior of the allocation parameters.

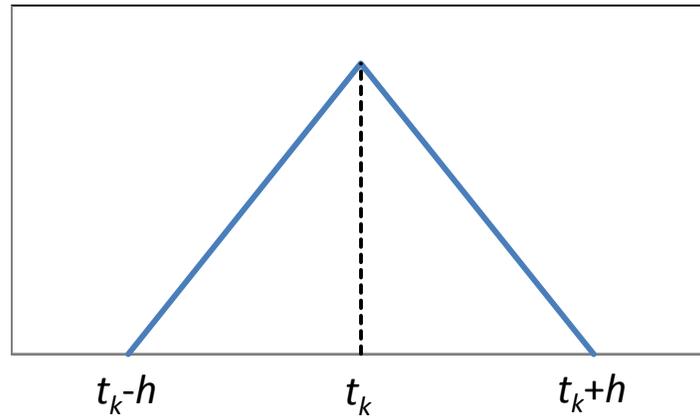


Figure 3.1: Allocation Parameter Illustration

Of course, other formulations could be used here as well, to allow for other shapes (e.g., normal or uniform) or for skew. However, adding further complexity to the model may add estimation challenges. Another option is to specify α as constant across alternatives present in each nest. While this formulation seems simpler, it would not reduce the number of parameters to be estimated in the model (the nest size, h , must still be estimated) and allocation parameters for alternatives at nest boundaries would not be well-defined¹⁶.

To facilitate comparisons with the continuous logit model, equation 3.8's and 3.9's same systematic utility functions are used for the CCNL. In the next section, properties of the CCNL are investigated and model behavior is compared to that of the continuous logit, to illustrate the roles of the allocation and inclusive value parameters.

¹⁶ For instance, if the nest parameter h is 1, then the allocation parameter for the 8 am alternative in the nest centered at 9 am is ill-defined. For the left, it would take a value of zero, while from the right, it would take a value of α . Since numerical integration is required in model estimation, which relies on evaluation of discrete density values, either treatment of such boundary points will affect estimation results, and neither will be correct.

3.2.3 CCNL Model Behavior and Properties

Recently, several papers have investigated the correlation structure implied by the (discrete) CNL model. Here, those results are applied and extended in the continuous context to illuminate correlation structures of the CCNL model.

The covariance between any two discrete CNL random error components can be expressed as follows (for derivation, see, e.g., Papola and Marzano 2005 and Abbe et al. 2007):

$$\text{Cov}(\varepsilon_k, \varepsilon_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \varepsilon_k \varepsilon_j \frac{\partial F(\varepsilon_k, \varepsilon_j)}{\partial \varepsilon_k \partial \varepsilon_j} d\varepsilon_k d\varepsilon_j - (\rho_0 \zeta)^2 \quad (3.16)$$

Here, $F(\varepsilon_k, \varepsilon_j)$ represents the joint cumulative distribution function (CDF) for error terms, ε_k and ε_j , ζ is Euler's constant, and ρ_0 is the scale parameter (set to one in this case as is typical to identify the model). This representation is rather complex, and Marzano and Papola (2008) have shown that equation 3.16 can be rewritten as follows:

$$\text{Cov}(\varepsilon_k, \varepsilon_j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(\varepsilon_k, \varepsilon_j) - F(\varepsilon_k)F(\varepsilon_j)] d\varepsilon_k d\varepsilon_j \quad (3.17)$$

While equation 3.17 still involves integration over the domain of the error terms, the partial derivative of the joint CDF has been eliminated (which is particularly important for extending these results for application in the CCNL context). Here, $F(\varepsilon_k)$ and $F(\varepsilon_j)$ denote the marginal CDFs of the error terms, which, thanks to the normalization of the allocation parameters (so that they sum to 1 across all nests for each alternative), can be written as follows:

$$F(\varepsilon_k) = \exp(-\exp[-(\varepsilon_k - \ln(\sum_{m=1}^M \alpha_{km}))]) = \exp(-\exp[-\varepsilon_k]) \quad (3.18)$$

The right-hand side of equation 3.18 results from the normalization of the allocation parameters. Note here that nests are indexed from 1 to M . The joint CDF can be written as follows:

$$F(\varepsilon_k, \varepsilon_j) = \exp\left(-\sum_{m=1}^M [(\alpha_{km} e^{-\varepsilon_k})^{\rho_m} + (\alpha_{jm} e^{-\varepsilon_j})^{\rho_m}]^{\frac{1}{\rho_m}}\right) \quad (3.19)$$

Here, ρ_m is the inclusive value parameter corresponding to nest m . Note that $\rho_m = \rho \forall m$ under the CCNL specification of Section 3.2.2. Moreover, the summation in equation 3.19 is analogous to an integral under the CCNL specification, which allows ease in writing the joint CDF for the CCNL as follows:

$$F(\varepsilon_k, \varepsilon_j) = \exp\left(-\int_{t_k-h}^{t_j+h} [(\alpha(t_k, t_m) e^{-\varepsilon_k})^\rho + (\alpha(t_k, t_m) e^{-\varepsilon_j})^\rho]^{\frac{1}{\rho}} dt_m\right) \quad (3.20)$$

Here, the allocation parameters, $\alpha(q, r)$, take the functional form shown in equation 3.15. Remember that under the CCNL representation, ε_k and ε_j represent the error components related to elemental alternatives k and j , where alternatives k and j represent departure times of t_k and t_j . Moreover, $t_k < t_j$ ¹⁷ is an implicit assumption.

With this information, it is not difficult to numerically compute correlation coefficients across random error components of the CCNL. However, a couple of observations can be made without computation. First, suppose $j = k + 1$, and t_j is one elemental time unit greater than t_k . In words, suppose the limit as $t_j \rightarrow t_k$ is taken. The joint CDF then reduces to the following (thanks to the normalization of allocation parameters):

¹⁷ While t_k need not be less than t_j , this is assumed to ensure bounds on the integral in equation 3.20 are stated correctly. This does not result in a loss of generality.

$$F(\varepsilon_k, \varepsilon_j) = \exp\left(-[e^{-\rho\varepsilon_k} + e^{-\rho\varepsilon_j}]^{\frac{1}{\rho}}\right) \quad (3.21)$$

This is exactly the joint CDF of nested logit random errors sharing a common nest, with correlation coefficient given by $(1 - \rho^{-2})$. And, since the correlation across any other alternatives separated by a distance greater than zero must be smaller, $(1 - \rho^{-2})$ represents the maximum correlation across any pair of alternatives under the CCNL specification. Of course, one would expect near perfect (or perfect) correlation between alternatives separated by an infinitesimally small time step. One could accommodate such correlations by setting ρ to be arbitrarily large, and not estimating it at all. For instance, if ρ was set to a value of 10 or more, the maximum correlations would be 0.99 or greater.

Another observation that can be made is that the time interval between alternatives need only be measured in units of h (where $2h$ represents the nest interval size). In other words, the correlation across any two alternatives separated by a distance ah , where a is some constant, will be the same, even if h changes. This results from the joint CDF (equation 3.20) only depending on the shared area between the respective allocation parameter functions.

To illustrate the range of correlation coefficients one can achieve under the CCNL model, numerical computation was used. Table 3.1 shows the correlation coefficient for alternatives separated by a variety of distances (measured in units of h) for a variety of inclusive value parameters, ρ .

Table 3.1: Correlation Coefficients for Varying ρ and Distance Variables¹⁸

		ρ						
		1.1	1.25	1.5	2	3	5	10
Distance between Alternatives	0	0.173	0.360	0.555	0.750	0.889	0.960	0.990
	0.2h	0.165	0.341	0.524	0.705	0.831	0.894	0.920
	0.4h	0.145	0.299	0.457	0.610	0.713	0.763	0.782
	0.6h	0.119	0.245	0.372	0.491	0.571	0.607	0.622
	0.8h	0.091	0.186	0.281	0.368	0.425	0.451	0.461
	h	0.064	0.129	0.194	0.254	0.292	0.309	0.315
	1.2h	0.041	0.082	0.123	0.160	0.184	0.195	0.199
	1.4h	0.023	0.046	0.069	0.089	0.102	0.108	0.110
	1.6h	0.010	0.020	0.030	0.039	0.045	0.048	0.049
	1.8h	0.002	0.005	0.008	0.010	0.011	0.012	0.012
2h	0	0	0	0	0	0	0	

Clearly, as the distance between alternatives grows, the correlations shrink, and correlations grow with increasing ρ . One may note that even for very large values of ρ , error term correlations die off rather quickly as the distance between the alternatives exceeds h units.

To illustrate differences between the continuous logit and CCNL specifications, suppose that all parameters of the systematic utility are known (and are the same under the continuous logit and CCNL), and temporarily exclude time-varying covariates from the utility specification. Further, suppose that the only covariate in X is a constant and $\gamma(t)$ is given by the following:

$$\gamma(t) = \left[\sin\left(\frac{2\pi t}{24}\right), \sin\left(\frac{4\pi t}{24}\right), \sin\left(\frac{6\pi t}{24}\right), \cos\left(\frac{2\pi t}{24}\right), \cos\left(\frac{4\pi t}{24}\right), \cos\left(\frac{6\pi t}{24}\right) \right]'$$

Two examples are presented here. In Example 1, the parameter vector is assumed to be $\beta = [2, -1, 1, -2, -0.2, 1]$, and in Example 2, the parameter vector is assumed to be

¹⁸ While not shown in Table 3.1, when $\rho = 1$, the model reduces to the continuous logit and correlations are zero for all distances between alternatives.

$\beta = [2, 0.5, -0.5, -2, 0.5, 0.7]$. Figure 3.2 shows the systematic utility profiles over departure time alternatives for Examples 1 and 2. In addition, 90% confidence intervals around the systematic utilities are also plotted under the assumption of standard Gumbel distributed error terms¹⁹.

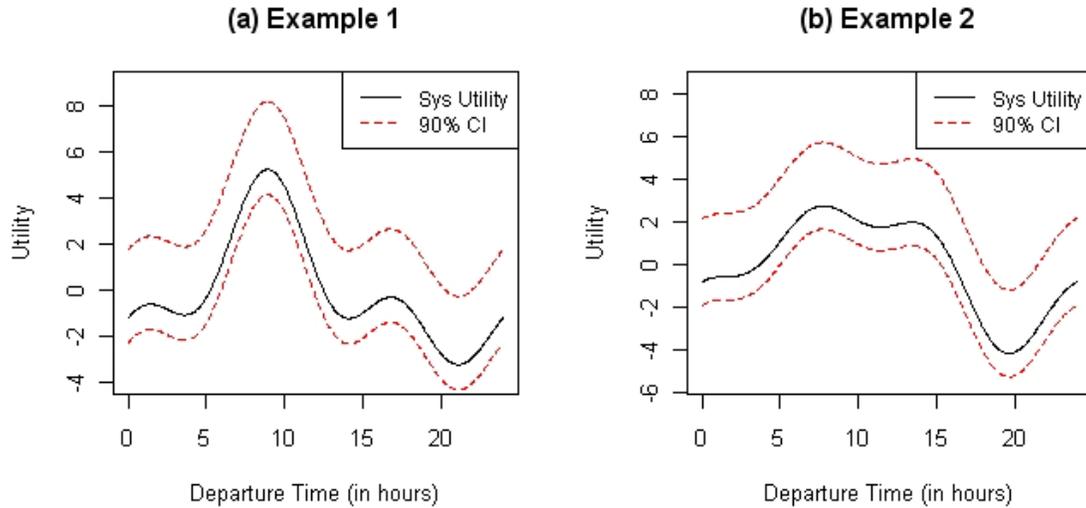


Figure 3.2: Utility Profiles and 90% Confidence Intervals for Examples 1 and 2

Figure 3.3 presents differences in predictive densities for departure time choice between the continuous logit and CCNL under Example 1's utility specification for $h = 0.5$ (30 minutes) and ρ varying from 1 to 15. Clearly, there is very little difference in the predictive densities between the two models, which is largely due to the relatively low value of h here. However, some important conclusions can be drawn from this example. For instance, under this utility specification (for both models), this individual would be very unlikely to choose departure times outside of 6 am to 11 am. In addition, even though the utility function (Figure 3.2a) has three local maxima (occurring at about 2 am, 9 am, and 5 pm), only the 9 am peak has substantial effects on the predictive distribution. This is because the 9 am utility peak represents utility values that are much greater (in a

¹⁹ Note that error terms in the random utilities have variances given by $\frac{\pi^2}{6}$.

relative sense) than other utility values. So, while a 5 pm departure may be more probable for this individual than a 9 pm departure, a 9 am departure is much more probable than both. Of course, if the size of utility values during the AM peak period were closer in magnitude to utility values in the early morning and/or midday or evening, then very different predictive densities would emerge. In Chapter 5's empirical analysis, the size of these utility values emerges via parameter estimation.

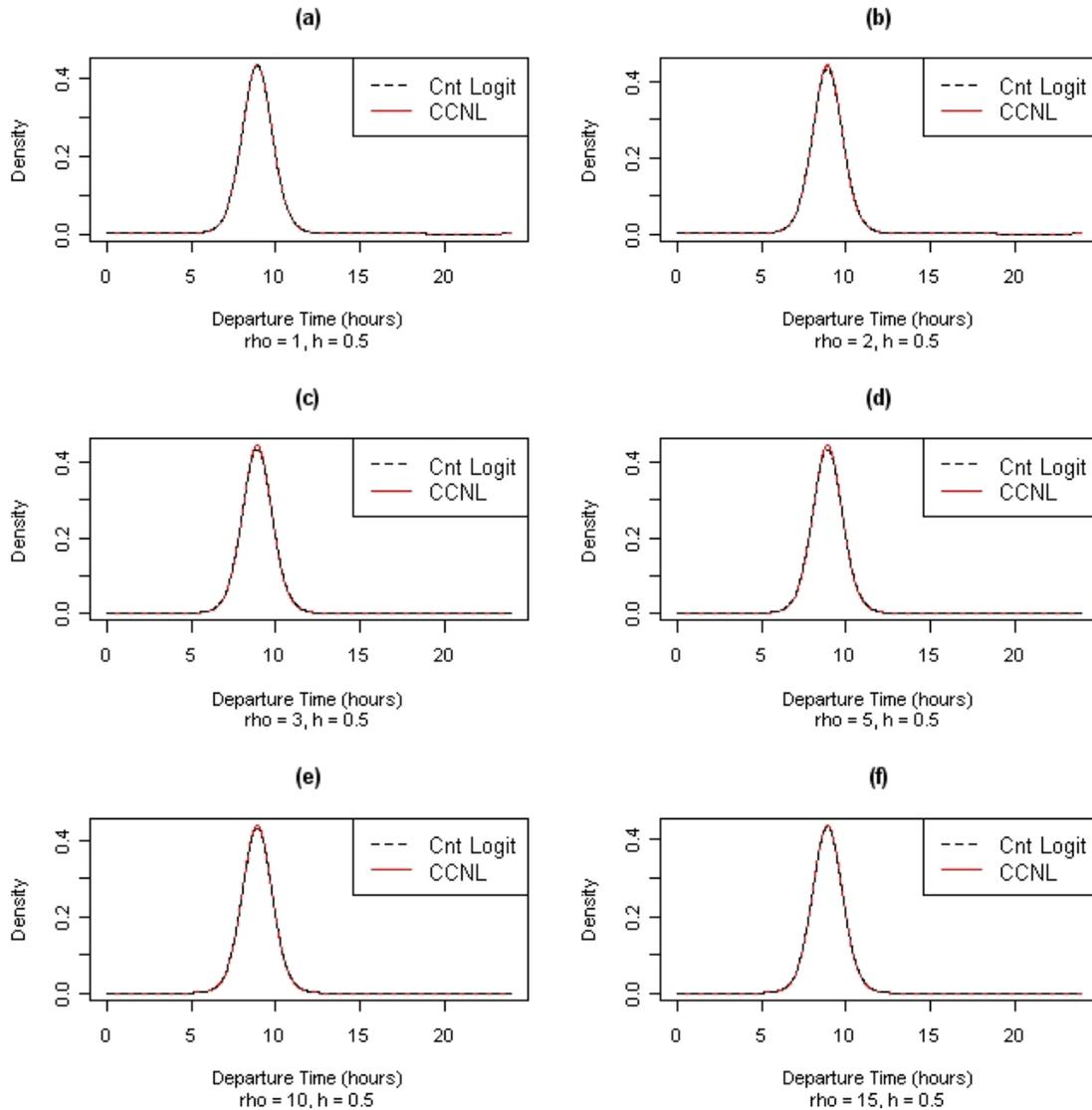


Figure 3.3: Predictive Densities with h of 0.5 and Varying Values of ρ for Example 1

Figure 3.4 shows similar predictive densities, but with h chosen as 1.0 (rather than Figure 3.3's $h = 0.5$). Like Figure 3.3, the differences in predictive densities between the continuous logit and CCNL are not very large, though some difference is perceivable. In particular, the heights of the CCNL's predictive density peaks are slightly higher as ρ grows in value. This is, in fact, a property of the CCNL. One cannot achieve peak

heights less than the continuous logit's peak heights for identical systematic utility functions.

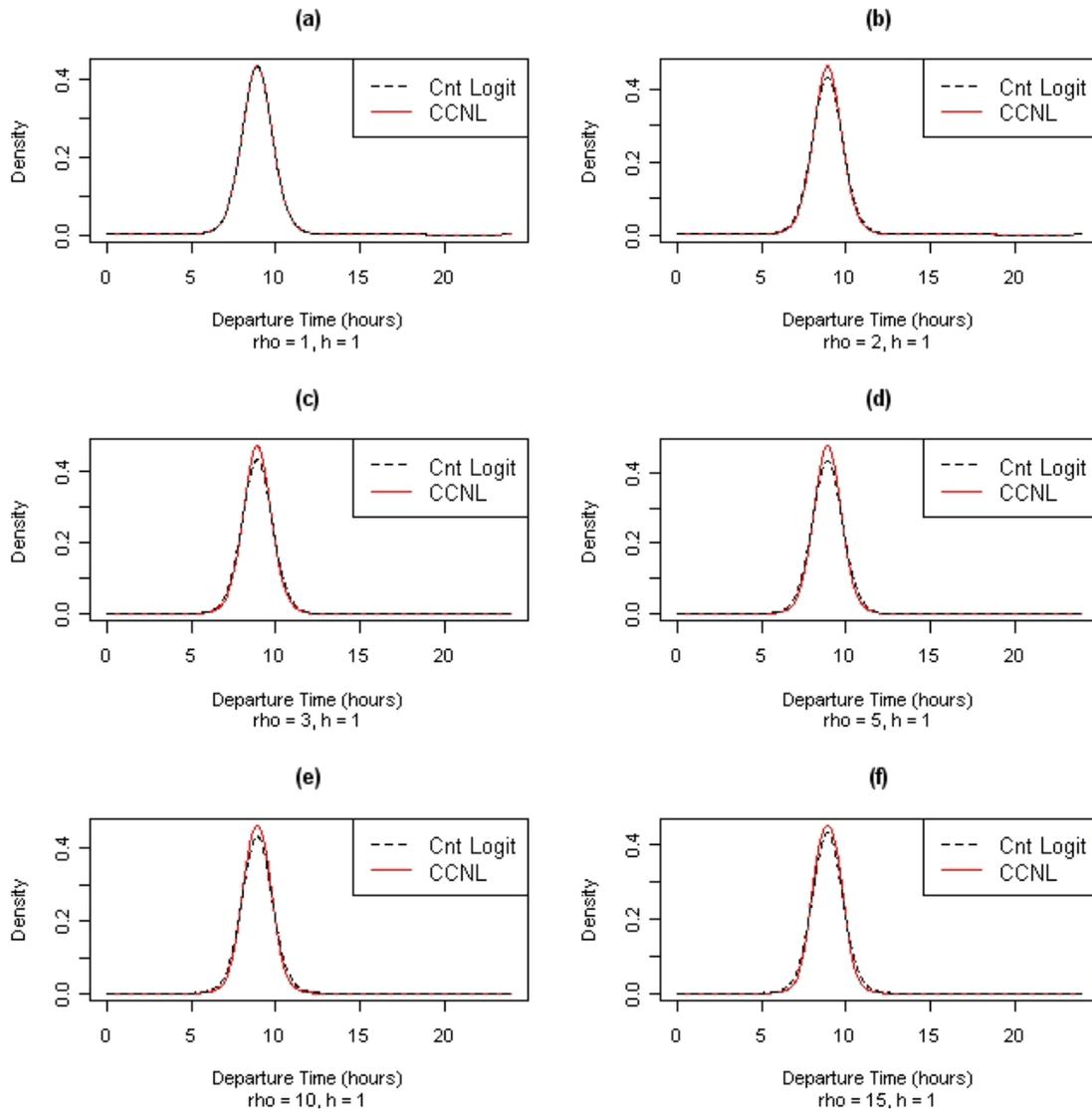


Figure 3.4: Predictive Densities with h of 1.0 and Varying Values of ρ for Example 1

Figure 3.5 plots predictive densities for Example 1's utility specification with h chosen as 2.0. With the larger value of h here, very noticeable differences between continuous logit and CCNL predictive densities emerge. But again, like Figure 3.4, predictive density

peaks only grow (with increasing values of ρ) in relation to the continuous logit density peaks. In addition, Figure 3.5 depicts how, relative to the continuous logit, the base of the peak density shrinks in size with increasing values of ρ , and the peaks become more rounded. In fact, with $\rho = 15$, there is a noticeable valley within the peak portion of the density.

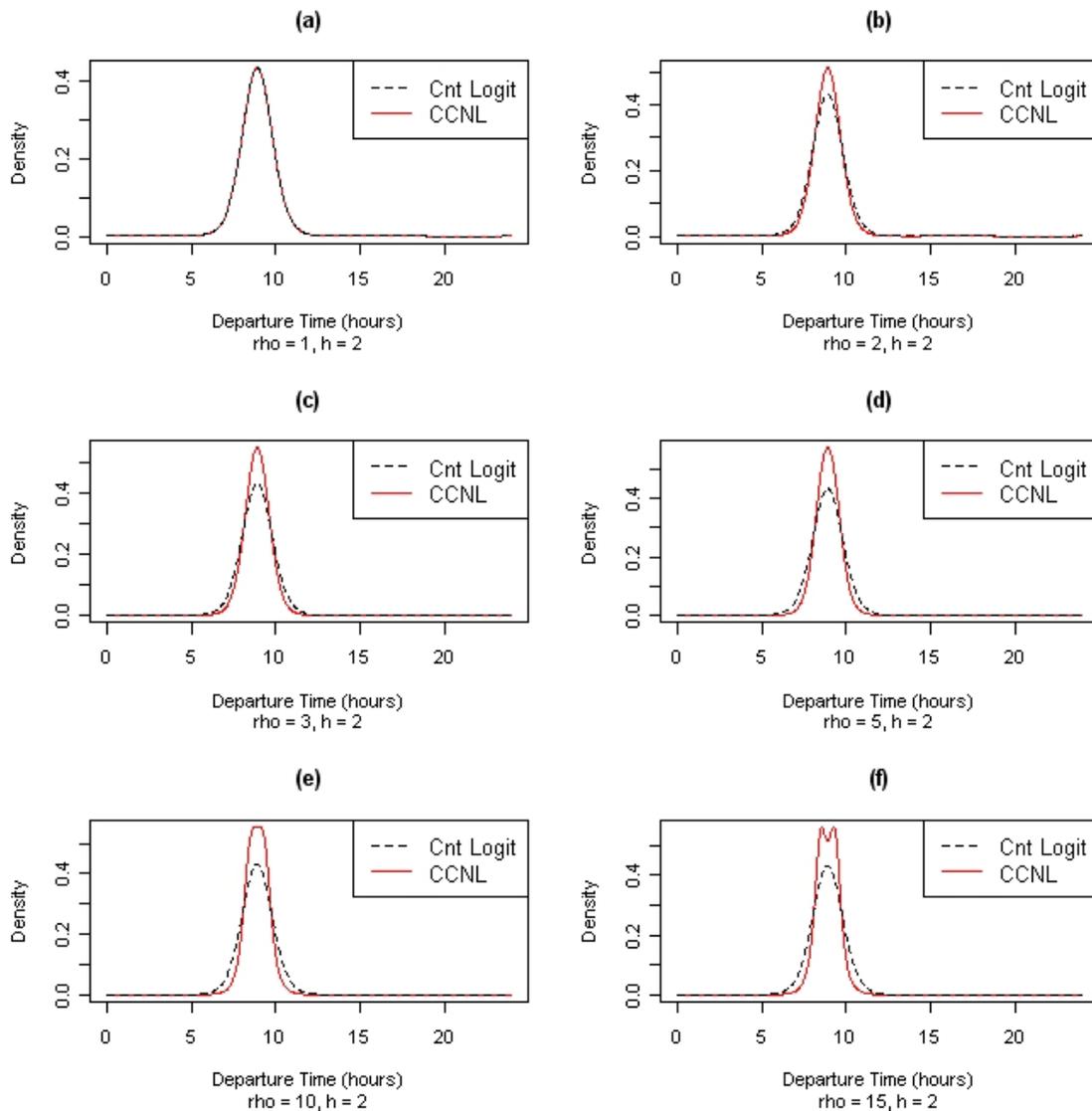


Figure 3.5: Predictive Densities with h of 2.0 and Varying Values of ρ for Example 1

Example 2 illustrates a slightly different utility specification. Figure 3.6 shows predictive densities for the continuous logit and CCNL (with h of 2.0 and a variety of ρ values). The only difference between Examples 1 and 2 is the utility profile, as shown in Figure 3.2. Example 2's utility profile represents a bimodal peak utility situation.²⁰ This is evident in the predictive density profiles of Figure 3.6. As with Example 1, one can again notice that the two peaks in the predictive densities are higher under the CCNL specification (with increasing values of ρ), not surprisingly. Moreover, as in Example 1, the base (or spread) of the peaks shrinks in size with increasing values of ρ , unlike the continuous logit.

²⁰ While Example 1's utility profile also exhibits bimodality, because the relative difference between utilities from the highest peak and the other peaks were so large, they did not contribute in any meaningful way to the shape of predictive densities.

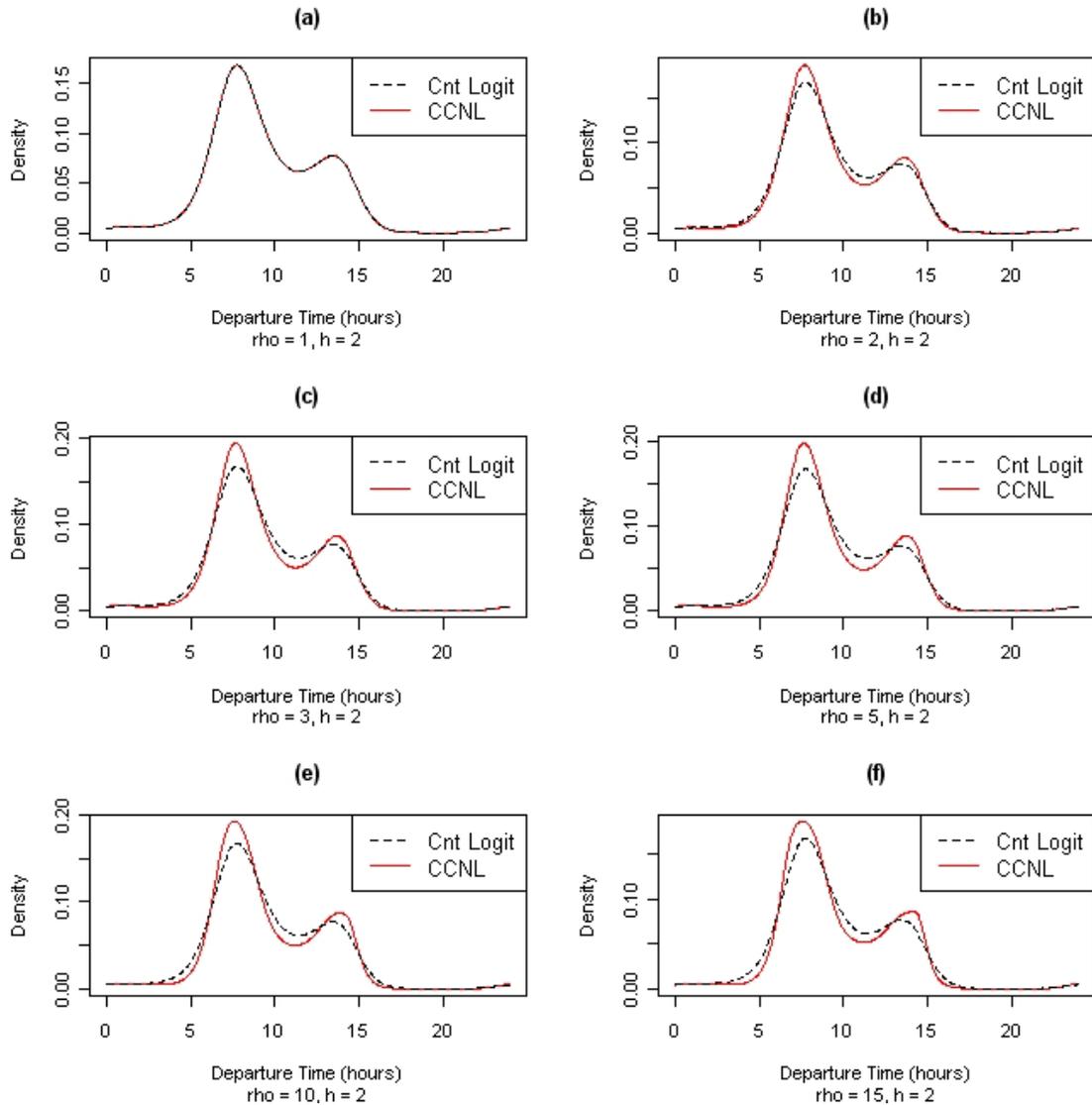


Figure 3.6: Predictive Densities with h of 2.0 and Varying Values of ρ for Example 2

3.2.4 CCNL Parameter Estimation via MCMC Simulation

Like the continuous logit of Section 3.1, the CCNL is estimated here using Bayesian techniques. And like the continuous logit, the β prior is chosen to be independent and normally distributed and the η prior follows a log-normal distribution multiplied by minus one, both with vague (i.e., large variance) parameters. The CCNL contains two additional parameters (h and ρ) for which a normal prior would not be reasonable.

Instead, priors for h and ρ are chosen as independent gamma distributions, bounded *below* by 0.25 and 1.0, respectively²¹. Shape and scale parameters for the prior distributions are chosen as 1.0 and 0.5 for both. These priors offer the model information on h and ρ (unlike the vague priors selected for β and η), which has the effect of pulling them closer to their respective left-side bounds. This is quite reasonable for h , since h specifies the minimum time interval between uncorrelated alternatives (and one would not expect correlations between alternatives a great distance apart). While similar expectations may not exist for ρ , its prior can be viewed as follows: unless the data offer significant proof for another value, the prior guides ρ to a value of one, thereby reducing the model to the continuous logit. Also, while h is not required to be at least 0.25 in model formulation, this restriction aids in numerical integration computations. For small h , a large number of function evaluations would be required to obtain reasonable integral estimates using Simpson's rule (Press et al. 1989).²² As shown in Figure 3.3, when h is equal to 0.5, there appears to be very little difference between the continuous logit and CCNL anyway (though Figure 3.3 illustrates the difference only for a specific utility function specification). Additionally, since values of ρ close to 1 will effectively eliminate the impact of h on the model, this should not be a sizable issue in model estimation. However, this means that if correlations exist in departure time choice (i.e., $\rho > 1$), departure time alternatives a minimum of 0.5 hours apart will exhibit some amount of correlation.

The Bayesian inference proceeds in much the same way as the continuous logit, where a one-step MH algorithm is employed. Of course, with the CCNL there are two additional parameters, h and ρ . While it may seem logical to draw the utility parameters, β and η , separately from the structural parameters, h and ρ (i.e., a two-step algorithm with two MH proposals), this was not sought here because with each step of the algorithm, the

²¹ The lower bound on h was chosen to aid in numerical integration computations. In the case of ρ , values less than its lower bound of 1.0 are not consistent with random utility theory.

²² With fixed integration bounds and more function evaluations, errors in numerical integration estimates will be smaller.

likelihood must be computed, which requires a great deal of computational effort. The proposal density is again constructed as a multivariate normal density with mean equal to the current draw's parameter values (i.e., the normal random walk proposal). Also, like the continuous logit, the covariance matrix of the proposal is initially set to zeroes on the off-diagonal elements and very small values for the diagonal elements to ensure high acceptance rates early in the estimation process. After 2,000 draws are obtained, the covariance matrix is estimated from the previous draws (not including the current draw); and, after 5,000 draws are obtained, the covariance matrix is estimated from the previous 5,000 draws (not including the current draw). It has been shown that an adaptive MH algorithm of this kind should generate draws that converge to the correct posterior distribution (Holden et al. 2009).

It should be noted that using this proposal density, it is possible to draw values of ρ less than 1.0 and/or values of h less than 0.25. In such cases, the entire set of parameters is re-drawn from the same proposal density until an acceptable set of parameters is obtained. In other words, the proposal density essentially amounts to a truncated normal distribution, where ρ is truncated from below by 1.0 and h is truncated from below by 0.25.

Finally, like the continuous logit, only every 50th draw is used for inference after draws appear to have converged, which reduces the autocorrelation across draws. Again, convergence here is diagnosed by a combination of informal graphical techniques as well as Geweke's (1992) diagnostic.

3.3 Bivariate Multinomial Probit

Like the MNL, the multinomial probit (MNP) relies on a latent random utility specification. However, unlike the MNL, the random error terms are not distributed according to a type I extreme value distribution (i.e., a gumbel); instead, they follow a normal distribution. Of course, normality results in open-form expressions for alternative

probabilities (unlike the MNL), which is why the MNP has not been utilized to a greater extent in the literature. Thanks to Bayesian and other sophisticated statistical methods, one need not assume error terms are independent and identically distributed with the MNP. In this section, a bivariate MNP (BVMNP) model specification for tour TOD choice is formulated, where the twin variables of interest are a tour's home-to-work arrival time and work-to-home departure time. Unlike the continuous logit and CCNL models, these two components are treated as discrete alternatives rather than continuous ones. However, the BVMNP here offers the ability to capture both timing dimensions of a tour, rather than a single timing dimension characterized by the continuous logit and CCNL models developed in previous sections of this chapter.²³

3.3.1 Random Utility Framework and Model Specification

The MNP model is a random utility model, like the MNL and other GEV-type models. One assumes that each alternative has a (latent) random utility, and the decision-maker always chooses the alternative offering the greatest underlying utility value. However, unlike the GEV class of models, the random error components are distributed according to a multivariate normal distribution, rather than an extreme value distribution.

Moreover, the MNP model places no restrictions on the covariance structure of random error components across alternatives (though some restrictions are needed for statistical identification in model estimation). Most closed-form GEV models, on the other hand, rely on some specific covariance structures, generally predetermined by the analyst through specification of the GEV generating function. Of course, restrictions on the covariance structure are not always unwelcome. For instance, in the case of departure time choice with discrete time-interval alternatives, one expects adjacent alternatives to

²³ It should be noted that both the continuous logit and CCNL models could (theoretically) be reformulated to accommodate TOD choice for two dimensions of a travel tour. In the case of the continuous logit, model estimation for two dimensions may be feasible; however, model estimation of the CCNL with two timing dimensions would be computationally prohibitive due to its multidimensional integration required for each timing dimension.

be more correlated than non-adjacent ones, similar to the correlation structure imposed with the CCNL in a continuous choice setting (Section 3.2.3). Nonetheless, implied correlations for the GEV class of models are rarely straightforward, largely because model estimation utilizes the generating function to compute choice probabilities, so the underlying distribution of correlated error terms is unnecessary (not to mention that covariance between correlated extreme values does not in itself define a multivariate extreme value distribution). Conversely, the MNP model relies on explicit estimation of the covariance matrix. Thus, error term correlations are easily recognizable. This represents a key motive for choosing the MNP here, rather than a GEV-type model. The MNP allows the analyst to posit a specific relationship between error terms first, and then derive the implied covariance structure.

Since the MNP model developed here is intended to offer a format for the two-dimensional TOD choice of a travel tour, the choice context needs special attention. One reasonable way to approach the problem is to consider it in a single dimension. Instead of choosing tour arrival times and tour return times, one may assume that individuals jointly choose tour arrival and return times, and the analyst need only consider a single choice dimension. This is exactly how many researchers have modeled tour TOD choice (see, e.g., Vovsha and Bradley 2004, Abou Zeid et al. 2006, and Popuri et al. 2008, among others)²⁴. In a TOD choice context, this method makes good sense, since the analyst need only consider a single set of utility functions for a single choice context and can directly include tour duration-specific elements in the utility function. For instance, consider the following joint utility specification:

$$U(t_a, t_r) = V_1(t_a) + V_2(t_r) + V_3(t_r - t_a) + \varepsilon_{ar} \quad (3.22)$$

²⁴ Note that all existing models of this type have been estimated in a MNL setting, ignoring correlations across arrival and return times.

Here, V_1 is the systematic utility component related to arrival time t_a , V_2 is the component related to return time t_r , and V_3 is the component related to duration $t_r - t_a$. A key difficulty of this method is that one is usually interested in rather small time intervals as alternatives; and, in two dimensions, the number of alternatives can become quite large. For instance, if 1-hour intervals are used across the 24-hour day, the number of alternatives becomes $24 \times 25 / 2 = 300$ (assuming one cannot return earlier than one departs). If 30-minute intervals are used, one has 1,176 alternatives. For a MNP model, this produces a covariance matrix of size 1,176 x 1,176, presenting a number of computational difficulties in model estimation (which worsens if interval size is less than 30 minutes). With this in mind, a bivariate multinomial probit (BVMNP) model is developed here, where tour arrival time represents one choice dimension and tour return time represents another choice dimension. While the BVMNP model has been used in previous studies (see, e.g., Chib and Greenberg 1998, Golob and Regan 2002, and Zhang et al. 2008), no previous work has investigated choice contexts with more than 3 or 4 alternatives. In addition, the estimation procedure used here varies from traditional methods to accommodate the large number of alternatives. In this bivariate context, one must specify two separate utility functions (one for tour arrival and another for tour return), as follows:

$$U_{aj} = V_{aj} + \varepsilon_{aj} \quad (3.23)$$

$$U_{rl} = V_{rl} + \varepsilon_{rl} \quad (3.24)$$

Here, U_{aj} and U_{rl} denote latent utilities for arrival and return time alternatives j and l , V_{aj} and V_{rl} are systematic utility components, and ε_{aj} and ε_{rl} are random error components. The set of arrival time alternatives is identical to the set of return time alternatives, with arrival time alternatives indexed by $j = 1, \dots, J$ and return time alternatives indexed by $k = 1, \dots, J$. While this specification does not allow for a utility component specifically related to tour/activity duration, it does significantly reduce the number of choice alternatives. For instance, if time-of-day is modeled in 30-minute intervals over the 24-

hour day period (as it is here), this results in 96 alternatives (and utility values), rather than the 1,176 needed for the joint choice model. Note that no assumptions regarding the error terms (ε_{aj} and ε_{rl}) have yet been made here, other than normality.

3.3.2 Error Correlation Structure

Since one cannot reasonably assume independence of alternatives, the correlation structure of the error components deserves some attention. While it is theoretically feasible to estimate the entire covariance matrix without imposing any pre-specified structure, this is likely to result in some strange parameter estimates due to the high number of alternatives (relative to the sample size of 997 used here). In addition, there is a clear ordering of alternatives, which evokes certain expectations for covariance properties. With this in mind, a specific structure is imposed here.

One can imagine a variety of correlation structures. For instance, an autoregressive (AR) process suggests that each error term is a function of earlier (in time) error terms and a random noise component. While such a process suggests that the utility of each alternative is realized in sequence (rather than viewed simultaneously), it does offer a rather simple covariance matrix structure. However, it does not offer a method of introducing correlations between arrival time and return time error components. Alternatively, one could turn to simultaneous autoregressive (SAR) or conditional autoregressive (CAR) processes, both of which have been used in the spatial econometrics literature (see, e.g., Smith and LeSage 2004 and Kissling and Carl 2008 for SAR, and Cressie 1995, Lichstein et al. 2002, and Parent and LeSage 2008 for CAR). One could also turn to specification of the covariance matrix components explicitly (rather than relying on a data generating process, like AR, SAR, or CAR processes). Here, two formulations are pursued: the CAR specification and a pseudo AR specification of covariance components. The CAR specification is used (over AR or SAR) largely because it is simpler to interpret and it offers advantages for Bayesian

estimation. With the CAR specification, relationships between error components can be written as follows:

$$\varepsilon_{aj}|\{\varepsilon_{aq}, q \neq j\}, \{\varepsilon_{rl} \forall l\} = \lambda_{CAR,a} \sum_{q \neq j} w_{jq} \varepsilon_{aq} + \lambda_{CAR,d} \sum_l c_{jl} \varepsilon_{rl} + \kappa_{aj} \quad (3.25)$$

$$\varepsilon_{rl}|\{\varepsilon_{rq}, q \neq l\}, \{\varepsilon_{aj} \forall j\} = \lambda_{CAR,r} \sum_{q \neq l} w_{lq} \varepsilon_{rq} + \lambda_{CAR,d} \sum_j c_{jl} \varepsilon_{aj} + \kappa_{rl} \quad (3.26)$$

Here, $\lambda_{CAR,a}$, $\lambda_{CAR,r}$, and $\lambda_{CAR,d}$ are parameters to be estimated; and κ_{aj} and κ_{rl} are independent random error terms for alternatives j and l , respectively, distributed normally with zero means and variances of $\sigma_{CAR,a}^2$ and $\sigma_{CAR,r}^2$ for arrival and departure time alternatives, respectively. w_{pq} measures the degree of “closeness” between alternatives p and q , and c_{pq} measures the degree of “closeness” of the duration implied by arrival and departure time alternatives p and q to some baseline/preferred duration. Note that w_{pq} and c_{pq} represent weights chosen by the analyst. For w_{pq} , it is assumed that conditional on all error terms, only alternatives adjacent to the alternative of interest are needed. In other words, ε_{aj} is conditionally independent of all other arrival times except those adjacent to the j^{th} arrival time, and likewise for ε_{rl} and other return times. Formally, w_{pq} is expressed as follows:

$$w_{pq} = \begin{cases} 1 & \text{if } |p - q| = 1 \\ 0 & \text{otherwise} \end{cases} \quad (3.27)$$

Note that the conditional error structure is the same for arrival time and return time sets (i.e., w_{pq} is the same regardless of whether p and q denote arrival times or return times). Also, it is important to note that this specification does not mean that correlations between non-adjacent alternatives are zero. For c_{pq} , one does not necessarily expect conditional independence. Instead, c_{pq} is specified to fall as the time between the implied duration of the joint alternatives and the baseline/preferred duration. In addition,

these terms are set to zero for joint alternatives that are not possible (i.e., when $t_p > t_q$), as shown in equation 3.28.

$$c_{pq} = \begin{cases} 0 & \text{for } t_p > t_q \\ \left(\frac{1}{|(t_q - t_p) - (\mu_{CAR,1} + \mu_{CAR,2} t_p)| + 1} \right)^2 & \text{for } t_p \leq t_q \end{cases} \quad (3.28)$$

Here, $\mu_{CAR,1} - \mu_{CAR,2} t_p$ represents the baseline duration at arrival time choice p , and $\mu_{CAR,1}$ and $\mu_{CAR,2}$ are parameters to be estimated. For notational convenience, one can define the following:

$$W_a = \begin{bmatrix} \lambda_{CAR,a} w_{11} & \cdots & \lambda_{CAR,a} w_{1J} \\ \vdots & \ddots & \vdots \\ \lambda_{CAR,a} w_{J1} & \cdots & \lambda_{CAR,a} w_{JJ} \end{bmatrix}, W_r = \begin{bmatrix} \lambda_{CAR,r} w_{11} & \cdots & \lambda_{CAR,r} w_{1J} \\ \vdots & \ddots & \vdots \\ \lambda_{CAR,r} w_{J1} & \cdots & \lambda_{CAR,r} w_{JJ} \end{bmatrix},$$

$$C = \begin{bmatrix} \lambda_{CAR,d} c_{11} & \cdots & \lambda_{CAR,d} c_{1J} \\ \vdots & \ddots & \vdots \\ \lambda_{CAR,d} c_{J1} & \cdots & \lambda_{CAR,d} c_{JJ} \end{bmatrix}, C' = \begin{bmatrix} \lambda_{CAR,d} c_{11} & \cdots & \lambda_{CAR,d} c_{J1} \\ \vdots & \ddots & \vdots \\ \lambda_{CAR,d} c_{1J} & \cdots & \lambda_{CAR,d} c_{JJ} \end{bmatrix},$$

$$\varepsilon = \begin{bmatrix} \varepsilon_{a1} \\ \vdots \\ \varepsilon_{aJ} \\ \varepsilon_{r1} \\ \vdots \\ \varepsilon_{rJ} \end{bmatrix}, \kappa = \begin{bmatrix} \kappa_{a1} \\ \vdots \\ \kappa_{aJ} \\ \kappa_{r1} \\ \vdots \\ \kappa_{rJ} \end{bmatrix}$$

Here, W_a and W_r represent the matrices of w_{pq} elements multiplied by $\lambda_{CAR,a}$ or $\lambda_{CAR,r}$, C is the matrix of c_{pq} elements multiplied by $\lambda_{CAR,d}$, ε is a vector of correlated random error terms beginning with arrival time alternatives (ordered from the first to the last), and κ is a vector of independent random error terms (ordered in the same way as ε). J denotes the number of 30-minutes intervals. With the CAR specification, the distribution of ε can now be written as follows:

$$\varepsilon \sim N(0, \Sigma_{CAR}) \quad (3.29)$$

$$\text{where } \Sigma_{CAR} = [I_{2J} - B]^{-1}S, \quad S = \begin{bmatrix} \sigma_{CAR,a}^2 I_J & 0 \\ 0 & \sigma_{CAR,r}^2 I_J \end{bmatrix},$$

$$B = \begin{bmatrix} W_a & \frac{1}{\sigma_{CAR,r}^2} C \\ \frac{1}{\sigma_{CAR,a}^2} C' & W_r \end{bmatrix} \quad (3.30)$$

I_J is the identity matrix of dimension J and I_{2J} is the identity matrix of dimension $2J$. Note that $\sigma_{CAR,a}^2$ and $\sigma_{CAR,r}^2$ are included in the diagonal elements of B to ensure that Σ_{CAR} is symmetric, since Σ_{CAR} is only symmetric if $b_{pq}s_{qq} = b_{qp}s_{pp}$ (Parent and LeSage 2008). Furthermore, Σ_{CAR} must be positive definite and is only positive definite for certain ranges of $\lambda_{CAR,a}$, $\lambda_{CAR,r}$, and $\lambda_{CAR,d}$. Parent and LeSage (2008) show that $\lambda_{CAR,a}$ and $\lambda_{CAR,r}$ must lie between q_{\min}^{-1} and q_{\max}^{-1} , where q_{\min} and q_{\max} are the smallest and largest eigenvalues of W . Since only positive correlations are expected here, q_{\min}^{-1} is replaced with 0. For $\lambda_{CAR,d}$, there is seemingly no simple restriction like this. But it can be shown that, if $\lambda_{CAR,a} = \lambda_{CAR,r} = q_{\max}^{-1}$, then $\lambda_{CAR,d} = 0$, and, if $\lambda_{CAR,a} = \lambda_{CAR,r} = \mathcal{B}_{\max}^{-1}$ (where \mathcal{B}_{\max} is the maximum eigenvalue of B), then the maximum value of $\lambda_{CAR,d}$ is \mathcal{B}_{\max}^{-1} .²⁵ The issue of parameter value bounds for $\lambda_{CAR,a}$, $\lambda_{CAR,r}$, and $\lambda_{CAR,d}$ is discussed in more detail in the next section.

One last element needing attention here is the role of $\mu_{CAR,1}$ and $\mu_{CAR,2}$ (which define the “baseline” duration on which elements of C are based). Essentially, the model posits that some activity duration may be highly desired (e.g., 8 hours for full-time workers), and this term allows correlations across arrival and return time utilities to be highest for such durations. It is not reasonable to view $\mu_{CAR,1}$ and $\mu_{CAR,2}$ as only two parameters, since one expects differences across individuals or classes of individuals. This is particularly important since duration is ignored in the systematic utility equations. Here, $\mu_{CAR,1}$ and $\mu_{CAR,2}$ are taken to be two separate parameters each, two for full-time workers making no

²⁵ Here it is assumed that $\lambda_{CAR,d} > 0$.

additional tours during the day ($\mu_{CAR,1,full}$ and $\mu_{CAR,2,full}$) and two for part-time workers and/or those making additional tours ($\mu_{CAR,1,part}$ and $\mu_{CAR,2,part}$), adding a layer of observed heterogeneity to the model.²⁶ Of course, this is a rather simplistic formulation since one may expect preferred durations to vary with other traveler attributes. The reason $\mu_{CAR,1}$ and $\mu_{CAR,2}$ are only differentiated between full-time workers with no additional tours and part-time workers and those making additional tours is that this distinction seems most important. Adding other segments and/or estimating $\mu_{CAR,1}$ and $\mu_{CAR,2}$ as functions of individual attributes is preferred, but will add complication to the estimation process. For instance, if $\mu_{CAR,1}$ and $\mu_{CAR,2}$ differed for each individual, Σ_{CAR} would also differ for each individual, requiring computation of distinct Σ_{CAR} 's for each observation, which can be computationally expensive. By allowing $\mu_{CAR,1}$ and $\mu_{CAR,2}$ to vary over only two traveler groups, the estimation process is more streamlined, with only two covariance matrices, $\Sigma_{CAR,full}$ and $\Sigma_{CAR,part}$, and facilitates demonstration of the model here.

In the second specification, components of the covariance matrix are formulated directly, with the upper left and lower right quadrants taking on forms *similar to* a typical AR1 process (though it is worth noting that the formulation cannot be directly interpreted as an AR1 process). Off-diagonal quadrants are formulated slightly different, though covariance components appear similar to those of an AR1 process. While the specification is not a typical AR1 process, it will be referred to as the AR1 specification here. The covariance matrix is specified as follows:

$$\Sigma_{AR1} = \begin{bmatrix} \mathcal{W}_a & c \\ c' & \mathcal{W}_r \end{bmatrix} \quad (3.31)$$

²⁶ The distinction here was chosen because average travel durations for full-time workers in the data sample were found to be about 8.2 hours, while average travel durations for both part-time workers and those making additional tours during the day were found to be about 6 hours.

$$\mathcal{W}_a = \begin{bmatrix} w_{a11} & \cdots & w_{a1J} \\ \vdots & \ddots & \vdots \\ w_{aJ1} & \cdots & w_{aJJ} \end{bmatrix}, w_{apq} = \lambda_{AR1,a}^{|t_p-t_q|} \sigma_{AR1,a}^2$$

$$\mathcal{W}_r = \begin{bmatrix} w_{r11} & \cdots & w_{r1J} \\ \vdots & \ddots & \vdots \\ w_{rJ1} & \cdots & w_{rJJ} \end{bmatrix}, w_{rpq} = \lambda_{AR1,r}^{|t_p-t_q|} \sigma_{AR1,r}^2$$

$$\mathcal{C} = \begin{bmatrix} c_{11} & \cdots & c_{1J} \\ \vdots & \ddots & \vdots \\ c_{J1} & \cdots & c_{JJ} \end{bmatrix}, c_{pq} = \begin{cases} 0 & \text{for } p > q \\ \sigma_{AR1,a} \sigma_{AR1,r} \lambda_{AR1,d}^{|(t_q-t_p)-(\mu_{AR1,1}+\mu_{AR1,2}t_p)|+1} & \text{for } p \leq q \end{cases}$$

Like in the CAR specification, $\mu_{AR1,1}$ and $\mu_{AR1,2}$ are segmented by traveler type (i.e., full-time workers with no additional tours and part-time workers and/or those with additional tours). While the AR1 specification does not offer a simple relationship between error components (like the CAR specification does), it does offer greater ease in understanding the covariance matrix components. In addition, the structure implies that correlations between arrival and return time alternatives that are not possible (i.e., when return time is earlier than arrival time) will be zero. With the CAR specification, these alternatives will exhibit some amount of correlation, though it will be much lower than correlations between alternative combinations that are feasible.

Finally, the systematic utility specification for arrival time and return time utilities takes the form shown in equation 3.8. Since each alternative represents a discrete time interval, t in the utility equation is taken to be the midpoint of the time interval. For notational convenience, the systematic utilities for arrival time and return time alternatives are rewritten as follows:

$$V_{i,aj} = X_{ij}\beta_a + \sum_{p=1}^{\mathcal{P}} \eta_{ap} g_{i,ap}(t_j) \quad (3.32)$$

$$V_{i,rl} = X_{il}\beta_r + \sum_{p=1}^{\mathcal{P}} \eta_{rp} g_{i,rp}(t_l) \quad (3.33)$$

Here, $g_{i,ap}(t_j)$ and $g_{i,rp}(t_l)$ represent network characteristics of type p (such as travel time and reliability) for arrival and return time intervals t_j and t_l ; and β_a , β_r , η_a , and η_r

are parameters to be estimated. \mathcal{X}_{ij} and \mathcal{X}_{il} represent row vectors of individual-specific attributes interacted with cyclical functions of the form shown in equation 3.9 (like the utility formulations of Popuri et al. 2008). That is, \mathcal{X}_{ij} and \mathcal{X}_{il} have the following forms:

$$\begin{aligned}
\mathcal{X}_{ij} &= \begin{bmatrix} X_{i1} \sin(2\pi t_j/24) \\ X_{i1} \sin(4\pi t_j/24) \\ \vdots \\ X_{i1} \sin(2Q_1\pi t_j/24) \\ X_{i1} \cos(2\pi t_j/24) \\ \vdots \\ X_{i1} \cos(2Q_1\pi t_j/24) \\ X_{i2} \sin(2\pi t_j/24) \\ \vdots \\ X_{i2} \cos(2Q_2\pi t_j/24) \\ \vdots \\ X_{iK} \sin(2\pi t_j/24) \\ \vdots \\ X_{iK} \cos(2Q_K\pi t_j/24) \end{bmatrix}', & \mathcal{X}_{il} &= \begin{bmatrix} X_{i1} \sin(2\pi t_l/24) \\ X_{i1} \sin(4\pi t_l/24) \\ \vdots \\ X_{i1} \sin(2Q_1\pi t_l/24) \\ X_{i1} \cos(2\pi t_l/24) \\ \vdots \\ X_{i1} \cos(2Q_1\pi t_l/24) \\ X_{i2} \sin(2\pi t_l/24) \\ \vdots \\ X_{i2} \cos(2Q_2\pi t_l/24) \\ \vdots \\ X_{iK} \sin(2\pi t_l/24) \\ \vdots \\ X_{iK} \cos(2Q_K\pi t_l/24) \end{bmatrix}', \\
\mathcal{X}_i &= \begin{bmatrix} \mathcal{X}_{i1} & 0 & g_{i,a}(t_1) & 0 \\ \vdots & 0 & \vdots & 0 \\ \mathcal{X}_{ij} & 0 & g_{i,a}(t_j) & 0 \\ 0 & \mathcal{X}_{i1} & 0 & g_{i,r}(t_1) \\ 0 & \vdots & 0 & \vdots \\ 0 & \mathcal{X}_{ij} & 0 & g_{i,r}(t_j) \end{bmatrix} \\
\begin{bmatrix} V_{i,a} \\ V_{i,r} \end{bmatrix} &= \mathcal{X}_i \begin{bmatrix} \beta_a \\ \beta_r \\ \eta_a \\ \eta_r \end{bmatrix} & & (3.34)
\end{aligned}$$

The number of individual-specific attributes is K , with each individual attribute interacted with $2Q_k$ cyclical functions (Q_k for sine functions and Q_k for cosine functions). By construction, the systematic utilities are linear in unknown parameters. The following section details the estimation procedure for the BVMNP model.

3.3.3 BVMNP Parameter Estimation via MCMC Simulation

Estimation of the BVMNP model is performed via MCMC simulation, like the CCNL and continuous logit models. For brevity and since both estimation procedures are identical in nearly every way, no distinction between the CAR and AR1 covariance specifications is made here. Bayesian techniques are particularly well-suited for estimation of the BVMNP (or any MNP for that matter) since classical methods generally rely on simulated maximum likelihood estimation (MSLE) to avoid numerical evaluation of multi-dimensional integrals involved in the likelihood (McFadden 1989 and Geweke et al. 1994). McCulloch and Rossi (1994) note that MSLE approaches have been found to be sensitive to choice probability estimation methodology.

In the standard Bayesian construction of the MNP model (see, e.g., Albert and Chib 1993, McCulloch and Rossi 1994, and Zhang et al. 2008, among others), one need not evaluate choice probabilities at all. For the MNP model, the dependent variable, Y_i , can take on values 1, 2, ..., J , where Y_i 's value simply indexes the chosen alternative. With the latent random utility specification of the model, the probability of Y_i taking on a value q is given by the following:

$$P(Y_i = q) = P(U_{iq} \geq \max_{p \in J} U_{ip}) \quad (3.35)$$

In other words, the choice probability of alternative q is equivalent to the probability that the latent utility associated with alternative q is the maximum utility value. Here, U_i is treated as a random (nuisance) parameter to be estimated and is normally distributed (under the MNP model specification), with mean given by the systematic utility, V_i , and variance given by Σ . For the BVMNP model, Y_i is simply taken to be bivariate, with joint choice probability of arrival time q_1 and return time q_2 given by the following:

$$P(Y_i = [q_1, q_2]) = P(U_{i,aq_1} \geq \max_{p \in J} U_{i,ap} \cap U_{i,rq_2} \geq \max_{p \in J} U_{i,rp}) \quad (3.36)$$

The joint choice probability of arrival time q_1 and return time q_2 is equivalent to the probability that the latent utility associated with arrival time alternative q_1 is the maximum utility across all arrival time alternatives and that the latent utility associated with return time alternative q_2 is the maximum utility across all return time alternatives. Bayesian estimation (for both covariance matrix specifications) proceeds via a three-step Gibbs sampler as follows:

Step 1: Draw $U_i \mid V_i, \Sigma_i, X_i, Y_i \forall i$

Step 2: Draw $\lambda_a, \lambda_r, \lambda_d, \mu_{1,\text{full}}, \mu_{2,\text{full}}, \mu_{1,\text{part}}, \mu_{2,\text{part}} \mid U_i, V_i, X_i, Y_i \forall i$

Step 3: Draw $\beta_a, \beta_r, \eta \mid \Sigma_{\text{full}}, \Sigma_{\text{part}}, U_i, X_i, Y_i \forall i$

Here, Σ_i is taken to be either Σ_{full} or Σ_{part} depending on whether individual i is a full-time worker with no additional travel tours (Σ_{full}) or a part-time worker or an individual with additional tours (Σ_{part}). In addition, the Gibbs sampler does not generate draws for σ_a^2 or σ_r^2 here. It is well known that the MNP requires one element of Σ to be fixed for identification purposes (see, e.g., McCulloch and Rossi 1994). However, with the BVMNP, one element of Σ must be fixed for each nominal measure (Zhang et al. 2008). Thus, σ_a^2 or σ_r^2 are fixed at 1 for identification purposes.

In step 1, drawing from the conditional distribution of U_i is no trivial task. While U_i is distributed multivariate normal, when one conditions on Y_i , the elements of U_i that must represent the maximum utilities are known. Thus, conditional on Y_i , (the actual choice), the distribution of U_i is truncated multivariate normal (i.e., $U_i \sim \text{TMVN}(V_i, \Sigma_i)$). Drawing from a multivariate truncated normal distribution is no simple task. However, if instead one draws each element of U_i individually (conditional on all other elements), it can be shown that each is from a truncated univariate normal distribution. Typically, when MNP models are estimated in a Bayesian setting, each alternative-specific utility value is drawn sequentially, conditional on all other alternatives' utility values (see, e.g., Albert

and Chib 1993, McCulloch and Rossi 1994, and Zhang et al. 2008). However, this method is not used here because it was found to be computationally unstable, probably due to the large number of alternatives. Instead, a normal random walk MH step is used here to draw an individual's utility values simultaneously.

The proposal density for the MH step is a multivariate normal, with mean equal to the current utility values, and covariance given by $\kappa \Sigma_i$. Here, Σ_i is the utility covariance matrix for individual i , computed from the current values of covariance parameters, and κ is a deflation factor to increase the probability of proposal acceptance. The deflation factor was set to $\kappa = 0.05$, after calibrating the parameter to achieve approximately 25% proposal acceptance. The reason for the relatively low value of κ is because of the truncation of utilities. When utility values are allowed to shift by larger amounts, there is a greater likelihood that truncation restrictions will be violated, resulting in proposal rejection. It was found that this MH algorithm is much more computationally stable than the typical Gibbs sampling algorithm, and it also reduces computation time per iteration by nearly one half. Unfortunately, since utility values are more restricted in their movements from one iteration to the next, the algorithm is slow to converge.

In the second step of the Gibbs sampler, a draw of the covariance matrix parameters is generated. This procedure is the same for both covariance matrix specifications, so no distinction is made here. Priors on λ_a , λ_r , and λ_d (for both specifications) are specified to be independent uniform distributions over the interval from 0 to 1, reflecting a belief that there should be positive correlation across alternative utilities. Priors on $\mu_{1,\text{full}}$ and $\mu_{1,\text{part}}$ are specified to be independent normal distributions with means of 9 and 6 (hours), respectively, and variances of 2 each, while priors on $\mu_{2,\text{full}}$ and $\mu_{2,\text{part}}$ are specified to be normal distributions, each with means and variances of 0 and 1, respectively. Thus, the

full conditional posterior distribution of the variance parameters can be written as follows²⁷:

$$p(\lambda_a, \lambda_r, \lambda_d, \mu_{\text{full}}, \mu_{\text{part}} \mid U_i, V_i \forall i) \propto |\Sigma_{\text{full}}|^{-n_{\text{full}}/2} |\Sigma_{\text{part}}|^{-n_{\text{part}}/2} \exp\left(-\frac{1}{2}\left(\sum_{q=1}^4 \frac{(\mu_q - \bar{\mu}_q)^2}{\sigma_q^2} + \sum_i (U_i - V_i)' \Sigma_i^{-1} (U_i - V_i)\right)\right) \quad (3.37)$$

where,

$$\sum_{q=1}^4 \frac{(\mu_q - \bar{\mu}_q)^2}{\sigma_q^2} = \frac{(\mu_{1,\text{full}} - 9)^2}{4} + \frac{(\mu_{1,\text{part}} - 6)^2}{4} + (\mu_{2,\text{full}})^2 + (\mu_{2,\text{part}})^2$$

Since the density here is not in any standard form (with respect to the parameters), a MH step is used to draw these parameters. The proposal density is assumed to be normal, with mean given by the current draw of the parameters (i.e., a normal random walk) and variance initially taken to be very small. Like the MH steps for the CCNL and continuous logit, the covariance matrix for the proposal density is updated during the estimation process to aid in generating good proposals. While one would not reasonably expect λ_a , λ_r , and λ_d to be from a normal distribution, the normal proposal offers ease in generating draws and capturing correlations across parameters in the proposal. Furthermore, there are certain restrictions on these parameters to ensure Σ_{full} and Σ_{part} are positive definite matrices. To enforce these restrictions, if a draw of λ_a , λ_r , λ_d , $\mu_{1,\text{full}}$, $\mu_{1,\text{part}}$, $\mu_{2,\text{full}}$, and $\mu_{2,\text{part}}$ generate Σ_{full} or Σ_{part} that are not positive definite, a new draw is generated²⁸.

In the last step of the Gibbs sampler, a draw of β_a , β_r , and η is generated from the full conditional posterior distribution. For notational convenience, write $\beta = [\beta_a', \beta_r', \eta']'$. Here, the prior for these parameters is chosen to be multivariate normal with mean $\bar{\beta}$ and

²⁷ When conditioned on U_i and V_i , λ_a , λ_r , λ_d , μ_{full} , and μ_{part} are independent of X_i and Y_i .

²⁸ This is not considered a rejected draw for the MH step. The proposal for the MH step is the first draw from the proposal that results in Σ_{full} and Σ_{part} being positive definite.

covariance matrix Σ_β . Thus, the full conditional posterior is proportional to the following²⁹:

$$p(\beta \mid \Sigma_{\text{full}}, \Sigma_{\text{part}}, U_i, X_i \forall i) \propto \exp\left(-\frac{1}{2}\left((\beta - \bar{\beta})' \Sigma_\beta^{-1} (\beta - \bar{\beta}) + \sum_i (U_i - X_i \beta)' \Sigma_i^{-1} (U_i - X_i \beta)\right)\right) \quad (3.38)$$

Suppose Ω and Λ are given by the following:

$$\Lambda = \Sigma_\beta^{-1} \bar{\beta} + \sum_i X_i' \Sigma_i^{-1} U_i \quad (3.39)$$

$$\Omega = \left(\Sigma_\beta^{-1} + \sum_i X_i' \Sigma_i^{-1} X_i\right)^{-1} \quad (3.40)$$

Expression 3.38 suggests that β is proportional to a multivariate normal distribution with mean given by $\Omega\Lambda$ and covariance matrix Ω . Thus, β is drawn from a multivariate normal distribution. Here, vague prior parameters are specified, with $\bar{\beta}$ taken to be a vector of zeroes, off-diagonal elements of Σ_β taken to be zeroes, and diagonal elements of Σ_β set to be very large. That is all that is needed to generate the MCMC draws for this BVMNP model.

3.4 Chapter Summary

This chapter detailed three models for time-of-day choice, including the Bayesian methods needed for their parameters' estimation. Each of the three models is based in random utility theory, a key objective of this dissertation's model development. However, none of the three meets each of the other four goals here. Nonetheless, each offers something new relative to existing models. No existing continuous TOD models are derived from random utility theory, like the continuous logit and CCNL. In addition, the CCNL allows for correlations across alternatives, something not yet realized in any

²⁹ Note that conditional on $\Sigma_{\text{full}}, \Sigma_{\text{part}}, U_i$, and X_i , β is independent of Y_i .

continuous random utility model. While the BVMNP model may not represent time in a continuous context, it does allow correlation across alternatives in a two-dimensional tour TOD choice setting. No existing two-dimensional tour TOD choice models allows for such correlations. Each of the methods described in this chapter offer promise for advancing the state-of-the-art of TOD modeling. In Chapter 4, methods for imputing continuously varying network attributes are discussed.

CHAPTER 4: IMPUTING TIME-VARYING NETWORK VARIABLES

In practice, many advanced tour- and activity-based travel demand model systems use only a handful of broad time-of-day periods to generate network-level variables in model application (see, e.g., San Francisco's SFCTA model [Jonnalagadda et al. 2001], New York's NYBPM [PB Consult 2005a], Columbus, Ohio's MORPC model [PB Consult 2005b], and Sacramento's SACOG model [Bowman et al. 2006], among others). However, many of these model systems treat time-of-day (TOD) choice at a much finer resolution (like the 1-hour intervals of the MORPC model [PB Consult 2005b] and the 30-minute intervals of the SACOG model [Bowman et al. 2006]). Network attributes can only play a limited role in such TOD models since they only vary across particular subsets of the alternative choice set. Better results could be obtained if network attributes varied over all alternatives, either by directly modeling these over time or by assuming some distribution for them. This dissertation seeks to employ continuously varying (over time) network attributes (at least by automobile, for which many data points exist). This chapter details the methods, data, and empirical findings of this work.

4.1 Methodology

The data used for this dissertation's empirical work does not contain network attributes at the desired temporal resolution. In fact, the only network attributes contained in the 2000 San Francisco Bay Area data are travel time, distance, and tolls for five broad times of day (and no reliability information is available). However, one key element of incorporating network attributes as explanatory variables in this dissertation's activity timing models lies in deriving reasonable measures from available data. This section details the methodology used for imputing such measures on a continuous scale from individual trip-making reports.

4.1.1 Automobile Travel Times

The literature offers several instances where ordinary least squares (OLS) regression models have been employed for travel time prediction of automobile trips on a continuous scale (see, e.g., Cambridge Systematics 2005, Abou Zeid et al. 2006, Popuri et al. 2008, and Komma and Srinivasan 2008). This method assumes that reported travel times are correct, on average.³⁰ Komma and Srinivasan (2008) created a dependent variable equal to the ratio of reported travel times to free-flow travel times, while Abou Zeid et al. (2006) used the ratio of reported to free-flow speeds as the dependent variable. To ensure non-negativity, Popuri et al. (2008) modeled the natural logarithm of the ratio of reported to free-flow automobile trip speeds. In each case, a number of origin- and destination-specific variables were used as explanatory variables (such as distance, origin and destination area types, and peak travel delay). A similar approach is employed here for automobile trips.

Like Popuri et al.'s (2008) models, the dependent variable here is chosen as the natural logarithm of the ratio of reported to free-flow trip speeds. The model is formulated as follows:

$$\ln\left(\frac{[\text{Reported Speed}]_{iq}}{[\text{FF Speed}]_q}\right) = X_{iq}'v_1 + (t_1 + \sum_{j=1}^4[\sum_{l=1}^{L_1} \omega_{1jl}(\psi_j(t))^l]) * (\text{Delay})_q + \epsilon_i \quad (4.1)$$

Here, X_{iq} is a vector of covariates related to trip i and origin-destination (OD) pair q , which includes a constant, the natural logarithm of trip distance, area type indicator variables for origin and destination zones (with rural area type as the base), day-of-week indicator variables (with Monday as the base), and mode indicator variables³¹ (with drive

³⁰ It is well known that reported travel times are often not that accurate due to misreporting and rounding error (Stopher et al. 2008). In fact, the 1996 North Central Texas Council of Governments Household Travel Survey suggests that only 13% of survey respondents report travel times to the nearest minute, while over 50% report times to the nearest 30 minutes or more (Stopher et al. 2008).

³¹ Note that these models only consider automobile trips, so the mode indicator indicates vehicle occupancy (1, 2, and 3 or more occupants).

alone as the base). v_1 , t_1 , and ω_1 are parameters to be estimated and ϵ_{iq} is the random error term, assumed to be independent and identically distributed across observational units (reported auto trips). The $(\text{Delay})_q$ variable is very important since it describes the typical peak level of congestion versus the analyst's estimate of free-flow conditions. For instance, if a particular origin-destination pair sees little or no congestion across the day, the $(\text{Delay})_q$ variable takes a value close to zero, and predicted speeds will exhibit very little variation across different TODs. Similar to Popuri et al. (2008), the delay variable is defined as follows:

$$(\text{Delay})_q = 1 - \frac{(\text{peak speed})_q}{(\text{FF speed})_q} \quad (4.2)$$

Here, peak speed represents the analyst's estimate of speed when congestion is at its worst for OD pair q . The main reason for this variable (rather than simply a constant) is to account for differences in speed profiles across different OD pairs. For instance, network links connecting some OD pairs may enjoy little or no delay across the entire day, while links connecting other OD pairs may exhibit extreme congestion during peak periods. The $(\text{Delay})_q$ variable accounts for these differences in a systematic way. As shown in equation 4.1, the coefficient on this delay variable is made up of a number of terms, and describes how speeds vary over time t . The functions $\psi_1(t)$, $\psi_2(t)$, $\psi_3(t)$, and $\psi_4(t)$ are the following non-negative cyclical functions (similar to Popuri et al.'s [2008] formulation):

$$\begin{aligned} \psi_1(t) &= \exp \left[\sin \left(\frac{2\pi t}{24} \right) \right], \psi_2(t) = \exp \left[\cos \left(\frac{2\pi t}{24} \right) \right], \\ \psi_3(t) &= \exp \left[\sin \left(\frac{4\pi t}{24} \right) \right], \psi_4(t) = \exp \left[\cos \left(\frac{4\pi t}{24} \right) \right] \end{aligned} \quad (4.3)$$

Clearly, this specification ensures smooth speed prediction profiles across time, which is a desirable property. In addition, there are a total of L_1 sets of these cyclical functions in

the regression equation (as shown in equation 4.1), where $\psi_1(t)$, $\psi_2(t)$, $\psi_3(t)$, and $\psi_4(t)$ are taken to powers 1, 2, ..., L_1 with a parameter related to each. The value of L_1 in equation 4.1 is determined through empirical investigation by comparing the reasonableness of the implied speed profiles for different L_1 values (and the statistical significance of associated parameters).

After estimates of equation 4.1's parameters are obtained, predictions for the dependent variable can be made and corresponding travel times computed. These travel time predictions can be best viewed as average (or expected) travel times for a particular OD pair at a particular TOD (rather than predictions of actual travel times). This notion leads to formulation of travel time variability measures developed in the following section.

4.1.2 Automobile Travel Time Variability

A natural measure of travel time variability/unreliability is the variance or standard deviation of travel times for a given OD pair at a given TOD (i.e., distributional measures). The methods of Section 4.1.1 offer travel time predictions, which represent predictions of average travel time. Assuming reported travel times, on average, are equal to actual travel times (as before), one can reasonably argue that travel time residuals³² provide a measure of reliability. As discussed in Chapter 2, schedule delay methods may really be preferred in a TOD model context³³, but without appropriate data to measure schedule delay (i.e., travelers' preferred arrival times), the distributional approach seems reasonable.

Here, the dependent variable is computed as the natural logarithm of squared residuals (or the difference between reported and predicted travel times³⁴). The set of explanatory

³² Note that travel time residuals are not direct residuals from the model, but rather residuals computed as the difference between reported and predicted travel times. Model residuals would be in the form of the natural logarithm of reported to free-flow speeds.

³³ For instance, schedule delay measures reflect the real source of disutility associated with reliability.

³⁴ Predicted travel times come from the speed regression model of Section 4.1.1.

variables for automobile travel time variability (i.e., unreliability) are very similar to those used for the speed regression, detailed in Section 4.1.1. Moreover, the model can be estimated using OLS techniques. The model is formulated as follows:

$$\ln \left([TT_{rep} - TT_{pred}]_{iq}^2 \right) = X_{iq}' v_2 + (\iota_2 + \sum_{j=1}^4 [\sum_{l=1}^{L_2} \omega_{2jl} (\psi_j(t))^l]) * (\text{Delay})_q + \epsilon_i \quad (4.4)$$

Here, X_{iq} is an identical set of covariates used in equation 4.1, $(\text{Delay})_q$ is the same as in equation 4.2, $\psi_j(t)$ is the same as in equation 4.3, v_2 , ι_2 , and ω_2 are parameters to be estimated, and ϵ_{iq} is an independent and identically distributed (across observational units) error term. As in the speed regression (equation 4.1), L_2 is determined by comparing the reasonableness of the implied TOD effect on travel time variance.

Since equation 4.4's dependent variables (i.e., the travel time residuals) rely on reported travel times, they not only capture variations in travel times, but also respondent mistakes/errors in reporting (due to misreporting and rounding [see, e.g., Stopher et al. 2008]). Thus, model estimates will likely be biased a bit high. However, a great deal of uncertainty is controlled for via explanatory variables specific to the OD pair (including travel distance, origin and destination characteristics, vehicle occupancy, and day-of-week), which many travelers may not perceive, thereby lowering these estimates.

One key attribute of the models in equations 4.1 and 4.4 is that travel time and variance profiles of different trips will share the same underlying shape (as predicted by the delay coefficient). While relative heights of these profiles' peaks and valleys will differ (depending on the value of $(\text{Delay})_q$), the location and relative shape will be the same. For instance, under the variance regression, all trips will share the feature that the maximum variance occurs at some specific time (presumably during either the AM or PM peak). Of course, the models' search for a set of single parameters neglects the fact that different trips (with different free-flow speeds, travel distances, OD attributes, etc.)

may have wholly different travel time and variance profiles – due to specific link characteristics and demand profiles over times. Nonetheless, without actual network data, quantifying such differences in a meaningful way is difficult (e.g., inbound versus outbound to the central business district [CBD]). Such link-specific delay and travel time information, by TOD, may emerge through better instrumentation of highways, detailed dynamic traffic assignment estimates, controlled route-choice experiments, and the like. Alternatively, one could investigate variable interaction effects with the cyclical TOD functions.³⁵ However, investigating such effects could be quite time-consuming, depending on how many interactions the analyst chooses to examine, though new emerging methods exist for just this purpose. Bayesian Additive Regression Trees (BART) are one such method that have been used in just such an application (see, Chipman et al. 2009).

4.1.3 Transit Level-of-Service Attributes

While the previous two sections examined trips made by the automobile mode, this section details methods used for transit trips. Unlike the automobile models, where network attributes are modeled continuously over time, transit attributes are assumed constant over broad TOD periods. The reason for taking a different approach here is two-fold. First, the data used here (and described in Section 4.2) do not contain information that could be regarded as free-flow transit times. So, one cannot formulate equation 4.1 as a model for transit trips. Second, the dataset’s observational units do not offer information on the form of transit used (e.g., train, express bus, local bus), though (at least) four different transit modes are available to San Francisco Bay Area travelers. Presumably each form of transit has its own unique characteristics, which could potentially lead to bias in estimation.

³⁵ For instance, one may interact origin- or destination-specific variables with those TOD functions. This would allow for differential TOD predictions across OD pairs.

Since the data contain transit level-of-service (LOS) attributes for each of five TOD periods, those are used here for travel times, distances, and fares. However, these attributes are available for each of four transit modes. In the absence of information on the actual transit mode chosen from these four, the mode offering the lowest travel time was assumed to be the mode chosen for those travel units made by the transit mode. In addition, and like the automobile network attributes, reliability information is not present. Thus, transit travel time variance regression models are needed.

Two models are used here, one for drive-to-transit mode and one for walk-to-transit mode. Like the automobile variance regression, the response variable is chosen to be the squared difference between reported and skimmed travel times³⁶. Analysis of the San Francisco Bay Area transit trip data suggested that the drive-to-transit mode would be best modeled as in equation 4.5, while the walk-to-transit mode would be best modeled as shown in equation 4.6. Essentially, the response variable of travel time variance is normalized by the average *motorized* travel time. The main reason for the difference between the two models is that the majority of travel time variability for the walk-to-transit mode should be in-vehicle, since access and egress times should be fairly stable by the walk mode.

$$\ln \left(\frac{(TT_{rep} - TTT_{skim})_{iq}^2}{(TTT_{skim})_q^2} \right) = X_{iq}' v_3 + \iota_3 (TTT_{skim})_q + \epsilon_i \quad (4.5)$$

$$\ln \left(\frac{(TT_{rep} - TTT_{skim})_{iq}^2}{(IVTT_{skim})_q^2} \right) = X_{iq}' v_4 + \iota_4 (IVTT_{skim})_q + \epsilon_i \quad (4.6)$$

Here, TTT_{skim} refers to the network-based estimate (or average) of a trip's total travel time (including access and egress times), and $IVTT_{skim}$ refers to the network-based estimate (or average) of in-vehicle travel time (excluding access and egress times). As

³⁶ Here, skimmed travel times refer to network-based estimates of travel times. These are akin to the predicted travel times from speed regressions. However, instead of coming from a separate regression model, they come directly from the data.

described earlier, one may expect very different behaviors for different transit options. The models shown in equations 4.5 and 4.6 cannot capture such variations, which is an unfortunate limitation of the data. In Section 4.2, a more detailed description of the data, its limitations, and needed assumptions are presented.

4.2 Data Description

The data used here (and in the time-of-day models presented later) come from the 2000 San Francisco Bay Area Travel Survey (BATS). The survey collected travel information for roughly 17,000 households over a 2-day period. The observational unit of the data is the travel tour³⁷ (over 100,000 recorded tours), with network attributes provided for each of five TOD periods³⁸ and seven modes³⁹. In addition, network attributes are available for both the outbound tour journey and the inbound tour journey. In other words, each tour is associated with an origin (or base location) and a primary destination⁴⁰. The outbound tour journey corresponds to travel from the origin to the primary destination, while the inbound tour journey corresponds to travel from the primary destination to the origin. Network attributes include travel time, travel distance, and tolls for automobile modes, and access, egress, wait, transfer, and in-vehicle times along with fares for transit modes. Chosen modes reported in the data reflect the main transport mode used over all trip segments within a tour. And, as stated earlier, transit tour modes are coded as walk-to-transit and drive-to-transit, not by the actual form of transit used. Thus, the transit

³⁷ Travel tours represent a collection of trips beginning and ending at the same base location. The base location can either be home or work. Work-based tours represent sub-tours within a home-based tour. For instance, if an individual travels to work, then heads to lunch at noon and returns to work, before finally returning home that evening, a separate work-based sub-tour would be coded in the data along with the higher level home-based tour.

³⁸ These include early morning (EM) from 3 am to 6 am, AM peak (AM) from 6 am to 9 am, midday (MD) from 9 am to 3:30 pm, PM peak (PM) from 3:30 pm to 6:30 pm, and evening (EV) from 6:30 pm to 3 am.

³⁹ These include automobile, local bus, municipal bus, premium bus, and the Bay Area Rapid Transit (BART) rail mode.

⁴⁰ Since multiple activities can be undertaken on any tour, a set of rules is established to define activity characteristics of each tour. To this end, a hierarchy of activities was established and used to assign one activity on each tour as the primary activity. In general, mandatory type activities exist at the top of this hierarchy (including work, school, and university type activities). Thus, if a tour contains a work, school, or university activity, that activity is specified as the primary activity. The primary destination is defined as the location where the primary activity is undertaken.

network attributes were condensed into single measures by selecting the network attributes with the lowest travel times if multiple transit types were available for the observational unit.

Since the regression models of this chapter require trip-level data, special considerations are needed. Fortunately, network attributes recorded in the dataset reflect trip-based travel. Specifically, network attributes correspond to travel characteristics for a trip between the tour origin (or base location) and primary destination (regardless of whether stops are made on the actual tour journey). Thus, one can reasonably consider any tour journey for which no stops are made to be a trip between the tour origin and primary destination. Reported travel times for such tour journeys should correspond to trip travel times. Therefore, regression models of this chapter only consider tour journeys where no additional stops were made between the origin and primary destination.⁴¹

Several other considerations are needed. First, free-flow speeds are not available in the dataset. Instead, speeds during the early morning (EM) TOD period were substituted, since one would likely experience near free-flow speeds in the early morning hours of the day. Second, peak-period speeds are not explicitly reported in the data. Instead, peak-period speed was taken as the lowest across the five TOD periods for each OD pair (typically from the PM peak period). Third, to alleviate issues with misreporting and rounding of travel times, automobile mode observations were removed if the ratio of reported to free-flow speed exceeded 2.0 or if the ratio of reported to peak speed was less than 0.25 (totaling 17% of records). For transit variance regressions, observations were removed if the ratio of reported to average travel times was less than 0.5 or greater than 2.0. These cutoffs may seem (and are) rather arbitrary. Of course, there is no fool-proof way of determining the accuracy of reported travel times. The chosen cutoffs are not likely to result in many accurate reported speeds being dropped, though many inaccurate

⁴¹ Note that if an individual makes a stop on the way to the primary destination, but returns home directly, only the outbound tour journey need be dropped from the analysis.

reported speeds likely remain. Finally, only tours with relevant network attribute data and tours made on weekdays were considered.

The above considerations resulted in sample sizes of 86,358, 3,297, and 4,981 for automobile, drive-to-transit, and walk-to-transit modes, respectively. Table 4.1, Table 4.2, and Table 4.3 present descriptive statistics for a number of relevant variables in the automobile, drive-to-transit, and walk-to-transit mode regressions, respectively.

Table 4.1: Descriptive Statistics for Variables Used in Automobile Mode Regressions

Variable	Mean	Standard Deviation	Minimum	Maximum
Dependent Variable, $\ln\left(\frac{[\text{Reported Speed}]_{iq}}{[\text{FF Speed}]_q}\right)$	-0.366	0.484	-2.744	0.693
Departure Time (hours after midnight)	13.45	4.857	0.033	26.75
Arrival Time (hours after midnight)	13.82	4.857	0.083	26.92
ln(Travel Distance)	1.630	1.068	-2.996	4.882
Indicator for Tuesday	0.239	0.426	0	1
Indicator for Wednesday	0.217	0.412	0	1
Indicator for Thursday	0.192	0.394	0	1
Indicator for Friday	0.169	0.375	0	1
Indicator for Regional Core Origin Zone	0.008	0.087	0	1
Indicator for CBD Origin Zone	0.029	0.169	0	1
Indicator for Urban Business Origin Zone	0.070	0.255	0	1
Indicator for Urban Origin Zone	0.178	0.383	0	1
Indicator for Suburban Origin Zone	0.671	0.470	0	1
Indicator for Regional Core Destination Zone	0.009	0.096	0	1
Indicator for CBD Destination Zone	0.033	0.178	0	1
Indicator for Urban Business Destination Zone	0.075	0.264	0	1
Indicator for Urban Destination Zone	0.184	0.388	0	1
Indicator for Suburban Destination Zone	0.656	0.475	0	1
Indicator for Shared Ride Mode w/2 Occupants	0.246	0.430	0	1
Indicator for Shared Ride Mode w/3 or more Occupants	0.224	0.417	0	1
Delay Variable, $(\text{Delay})_q$	0.267	0.190	0	0.838
Observations	86,358			

Table 4.2: Descriptive Statistics for Variables Used in Drive-to-Transit Mode Regression

Variable	Mean	Standard Deviation	Minimum	Maximum
Dependent Variable, $\ln\left(\frac{(TT_{rep} - TTT_{skim})_{iq}^2}{(TTT_{skim})_q^2}\right)$	-2.473	2.579	-15.20	1.39
Indicator for Regional Core Origin Zone	0.271	0.444	0	1
Indicator for CBD Origin Zone	0.095	0.293	0	1
Indicator for Urban Business Origin Zone	0.079	0.270	0	1
Indicator for Urban Origin Zone	0.151	0.358	0	1
Indicator for Regional Core Destination Zone	0.279	0.449	0	1
Indicator for CBD Destination Zone	0.108	0.310	0	1
Indicator for Urban Business Destination Zone	0.073	0.260	0	1
Indicator for Urban Destination Zone	0.133	0.339	0	1
Indicator for AM Peak Period	0.445	0.497	0	1
Indicator for Midday Period	0.106	0.308	0	1
Indicator for PM Peak Period	0.355	0.478	0	1
Indicator for Evening Period	0.093	0.290	0	1
Total Travel Time	47.23	20.30	11.12	242.9
Observations	3,297			

Table 4.3: Descriptive Statistics for Variables Used in Walk-to-Transit Mode Regression

Variable	Mean	Standard Deviation	Minimum	Maximum
Dependent Variable, $\ln\left(\frac{(TT_{rep} - TTT_{skim})_{iq}^2}{(IVTT_{skim})_q^2}\right)$	-0.198	2.898	-20.94	8.48
Indicator for Regional Core Origin Zone	0.189	0.392	0	1
Indicator for CBD Origin Zone	0.179	0.384	0	1
Indicator for Urban Business Origin Zone	0.173	0.379	0	1
Indicator for Urban Origin Zone	0.260	0.439	0	1
Indicator for Regional Core Destination Zone	0.209	0.407	0	1
Indicator for CBD Destination Zone	0.190	0.393	0	1
Indicator for Urban Business Destination Zone	0.168	0.374	0	1
Indicator for Urban Destination Zone	0.243	0.429	0	1
Indicator for AM Peak Period	0.384	0.487	0	1
Indicator for Midday Period	0.223	0.416	0	1
Indicator for PM Peak Period	0.283	0.451	0	1
Indicator for Evening Period	0.104	0.305	0	1
In-Vehicle Travel Time	16.88	15.01	0.65	175.3
Observations	4,981			

Figure 4.1 shows the distribution of the dependent variables (excluding the automobile variance regression, which requires predictions from the automobile speed regression). As shown in Figure 4.1a, a large number of observations are positive, meaning reported speed was greater than free-flow speed. While this may seem strange, it is certainly possible, since every traveler will drive differently in free-flow conditions, with some exceeding assumed free-flow speeds. Of course, one would not expect many drivers to travel at twice free-flow speed. Some of these respondents have likely misreported their travel times, but others may be accurate as free-flow speeds only represent the analyst's estimates. Figure 4.1b and Figure 4.1c suggest a large portion of negative transit dependent variables. This makes sense, since the dependent variable is measured by the variance divided by squared in-vehicle or total travel time.

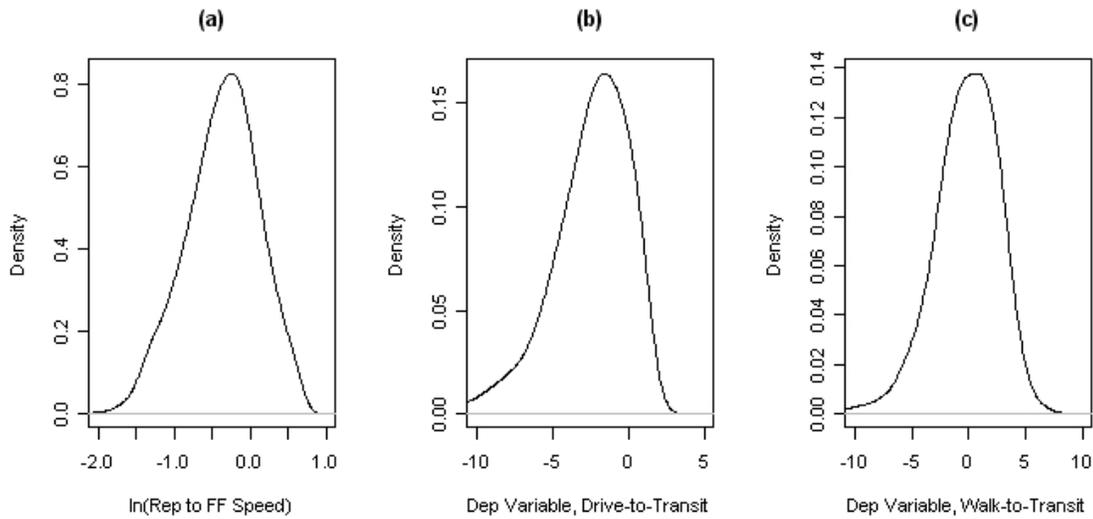


Figure 4.1: Distributions of Dependent Variables for (a) Automobile Speeds, (b) Drive-to-Transit Variances, and (c) Walk-to-Transit Variances

4.3 Empirical Results

The model estimates from equation 4.1 (i.e., automobile speed regressions) are shown in Table 4.4. Two models are shown here: one based on departure times and the other based on arrival times.⁴² While not all of the parameters are statistically significant, all have expected signs and similar effects for both models. Of the parameters not appearing in the delay coefficient, the most practically significant are indicators for regional core and central business district (CBD) zones (both origin and destination) and the natural logarithm of distance, not surprisingly, since longer distance trips will typically enjoy higher speeds.

⁴² The difference here is the time component entered into the cyclical functions of the model. The reason two separate models are desired here has to do with the activity timing models. For the BVMNP model, both outbound timing and return timing of travel (for a tour) are considered. Moreover, the outbound time is defined by the arrival time at the destination while the inbound time is defined by the departure time from the destination. In this way, the implied durations of stay at the primary destination are bounded below by 0.

Table 4.4: Automobile Mode Speed Regression Model Estimates for Departure Time- and Arrival Time-Based Models

Variable	Departure Time-Based Model		Arrival Time-Based Model	
	Coeff.	t-stat	Coeff.	t-stat
Constant	-0.3205	-30.08	-0.3195	-29.96
ln(Travel Distance)	0.2008	94.02	0.1987	93.16
Indicator for Tuesday	-0.0052	-1.10	-0.0055	-1.15
Indicator for Wednesday	-0.0143	-2.96	-0.0141	-2.89
Indicator for Thursday	-0.0154	-3.08	-0.0154	-3.08
Indicator for Friday	-0.0069	-1.33	-0.0054	-1.04
Indicator for Regional Core Origin Zone	-0.2015	-10.48	-0.1949	-10.13
Indicator for CBD Origin Zone	-0.1974	-16.66	-0.1875	-15.81
Indicator for Urban Business Origin Zone	-0.1608	-16.71	-0.1554	-16.14
Indicator for Urban Origin Zone	-0.1382	-16.32	-0.1346	-15.89
Indicator for Suburban Origin Zone	-0.0603	-7.75	-0.0611	-7.84
Indicator for Regional Core Destination Zone	-0.1998	-11.18	-0.2155	-12.06
Indicator for CBD Destination Zone	-0.2032	-17.63	-0.2083	-18.05
Indicator for Urban Business Destination Zone	-0.1529	-15.99	-0.1555	-16.26
Indicator for Urban Destination Zone	-0.1363	-16.01	-0.1381	-16.20
Indicator for Suburban Destination Zone	-0.0643	-8.16	-0.0633	-8.04
Indicator for Shared Ride Mode w/2 Occupants	-0.0361	-9.45	-0.0308	-8.04
Indicator for Shared Ride Mode w/3 or more Occupants	-0.0429	-10.71	-0.0392	-9.77
Delay	-3.5696	-7.46	-1.5538	-3.33
Delay* ψ_1	-0.2308	-0.30	5.0125	6.44
Delay* ψ_2	6.0955	6.64	-1.7428	-2.02
Delay* ψ_3	-1.9896	-7.50	-0.4177	-1.63
Delay* ψ_4	2.0779	5.17	-0.0788	-0.20
Delay*(ψ_1^2)	0.0906	0.18	-3.5170	-6.79
Delay*(ψ_2^2)	-3.5927	-5.73	1.4730	2.50
Delay*(ψ_3^2)	1.2948	7.25	0.4008	2.30
Delay*(ψ_4^2)	-0.8263	-3.75	-0.2123	-1.00
Delay*(ψ_1^3)	0.0015	0.02	0.7151	7.35
Delay*(ψ_2^3)	0.5920	4.90	-0.3310	-2.92
Delay*(ψ_3^3)	-0.2434	-6.95	-0.0932	-2.70
Delay*(ψ_4^3)	0.1438	3.51	0.0644	1.64
Observations	86,358		86,358	
R-Squared	0.137		0.136	

The shape of the delay coefficients, as they vary over departure and arrival time choices, appears reasonable, as shown in Figure 4.2. Also shown in Figure 4.2, speeds are predicted to be lowest (and travel times highest) during typical AM peak times (approximately 6 to 8 am) and PM peak times (approximately 3 to 6 pm), and the arrival time-based model predictions appear to mimic those of the departure time-based model with a rightward shift. While the slowest speeds are expected during these peak TODs, it is worth noting that the location of speed profile peaks and valleys are assumed to be the same across all OD pairs (i.e., the slowest speeds will occur around 5 pm for all OD pairs).

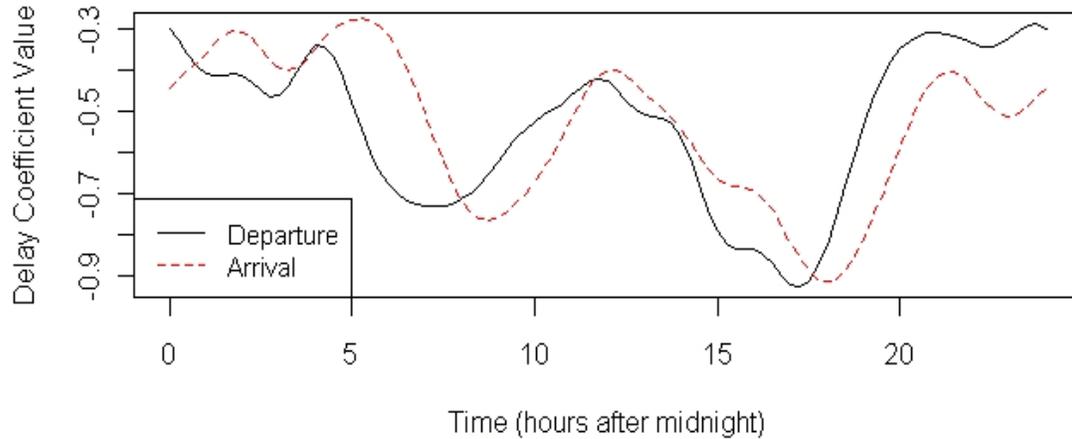


Figure 4.2: Delay Coefficient Variation in Automobile Mode Speed Regressions across Departure and Arrival Times

Figure 4.3 further details speed predictions from the model, by employing mean variable values for each time-invariant covariate. Profiles differ based on the value of the delay variable (i.e., the measure of typical peak congestion levels). Figure 4.3's y-axis represents coefficients by which *free-flow speed is multiplied* to obtain average speed estimates. When the delay variable is zero, the average speed prediction does not vary across time. As the delay variable increases from zero, peaks and valleys of the profiles become more pronounced, again with lowest speeds predicted near typical AM and PM

peak periods. As shown in Figure 4.3, speed coefficients have a maximum of about 0.8 (approximately a 20% speed reduction), even when no delay is present. This is an unfortunate consequence of the formulation here (free-flow speeds will not necessarily be realized at any TOD for a particular OD pair). It should be recognized, however, that this maximum coefficient will vary across OD pairs depending on origin and destination area types, travel distance, day-of-week, and vehicle occupancy. Figure 4.3 represents an average OD pair's attributes.

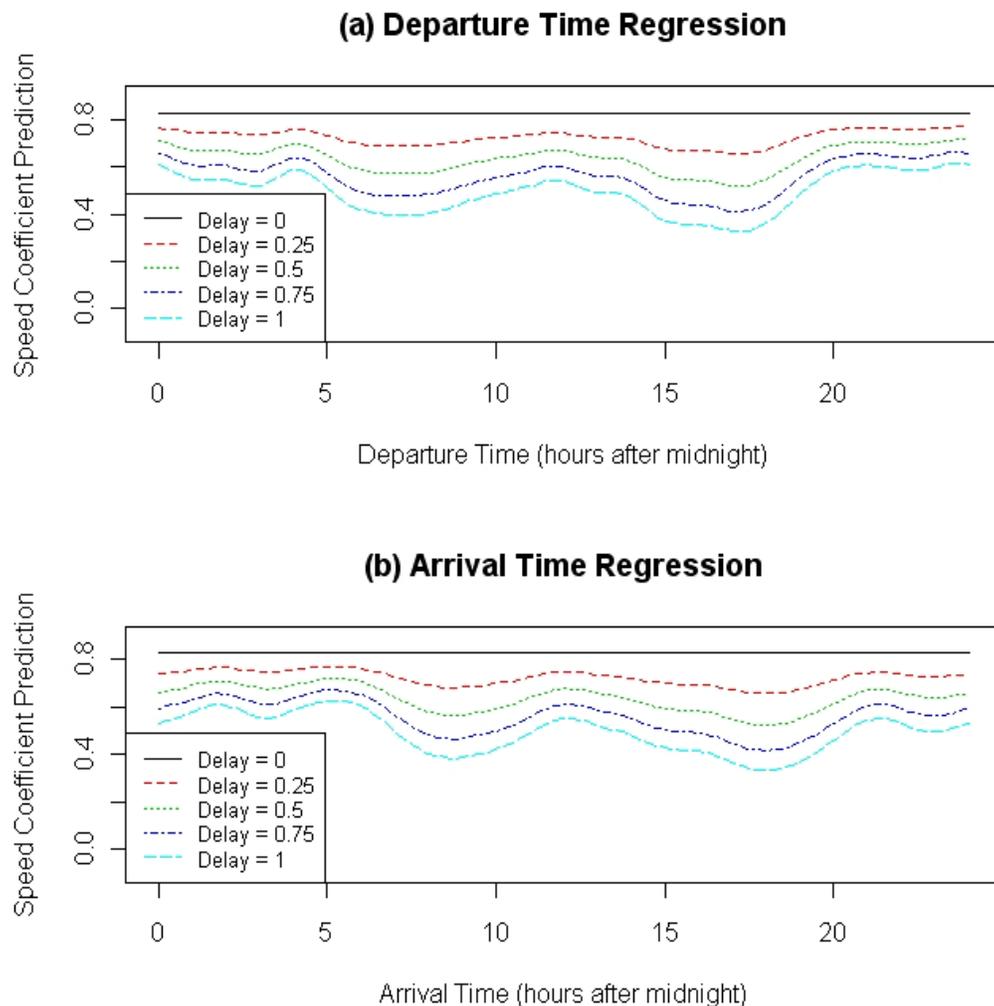


Figure 4.3: Automobile Mode Free-Flow Speed Coefficient Variation for Different Delay Variable Values

Parameter estimates of the model shown in equation 4.4 (i.e., the automobile travel time variance regressions) are shown in Table 4.5. Like the speed regressions, two models are estimated: a departure time-based model and an arrival time-based model. The most practically significant parameter is that associated with the natural logarithm of travel distance, which captures the effect of longer trips exhibiting higher travel time variance. As with the speed models, not all parameter estimates of the travel time variance models are statistically significant. In addition, several signs are not so intuitive. For instance, a trip originating in a regional core zone is predicted to exhibit greater travel time variability (less reliability) than one originating in a rural zone (all else equal), while trips originating in urban and suburban zones are predicted to enjoy lower travel time variance than those originating in rural zones (all else equal). Nonetheless, the results do not seem unreasonable, and more importantly, the fluctuations associated with the delay coefficient for these models (over TODs) appear as one would expect, as shown in Figure 4.4: travel time variances are highest during normal AM and PM peak periods. The differences between the departure time- and arrival time-based models shown in Figure 4.4 also appear reasonable, with the arrival time-based model predictions shifted to the right, relative to departure time.

Table 4.5: Automobile Mode Travel Time Variance Regression Model Estimates for
Departure Time- and Arrival Time-Based Models

Variable	Departure Time- Based Model		Arrival Time- Based Model	
	Coeff.	t-stat	Coeff.	t-stat
Constant	1.4239	25.64	1.4768	26.80
ln(Travel Distance)	0.7743	69.67	0.7774	70.65
Indicator for Tuesday	0.0036	0.15	0.0128	0.52
Indicator for Wednesday	-0.0175	-0.69	-0.0253	-1.01
Indicator for Thursday	0.0067	0.26	0.0091	0.35
Indicator for Friday	0.0161	0.60	-0.0032	-0.12
Indicator for Regional Core Origin Zone	0.5173	5.17	0.5809	5.85
Indicator for CBD Origin Zone	0.1732	2.81	0.1443	2.36
Indicator for Urban Business Origin Zone	0.0078	0.16	-0.0518	-1.04
Indicator for Urban Origin Zone	-0.0670	-1.52	-0.0976	-2.23
Indicator for Suburban Origin Zone	-0.0795	-1.96	-0.0823	-2.04
Indicator for Regional Core Destination Zone	0.4950	5.32	0.4823	5.22
Indicator for CBD Destination Zone	0.0619	1.03	-0.0015	-0.03
Indicator for Urban Business Destination Zone	-0.1760	-3.53	-0.2020	-4.09
Indicator for Urban Destination Zone	-0.1424	-3.21	-0.1948	-4.42
Indicator for Suburban Destination Zone	-0.1540	-3.75	-0.2087	-5.12
Indicator for Shared Ride Mode w/2 Occupants	0.1900	9.54	0.1828	9.25
Indicator for Shared Ride Mode w/3 or more Occupants	0.1617	7.75	0.1530	7.39
Delay	1.8612	13.97	3.0368	23.04
Delay* ψ_1	-0.1509	-4.57	-0.6832	-19.00
Delay* ψ_2	-0.7945	-14.17	-0.7361	-13.78
Delay* ψ_3	0.1838	4.53	-0.4294	-10.01
Delay* ψ_4	-0.3127	-8.73	-0.1838	-5.37
Observations	86,358		86,358	
R-Squared	0.147		0.151	

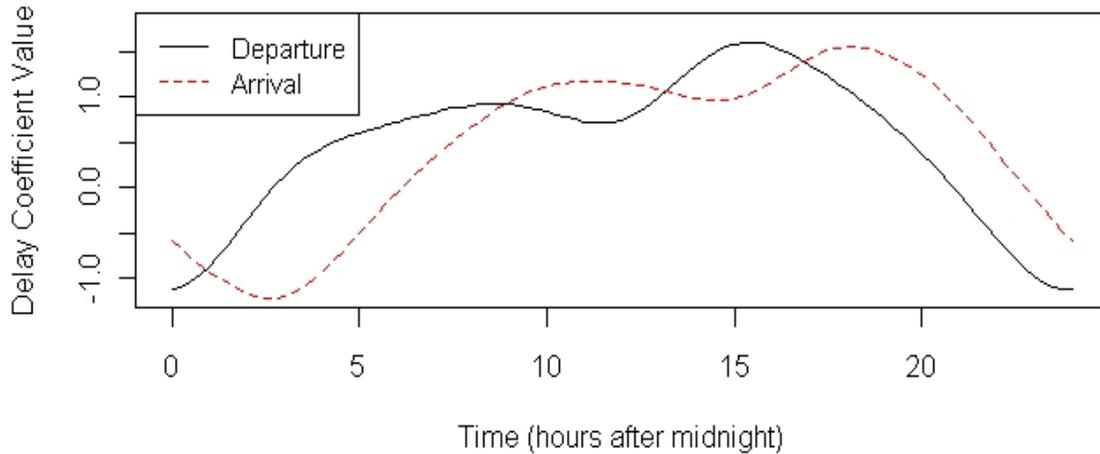


Figure 4.4: Delay Coefficient Variation in Automobile Mode Travel Time Variance Regressions across Departure and Arrival Times

Figure 4.5 shows travel time variance predictions for a trip with all time-invariant variables evaluated at mean covariate values, and a variety of delay coefficient values. Like the speed regressions, when the delay variable is zero, travel time variance does not vary over time. As the delay variable increases from zero, variations in travel time variance predictions become more pronounced. As with Figure 4.4, Figure 4.5's variance predictions are highest near the typical AM and PM peak periods.

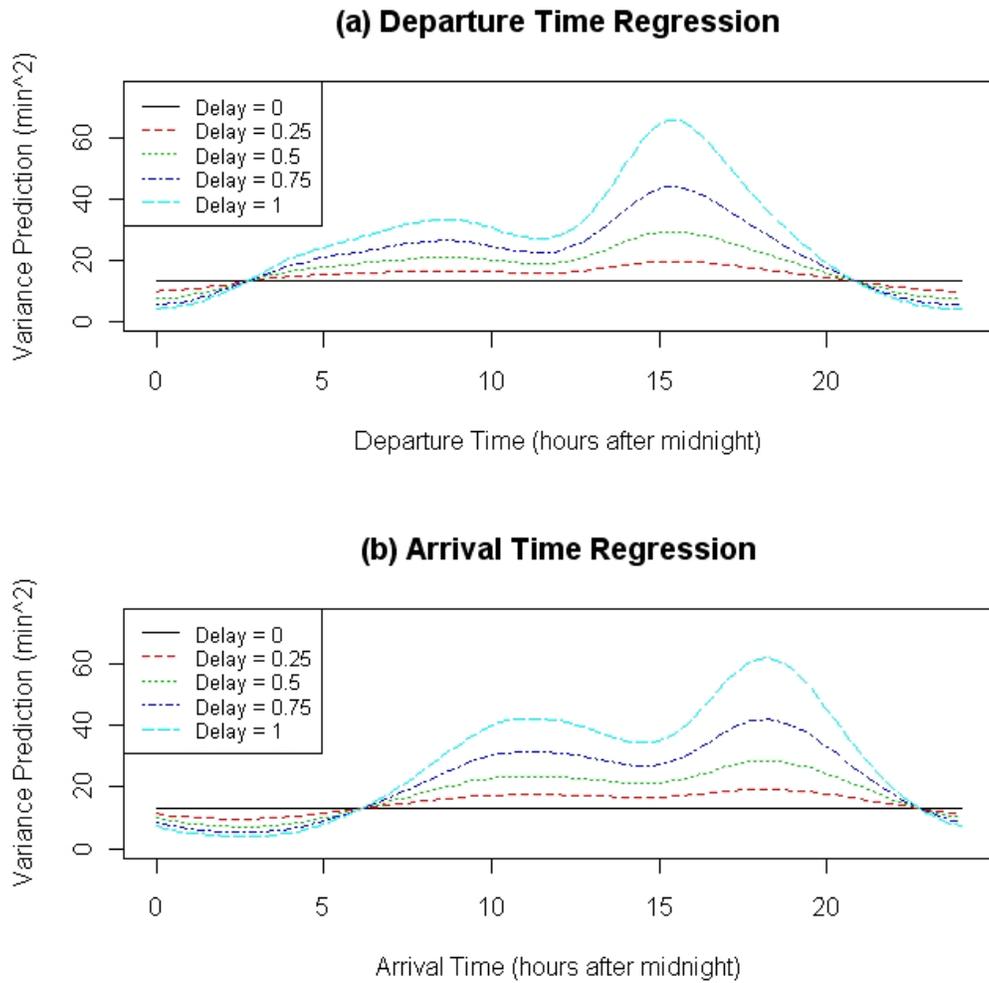


Figure 4.5: Automobile Mode Travel Time Variance Variation for Different Delay Variable Values and Mean Time-Invariant Covariates

As noted earlier, a key attribute of the automobile mode models presented in Section 4.1 is that travel time and variance profiles of different trips will share the same underlying shape (as predicted by the delay coefficients). Of course, different trips with different OD pairs may have very different profiles because of differing demand profiles and/or link characteristics. This is certainly a limitation of these models.

Model estimation results of the transit travel time variance regressions are shown in Table 4.6. It is clear that the smallest predicted travel time variances tend to occur during the early morning TOD period for drive-to-transit mode, and during early morning and evening periods for the walk-to-transit mode⁴³. For both models, travel time variances are predicted to be highest during the PM peak period, which seems very reasonable.

Table 4.6: Transit Mode Travel Time Variance Regression Model Estimates

Variable	Auto-Transit-Walk Model		Walk-Transit-Walk Model	
	Coeff.	t-stat	Coeff.	t-stat
Constant	-0.1656	-0.14	1.8097	10.71
Indicator for Regional Core Origin Zone	-0.8284	-5.56	-0.6344	-4.74
Indicator for CBD Origin Zone	-0.8248	-4.64	-0.7108	-5.56
Indicator for Urban Business Origin Zone	0.2899	1.68	-0.6192	-4.90
Indicator for Urban Origin Zone	0.2034	1.59	-0.5489	-4.92
Indicator for Regional Core Destination Zone	-0.7507	-4.99	-0.5245	-4.04
Indicator for CBD Destination Zone	-0.5704	-3.32	-0.4949	-3.91
Indicator for Urban Business Destination Zone	-0.0319	-0.17	-0.3951	-3.10
Indicator for Urban Destination Zone	0.1764	1.31	-0.1457	-1.27
Indicator for AM Peak Period	0.3246	0.28	0.3531	2.51
Indicator for Midday Period	0.5217	0.44	0.1685	1.20
Indicator for PM Peak Period	0.7685	0.65	0.3630	2.75
Indicator for Evening Period	0.5286	0.45	n/a	n/a
Total Travel Time (from skim)	-0.0493	-23.46	n/a	n/a
In-Vehicle Travel Time (from skim)	n/a	n/a	-0.0877	-35.23
Observations	3,297		4,981	
R-Squared	0.185		0.204	

⁴³ Note that an alternative specific constant for the evening TOD period was not included in the walk-to-transit mode model specification, since very few observations occurred during that period.

4.4 Chapter Summary

This chapter detailed the analytical methods and empirical results of the network variable imputation procedures. While the imputation techniques are not perfect, they do provide network variable measures on a continuous scale, which is an important input for the activity timing models developed in this dissertation. Existing methods were utilized to impute average travel times (by TOD), but travel time variance regressions represent a new technique for imputing travel time (un)reliability information using OLS regressions. Unfortunately, this technique also captures misreporting and rounding error (along with the actual variance), but considering the data limitations, the measures seem very reasonable.

In the following two chapters, the empirical results of the TOD models are presented. The empirical results of this chapter serve to inform the TOD models with network attribute variables.

CHAPTER 5: EMPIRICAL RESULTS OF CONTINUOUS TOD MODELS

This chapter details the empirical results of the continuous logit and continuous cross-nested logit (CCNL) model specifications. As discussed in Chapter 3, the models are estimated using Bayesian estimation techniques on work-tour outbound (i.e., from home to work) departure time choice data. For comparison purposes, the same data is used to estimate both the continuous logit and CCNL. The following section details the data used in this empirical work.

5.1 Work-Tour Departure Time Data

Like the data used for the speed and variance regressions of Chapter 4, the data here come from the 2000 Bay Area Travel Survey (BATS). Since observational units in the dataset represent travel tours (rather than trips), tour home-to-work departure time choice is analyzed here. Each tour record includes demographic variables (such as household and person characteristics), tour characteristics (such as origin and primary activity⁴⁴ destination attributes), tour mode chosen⁴⁵, and the timing of tour departure and return journeys (as well as the transportation level-of-service [LOS] data discussed earlier). The departure time analysis undertaken here uses only home-based tours made on weekdays and for which the primary activity was work⁴⁶. Further, if multiple home-based work tours existed for any individual in the data, only the first one was used in the sample⁴⁷. Thus, each record in the sample represents the first weekday work tour made over the 2-day survey period for all individuals. The sample was further restricted to observations where relevant transportation LOS data was available: tours with origins

⁴⁴ Primary activities were coded in the data using an activity hierarchy, with work-related activities being highest in this hierarchy for any tours involving multiple activities.

⁴⁵ Tour mode reflects the main transport mode used over all trip segments within a tour.

⁴⁶ Work activities were selected thanks to relatively predictable timing, and extensive existing literature on this type of trip-making. Of course, the methodology applies for any activity purpose.

⁴⁷ It is assumed here that the first work tour occurring in the survey period is scheduled first. To avoid correlations across tour records made by a single individual, only one work tour per individual is used.

and destinations outside the region did not have LOS data and so could not be used. After restricting attention to such tours, 17,820 tour records remained. However, because the CCNL model is computationally expensive, in terms of obtaining each draw, a random sample of 997 observations (about 6% of the 17,820 total) is used in the analyses here.

In this analysis, outbound work-tour departure time choice is considered. Alternatively, outbound arrival time choice could have been chosen as the dependent variable. Conditional on travel times, arrival times and departure times are essentially interchangeable if no intermediate stops are made between the home and work location. In addition, the inbound tour journey (i.e., from work to home) time-of-day (TOD) choice could have been considered. While the choice of outbound departure time choice is somewhat arbitrary, it does offer certain expectations in terms of the expected model behavior. For instance, one would expect individuals with longer travel distances to depart earlier and part-time workers to depart later. Of course, the techniques could apply equally well to other timing dimensions.

The explanatory variables used in this analysis include the (departure time-based) time-varying LOS attributes discussed in Chapter 4, as well as a time-varying cost variable. Of course, travel distances will not vary much across TODs, making it difficult to obtain time-varying costs. Here, travel costs vary (slightly) by TOD period (not continuously over time) and are computed as the travel distance multiplied by \$0.15/mile (the assumed operating cost of automobiles) plus any applicable tolls, and divided by the number of vehicle occupants for automobile modes. For transit modes, the transit fare (by TOD) is used, and for walk and bike modes, the cost is assumed to be zero.

In addition to time-varying LOS attributes, eight other individual/tour-specific variables plus a constant are interacted with the cyclical functions in the systematic utility equation. These variables include an indicator for males, age of the individual, an indicator for part-

time workers, an indicator for high income households (greater than \$75,000), household size, the number of tours undertaken by the individual over the entire day (excluding the modeled tour), travel distance to the primary destination, and a variable indicating whether the destination zone is coded as central business district (CBD). Table 5.1 offers descriptive statistics for these variables within the larger sample (of 17,820) of all work tours and within the smaller 6% sample (of 997) used in model estimation. As shown in Table 5.1, the differences between descriptive statistics for the two samples are very small. Figure 5.1 shows outbound departure time choice distributions for the full sample and the 6% sample. Again, there is little difference between the full and 6% samples, with departure times centered around 8 am, which is very reasonable.

Table 5.1: Descriptive Statistics of Explanatory Variables Used in Continuous TOD Models

Variable	Full Sample				6-percent Sample			
	Mean	St. Dev.	Min	Max	Mean	St. Dev.	Min	Max
Outbound Departure Time	8.03	2.15	3	23.2	8.08	2.22	3.25	18.5
Male Indicator	0.528	0.499	0	1	0.526	0.500	0	1
Age	42.7	11.8	18	97	42.8	11.8	18	80
Part-Time Worker Indicator	0.018	0.134	0	1	0.022	0.147	0	1
High-Income HH Indicator	0.535	0.499	0	1	0.547	0.498	0	1
Household Size	2.64	1.27	1	9	2.70	1.30	1	9
Number of Other Tours	0.345	0.624	0	5	0.337	0.622	0	4
Free-Flow Travel Distance	13.6	12.5	0.08	119.5	13.1	12.1	0.1	94.8
CBD Indicator	0.110	0.313	0	1	0.101	0.302	0	1

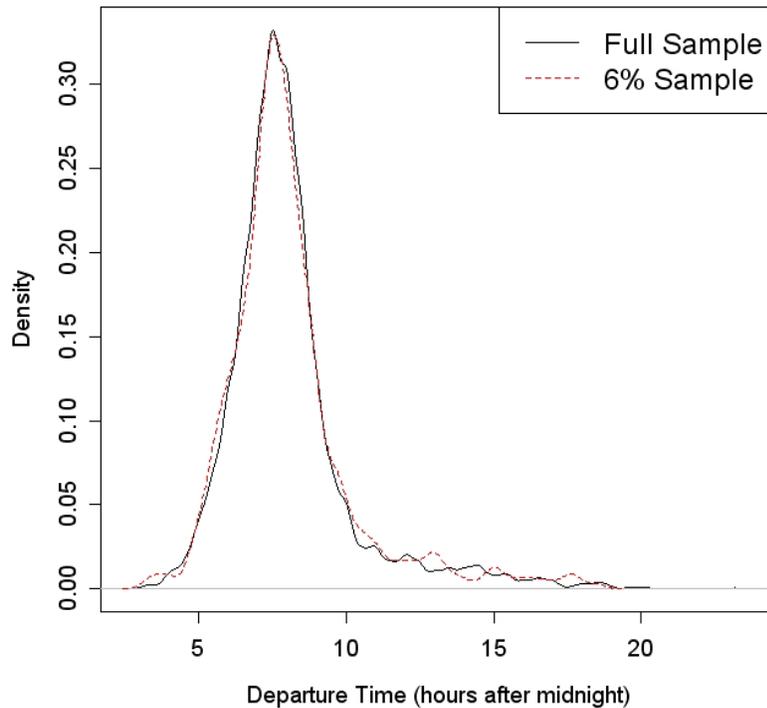


Figure 5.1: Estimated Density of Outbound Departure Time

5.2 Model Estimation Details

The Bayesian methods for model estimation were described in detail in Chapter 3. Here, some additional details are provided.

For the continuous logit, starting values of all parameters were set to zero. After approximately 200,000 draws from the posterior distribution of the model, it seemed fairly clear that draws had converged. Another 100,000 draws were obtained after convergence. As indicated in Chapter 3, only every 50th draw (from these 100,000) was used for inference to eliminate correlation between successive draws (which is inevitable using the MH algorithm here). Thus, the posterior distribution is characterized by 2,000 draws from that distribution. Using both informal graphical convergence diagnostics and Geweke's (1992) convergence diagnostic, convergence was checked and results were found to be acceptable.

Because obtaining draws from the CCNL's posterior is much more computationally burdensome (about 30 times slower), the CCNL Markov chain was initialized with mean parameter estimates obtained for the continuous logit model. Structural parameters, h and ρ , were initialized with values of 0.8 and 1.1, respectively. This allowed the CCNL's posterior draws to converge after only 100,000 draws. An additional 50,000 draws were obtained after convergence, again using only every 50th draw (from these 50,000) for inference (for a total of 1,000 posterior parameter draws). Like the continuous logit, convergence was checked using informal graphical techniques as well as Geweke's (1992) diagnostic, and results were found to be acceptable.

As the reader may note from Chapter 3, each time-invariant covariate (i.e., individual/tour-specific variables) is interacted with a collection of cyclical functions in the systematic utility equation. While no formal methods for variable selection were used here, initial results from the continuous logit estimation were examined to identify covariates with the most important effects. Those found to have the most substantial effects were then interacted with a larger collection of cyclical functions. This procedure was performed by estimating the continuous logit with each of the nine time-invariant covariates interacted with exactly eight cyclical functions. For those variables with less substantial effects in the model, either two or four of the interaction terms were removed, beginning with those cyclical interactions having the greatest variation over the 24-hour day. The following section describes the empirical results in detail.

5.3 Empirical Findings

Table 5.2 and Table 5.3 present the estimation results for the two models. Mean parameter estimates are generally consistent across the two models. The mean coefficient estimates on average travel time, travel time variance, and travel cost have negative signs for both models, by prior construction. Mean values of travel time (VOTTs) implied from the models' draws are \$7.34 and \$2.97 per hour for the

continuous logit and CCNL models, respectively. These values are rather low compared to expectations, since VOTTs are often estimated to be much greater (e.g., around \$8 to \$25 per hour [Brownstone and Small 2005]). Mean values of reliability (VOR) for the two models are almost identical, with estimates of \$0.028 and \$0.029 per squared minute (or \$9.95 and \$10.30 per hour of travel time's standard deviation).

The intervals are rather wide on all travel cost parameters, which is likely due to less variation in the travel cost variable across departure time alternatives, relative to average travel time and travel time variance. This is particularly true for the continuous logit estimates, and results in 90% VOTT intervals that range from \$0.03 to \$30.00 per hour (with a median of \$1.43 per hour) and VOR intervals from \$0.05 to \$22.13 per hour (with a median of \$2.64 per hour). The CCNL's parameter estimates suggest tighter 90% intervals, ranging from \$1.20 to \$6.01 per hour for VOTT (with a median of \$2.28 per hour) and from \$6.58 to \$14.04 per hour for VOR (with a median of \$9.81 per hour). Of course, people may not trade off time and money (or even reliability so much) in their departure time choices, due to ignorance or constraints on departures (like child drop off or fixed work times). So these VOTT and VOR estimates cannot really be compared to those coming from route or mode choice models. Moreover, these estimates are specific to the imputed network variables used here, which are imperfect. As discussed in Chapter 4, this dissertation's reliability measures not only capture travel time variances, but also misreporting and rounding errors in reported travel times, which could bias unreliability estimates high (or low potentially), resulting in lower or higher VOR estimates.

The mean estimates of the CCNL's correlation parameters appear quite reasonable, and statistically significant. For instance, h 's mean estimate is about 0.75, indicating that the minimum time between uncorrelated alternatives is 1.5 hours. While there is nothing special about this particular value, one certainly does not expect correlations to exist across alternatives very far apart (e.g., several hours in time). Of course, the mean

estimate of ρ (of 2.40) seems lower than expected. As noted in Chapter 3, the expression $(1 - \rho^{-2})$ defines the correlation between alternatives separated by an infinitesimally small distance, and one would expect near-perfect correlation between such alternatives. The correlation for such alternatives implied by the mean estimate of ρ is 0.83, not 1 (or something closer to one). However, given that ρ (along with h) defines correlations between other alternatives as well, the estimate seems reasonable. In order to truly achieve the near-perfect correlation between alternatives separated by a very small distance, a more flexible correlation structure would likely be needed.

Table 5.2: Continuous Logit and CCNL Model Estimation Results

Variable	Continuous Logit		CCNL	
	Mean Estimate	95% Interval	Mean Estimate	95% Interval
LOS Variables				
Average Travel Time	-0.00095	(-0.0178, 0)	-0.0042	(-0.0135, -0.0011)
Travel Time Variance	-0.00004	(-0.0063, 0)	-0.0025	(-0.0064, -0.0008)
Cost	-0.04754	(-0.3633, -0.0038)	-0.0954	(-0.1992, -0.0434)
Departure Time Functions				
Sin(2*pi*t/24)	3.6977	(1.7867, 6.0541)	2.1785	(1.2983, 3.0339)
Sin(4*pi*t/24)	2.4269	(0.2672, 5.0664)	0.6261	(0.2759, 0.937)
Sin(6*pi*t/24)	2.4645	(0.8885, 4.1858)	1.5397	(1.1058, 1.9953)
Sin(8*pi*t/24)	0.5528	(-0.2218, 1.3142)	0.4923	(0.1411, 0.8538)
Cos(2*pi*t/24)	-4.7262	(-8.0947, -1.9919)	-2.8292	(-3.391, -2.3379)
Cos(4*pi*t/24)	-0.5141	(-1.9731, 0.8283)	-0.3064	(-0.8347, 0.2194)
Cos(6*pi*t/24)	2.4113	(1.3118, 3.5476)	1.5987	(1.0785, 2.0579)
Cos(8*pi*t/24)	1.3434	(0.5838, 2.1487)	0.7941	(0.3889, 1.2113)
Male Indicator Interactions				
Sin(2*pi*t/24)	0.7372	(0.1533, 1.3411)	0.7580	(0.3246, 1.206)
Sin(4*pi*t/24)	0.6996	(0.2083, 1.1869)	0.7346	(0.4706, 1.0062)
Cos(2*pi*t/24)	-0.1548	(-1.2793, 0.9883)	-0.3178	(-0.7462, 0.1223)
Cos(4*pi*t/24)	0.2382	(-0.2816, 0.7504)	0.1479	(-0.2071, 0.4989)
Age Interactions				
Sin(2*pi*t/24)	0.0583	(-0.0073, 0.1269)	0.0735	(0.0342, 0.1178)
Sin(4*pi*t/24)	0.0600	(-0.0213, 0.1427)	0.0762	(0.0302, 0.1281)
Sin(6*pi*t/24)	0.0107	(-0.0437, 0.0656)	0.0193	(-0.0126, 0.0565)
Sin(8*pi*t/24)	0.0175	(-0.0035, 0.0387)	0.0138	(0.0011, 0.0279)
Cos(2*pi*t/24)	-0.2124	(-0.3635, -0.072)	-0.2027	(-0.3129, -0.1094)
Cos(4*pi*t/24)	-0.1242	(-0.2073, -0.0428)	-0.1037	(-0.1676, -0.0493)
Cos(6*pi*t/24)	-0.0666	(-0.1054, -0.0291)	-0.0433	(-0.0698, -0.0197)
Cos(8*pi*t/24)	-0.0267	(-0.046, -0.0081)	-0.0145	(-0.0262, -0.0033)
Part-Time Indicator Interactions				
Sin(2*pi*t/24)	-3.1017	(-4.8199, -1.4657)	-2.4492	(-3.032, -1.9058)
Sin(4*pi*t/24)	-2.3430	(-4.1796, -0.4361)	-1.5309	(-2.0617, -1.0047)
Sin(6*pi*t/24)	-0.7430	(-2.0028, 0.5482)	-0.4138	(-1.0535, 0.2374)
Cos(2*pi*t/24)	-0.3411	(-4.946, 3.4978)	-1.8135	(-2.3856, -1.324)
Cos(4*pi*t/24)	-0.8122	(-3.7006, 1.5505)	-1.5343	(-1.9832, -1.0824)
Cos(6*pi*t/24)	-0.6394	(-1.9494, 0.5628)	-0.7680	(-1.1835, -0.3127)
High Income HH Indicator Interactions				
Sin(2*pi*t/24)	-0.1619	(-0.6706, 0.3492)	-0.1105	(-0.4649, 0.2569)
Sin(4*pi*t/24)	-0.6585	(-1.0927, -0.2146)	-0.5793	(-0.8003, -0.3479)
Cos(2*pi*t/24)	0.3320	(-0.7165, 1.3286)	0.1772	(-0.1772, 0.5688)
Cos(4*pi*t/24)	0.0923	(-0.3726, 0.5582)	0.0300	(-0.2049, 0.2567)

Table 5.3: Continuous Logit and CCNL Model Estimation Results (Cont'd)

Variable	Continuous Logit		CCNL	
	Mean Estimate	95% Interval	Mean Estimate	95% Interval
HH Size Interactions				
Sin(2*pi*t/24)	-0.2642	(-0.4788, -0.0432)	-0.2304	(-0.3905, -0.0608)
Sin(4*pi*t/24)	-0.0519	(-0.2373, 0.1416)	-0.0158	(-0.1391, 0.1111)
Cos(2*pi*t/24)	0.4706	(0.039, 0.8958)	0.3622	(0.0602, 0.6185)
Cos(4*pi*t/24)	-0.0110	(-0.1986, 0.1645)	-0.0382	(-0.1763, 0.0965)
No. of Other Tours Interactions				
Sin(2*pi*t/24)	-0.8870	(-1.2404, -0.5448)	-0.8571	(-1.1497, -0.5676)
Sin(4*pi*t/24)	-0.4643	(-0.7912, -0.1508)	-0.3929	(-0.6774, -0.104)
Cos(2*pi*t/24)	0.7624	(-0.0748, 1.6177)	0.5722	(-0.0927, 1.1643)
Cos(4*pi*t/24)	0.2327	(-0.1372, 0.6069)	0.1470	(-0.1048, 0.3997)
Travel Distance Interactions				
Sin(2*pi*t/24)	0.0241	(-0.0126, 0.0683)	0.0293	(-0.0055, 0.072)
Sin(4*pi*t/24)	-0.0374	(-0.088, 0.0167)	-0.0292	(-0.079, 0.024)
Sin(6*pi*t/24)	-0.0053	(-0.0329, 0.0233)	-0.0065	(-0.0318, 0.0195)
Cos(2*pi*t/24)	0.0644	(-0.0359, 0.1499)	0.0652	(-0.0208, 0.1541)
Cos(4*pi*t/24)	-0.0035	(-0.0606, 0.0495)	0.0041	(-0.0451, 0.0554)
Cos(6*pi*t/24)	-0.0489	(-0.0695, -0.0288)	-0.0408	(-0.0645, -0.0179)
CBD Destination Zone Indicator Interactions				
Sin(2*pi*t/24)	-0.7472	(-1.6973, 0.1987)	-0.6528	(-1.1057, -0.2365)
Sin(4*pi*t/24)	-0.7457	(-1.4699, 0.0221)	-0.7339	(-1.0776, -0.3483)
Cos(2*pi*t/24)	1.6963	(0.3099, 3.0539)	1.4181	(0.8338, 2.0036)
Cos(4*pi*t/24)	0.6503	(0.0173, 1.3087)	0.5738	(0.3014, 0.8679)
CCNL Structural Parameters				
h	n/a	n/a	0.7504	(0.5, 1.4)
ρ	n/a	n/a	2.3958	(1.0675, 4.6205)
Observations				
	997		997	

To better understand behavioral differences suggested by the two models and the effects each time-invariant covariate has on departure time choice (since parameter estimates for these variables are difficult to interpret on their own), Figure 5.2 and Figure 5.3 show density profiles for example individuals. In each plot, covariate values are taken to be the average over all individuals within the sample for all but one of the covariates. For the final covariate, two representative values are chosen to demonstrate the covariate's effect

on departure time choice⁴⁸. Note here that level-of-service variable effects are identical across all plots.

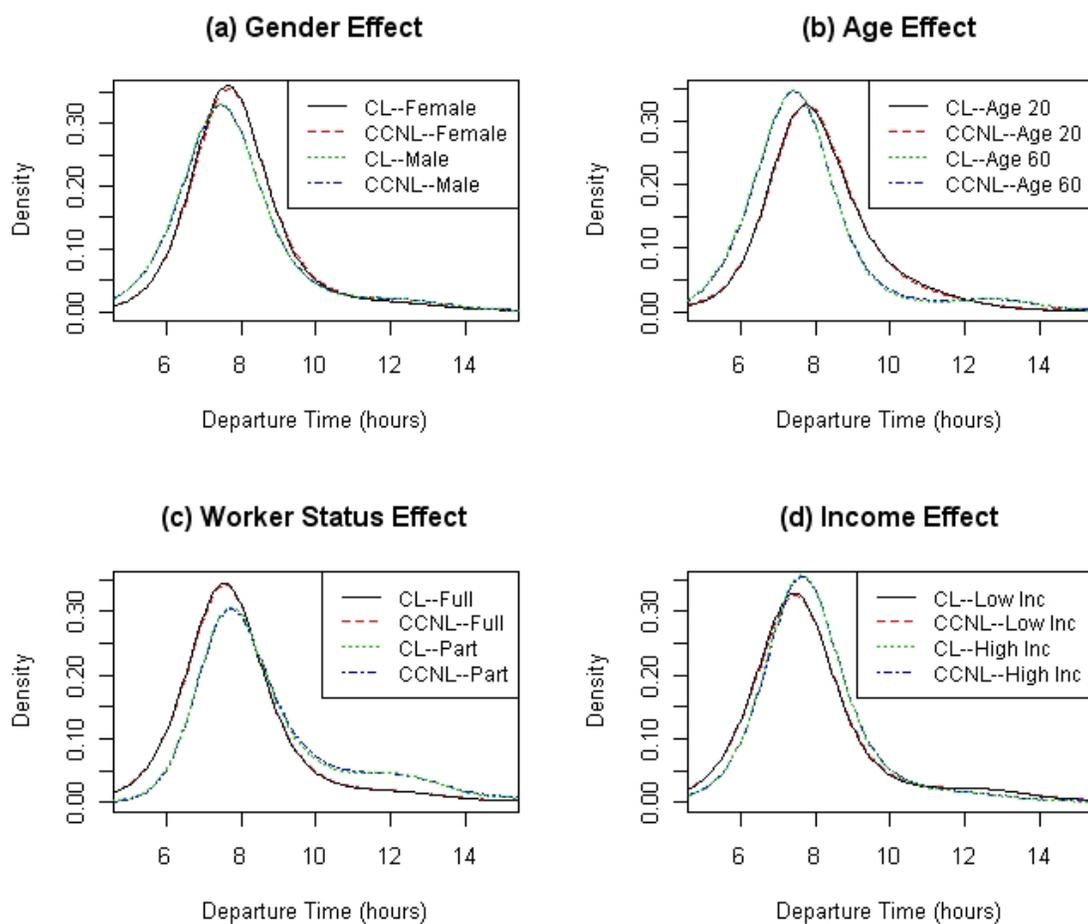


Figure 5.2: Gender, Age, Worker Status, and Income Effects on Average Individuals' Predictive Densities for Continuous Logit and CCNL

Figure 5.2a shows the differences in predictive densities for males and females, Figure 5.2b illustrates the age effect on departure time choice, Figure 5.2c reveals predictive densities for full-time versus part-time workers, and Figure 5.2d demonstrates departure time choice differences between individuals from high- and low-income households

⁴⁸ For instance, if the final covariate is an indicator variable, the two representative values are 0 and 1. For other covariates, a low value and a high value are chosen to illustrate differences across individuals.

(where high income is defined as \$75K per year or more). Clearly, men and older individuals tend to depart earlier than women and younger persons. In the case of the age variable, similar results were noted by Komma and Srinivasan (2008) and Gadda et al. (2009). Komma and Srinivasan (2008) also found those from high income households tend to depart later, similar to the results presented here⁴⁹. Presumably this results from such individuals having more flexible work schedules. Finally, both models predict part-time workers to depart later, relative to full-time workers, consistent with the findings of Abou Zeid et al. (2006), Popuri et al. (2008), and Gadda et al. (2009). Since part-time workers work schedules often do not conform to those of full-time workers, and they presumably work shorter shifts, in general, this result seems very reasonable.

Figure 5.3 shows the effect of the four remaining covariates on departure time choice, including household size (1 versus 5 in Figure 5.3a) the number of other tours (0 and 3 in Figure 5.3b), the travel distance (2 mi versus 35 mi in Figure 5.3c), and the effect of traveling to a central business district (CBD) destination (Figure 5.3d). Individuals from larger households tend to depart earlier (similar to results found by Komma and Srinivasan [2008]), which is not so surprising since they may have additional obligations, such as dropping off a child at school. The number of additional tours an individual undertakes tends to result in later departures, thanks to added scheduling constraints and shorter work-activity durations⁵⁰. Not surprisingly, the longer the distance traveled, the earlier one departs⁵¹, presumably in order to arrive on time. Travelers destined for the CBD tend to depart a bit later, which is in contrast to the findings of Komma and Srinivasan (2008). In fact, one may expect such individuals to depart earlier, since one would typically find lower speed roadways and higher congestion in such areas during

⁴⁹ Popuri et al. (2008) noted later departures for individuals from households with 2 or more vehicles, which are generally higher income households. Conversely, Gadda et al. (2009) found the opposite affect for high income households, though parameter estimates in that case were not statistically significant for the variable.

⁵⁰ Note that average work-activity durations for those with additional tours is only about 6 hours, while durations for those with no additional tours is just a bit greater than 8 hours.

⁵¹ Komma and Srinivasan (2008) and Gadda et al. (2009) obtained similar results.

the AM peak. However, the model controls for congestion and speed effects through the time-varying covariates. Those working in the CBD may have particular job types, which allow for later work start times. Komma and Srinivasan (2008) controlled for two occupation types, finding that those in “professional” occupations depart a fair bit later than others, so those working in the CBD may largely be working in “professional” occupations.

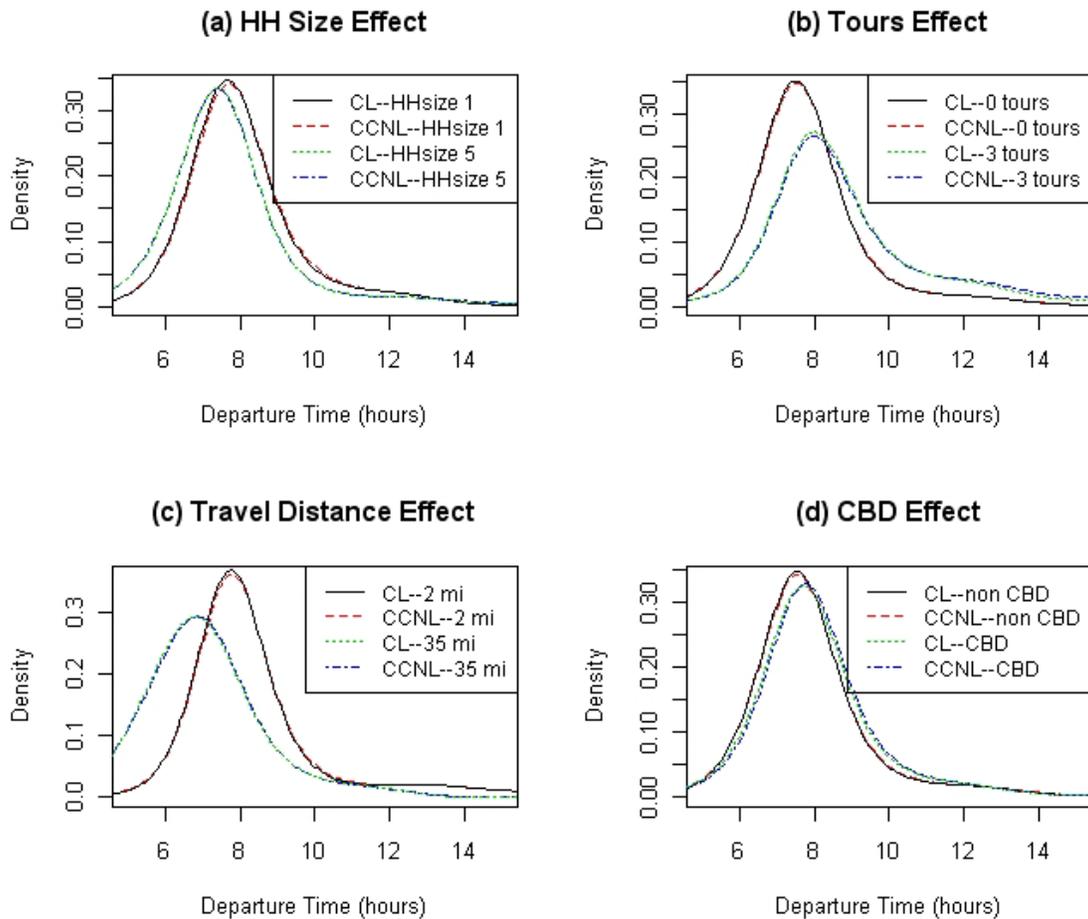


Figure 5.3: Household Size, Number of Tours, Travel Distance, and CBD Effects on Average Individuals’ Predictive Densities for Continuous Logit and CCNL

Interestingly, the differences in predictive densities from the continuous logit and CCNL appear negligible in each case. This is a bit odd since the continuous logit assumes

independent and identically distributed error terms. Perhaps the continuous form of the systematic utility equation plays a role here. Since the systematic utility equation is defined as a smoothly continuous function, systematic utilities for alternatives separated by only small distances cannot have largely different systematic components. And, since one would only expect error term correlations between alternatives when similarities between alternatives are not systematically controlled for, it is possible that the smoothly continuous utility specification actually controls for these similarities in a systematic way. Of course, this is only one explanation. Perhaps the departure time choice context here is not as suitable for demonstrating the usefulness of the CCNL specification as some other choice contexts (e.g., location or destination choice, activity durations, or vehicle usage). Alternatively, the data may be to blame. It is no secret that departure time data often suffers from rounding error and misreporting (Stopher et al. 2008), which could cause clustering in the data. Such clustering may cause difficulty in resolving the underlying correlations for the CCNL in model estimation. Further comparisons between the two models, via out-of-sample predictive accuracies, are provided in the following section.

5.3.1 Out-of-Sample Predictions

In order to evaluate the models' abilities to capture variation in departure time choice, out-of-sample prediction provides a number of benefits. Furthermore, it aids in illustrating the merits of Bayesian methods. The out-of-sample data is composed of 3,550 records, representing about 20% of the total data. Since Bayesian estimation output offers a collection of parameter draws from the posterior distribution, each draw is used to compute the likelihood each model would predict the actual departure time outcome for each individual. The distribution of individual-level likelihoods and total log likelihoods can then be characterized. Here, each of the 2,000 continuous logit and 1,000 CCNL posterior draws are employed. Figure 5.4a shows the total log likelihood

(across all individuals) distributions and Figure 5.4b shows the individual-level likelihood distributions for the two models⁵².

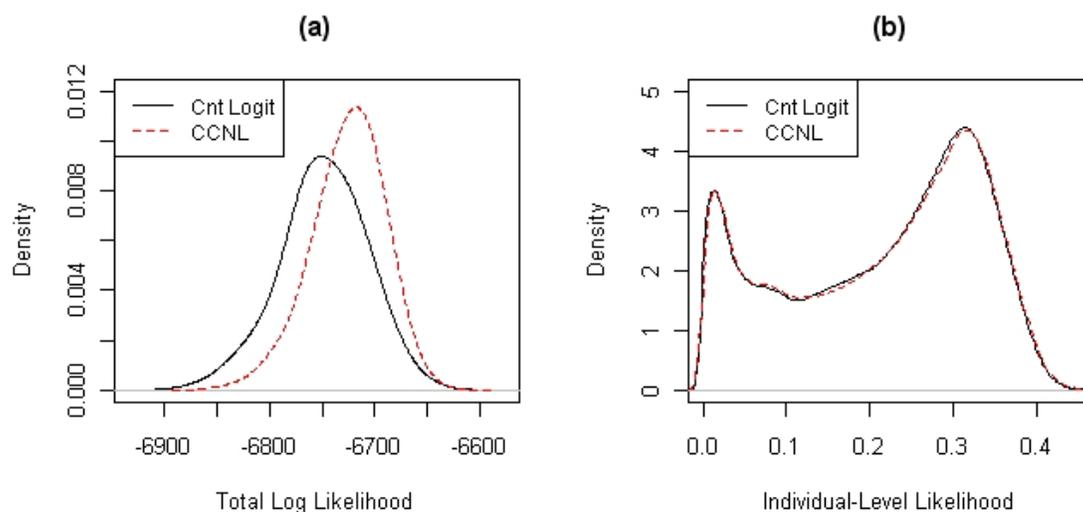


Figure 5.4: Out-of-Sample Likelihood Predictions for Continuous Logit and CCNL

As shown in Figure 5.4a, the CCNL out-performs the continuous logit, with a difference in mean log likelihoods of about 23. Figure 5.4b shows almost no differences between the two models, not surprisingly, since this plot is over individuals, and not the aggregate. If the average aggregate log likelihood difference is 23 across 3,550 individuals, this suggests just a 0.006 average contribution per individual for the CCNL over the continuous logit.

Good (1958) proposed Bayes factors (BFs) for testing whether differences between two models are significant.⁵³ And, the BF here is the exponent of the difference between the mean log likelihoods of Figure 5.4a. Kass and Raftery (1995) propose a test statistic of $2\ln(BF)$, suggesting values between 2 and 6 provide positive evidence of rejecting the

⁵² In other words, the density value for each individual was computed, and the distribution of these density values across all individuals and all parameter draws is plotted in Figure 5.4b.

⁵³ Other measures have been proposed to evaluate model fit in Bayesian contexts as well, including the Bayesian information criterion (Schwartz 1978), the Akaike information criterion (Akaike 1974), and the deviance information criterion (Spiegelhalter et al. 2002).

null hypothesis (i.e., the continuous logit in this case), values between 6 and 10 provide strong evidence, and values over 10 offer very strong evidence for rejecting the null hypothesis. In this case, the test statistic takes a value of 48, offering strong evidence that the continuous logit model is inferior to the CCNL in terms of model prediction. Kass and Raftery (1995) also note that $2\ln(BF)$ serves as the Bayesian equivalent of a likelihood ratio test statistic for classical methods. Thus, if the test statistic value of 48 is scaled for the observations in the estimation sample (i.e., multiplied by $\frac{997}{3550} = 0.281$), the equivalent likelihood ratio statistic here is about 13. With two degrees of freedom (since the CCNL adds two parameters to the continuous logit), the test statistic is statistically significant.

Alternatively, one could think of measuring how often the CCNL beats the continuous logit (in terms of its predictive ability). For instance, if random pairs of parameter values are drawn from their respective posterior distributions (i.e., one from the continuous logit and one from the CCNL), the CCNL's corresponding likelihood beats the continuous logit's about 65% of the time. Of course, the practical benefit of the CCNL over the continuous logit does not appear as meaningful as these results suggest (and as noted in the previous section). In the next section, some example policy simulations are examined, demonstrating how economic welfare calculations can be made for the two models.

5.3.2 Economic Welfare Demonstration

To illustrate how economic welfare change can be handled for the continuous logit and CCNL models, an example is formulated here. Ben-Akiva and Watanatada (1981) showed that consumer surplus (CS) for the continuous logit can be computed as the limiting formula for the MNL, as follows:

$$CS_i = \ln \left(\int_a^b \exp[V_i(t)] dt \right) \quad (5.1)$$

Moreover, one can convert differences in CS into monetary terms by dividing the term in equation 5.1 by the estimated cost coefficient. CS for the discrete cross-nested logit is computed as follows (Hunt et al. 2007):

$$CS_i = \ln \left(\sum_{m=1}^M [\sum_{j \in C_m} (\alpha_{jm} y_j)^{\rho_m}]^{\frac{1}{\rho_m}} \right) \quad (5.2)$$

It follows that the CS for the CCNL can be computed as follows:

$$CS_i = \ln \left(\int_a^b \left(\int_{q-h}^{q+h} [\alpha(r, q) y(r)]^\rho dr \right)^{\frac{1}{\rho}} dq \right) \quad (5.3)$$

As a base scenario, travelers are assumed to face the conditions provided in the data. Three tolling policy simulations are investigated here. In the first, it is assumed that \$0.15/mile tolls are assessed on all roads during the peak periods, resulting in peak-period travel time delay reductions of 50%. The \$0.15/mile toll represents a doubling in travel costs for most travelers (since operating costs are assumed to be \$0.15/mile), though some roadways in the San Francisco region are already tolled. So the additional \$0.15/mile toll represents less than a doubling of costs for travelers incurring other tolls. In the second simulation, \$0.15/mile tolls are assessed during peak periods, with assumed peak-period travel time delay reductions of only 10%. And in the final simulation, \$0.30/mile tolls during peak periods are assumed to reduce peak-period travel time delay by 50%. Clearly, there is some discrepancy here between the three scenarios (e.g., \$0.15/mile tolls cannot simultaneously reduce peak period travel time delay by 50% and 10%). However, these scenarios should be viewed simply as potential outcomes of the tolling policies. In order to truly understand the tolls' effects, a feedback mechanism would be required. Here, the purpose is not necessarily to obtain a good guess of the

tolls' effects, but instead to illustrate how CS change calculations are performed and to examine the continuous logit and CCNL model behavior.

Using these assumptions, an individual's CS change is measured by the difference in CS values across the base scenario and the tolling policy scenarios, divided by the cost coefficient from the model. Since this is a Bayesian analysis, CS change is computed for each of 100 parameter draws from the posterior distribution for the two models. The same random draws are used in computations for each of the three simulations. The sample used here is identical to the sample used in model estimation (with 997 tour observations). Because individuals choosing transit modes would not be expected to incur such tolls (nor would those traveling by walk and bike modes), only those sample tours made by the automobile mode are considered here (for 821 sample records).

Figure 5.5a shows the distribution of aggregate CS change (over all sample individuals) under the continuous logit specification for the three policy simulations, while Figure 5.5b shows the aggregate CS change distribution under the CCNL specification. Simulations 1 and 2 (with the only difference being the assumed reduction in peak travel time delay) have similar CS distributions under both model specifications, with simulation 1's density shifted to the right. Since travel times are lower and more reliable under simulation 1, it is not surprising that the estimated CS change is greater. The \$0.15/mile tolls during the peak periods (where the AM peak period is highly preferred) essentially result in a doubling of total costs for most individuals (since operating costs are assumed to be \$0.15/mile as well). Thus, it is not surprising to see the large drop in CS change for the first two simulations. It is also not surprising to see simulation 3's CS reduction is about twice as large as that of simulation 1, on average.

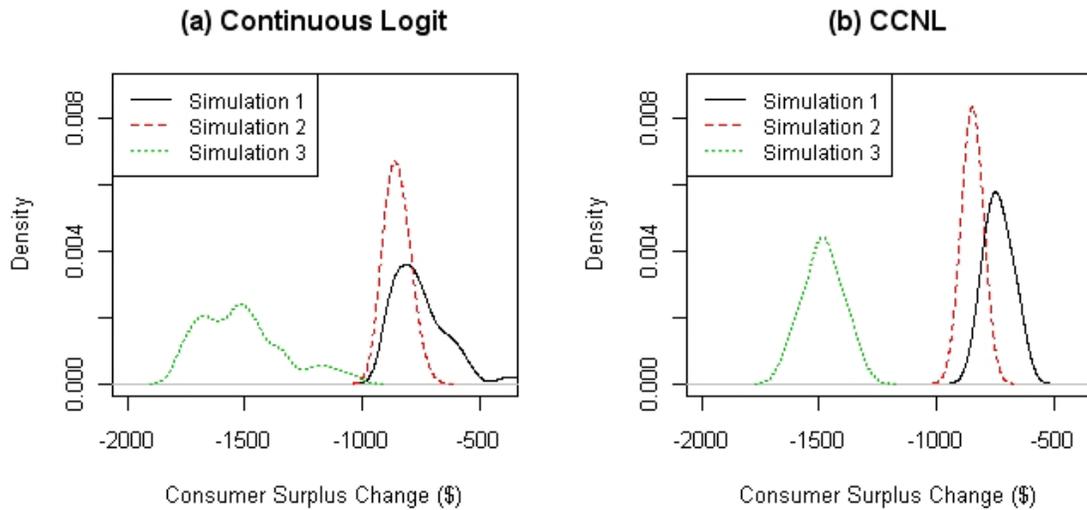


Figure 5.5: Distribution of Consumer Surplus Change for Continuous Logit and CCNL under Three Tolling Policy Simulations

Figure 5.6 shows the same information as Figure 5.5, but is presented by scenario to offer ease in comparisons across the two models. In each policy simulation, the general shape of CS change distributions is very similar for the two models, though the CCNLs' distributions have smaller variances than those of the continuous logit. Mean CS changes per traveler (across the 821 travelers) under the continuous logit are about $-\$0.88$, $-\$1.04$, and $-\$1.84$ for simulations 1, 2, and 3, respectively, while under the CCNL, mean CS changes are about $-\$0.89$, $-\$1.03$, and $-\$1.81$. Not surprisingly, these mean CS changes are very similar across the two models. However, standard deviations (at the traveler level) for the three policy simulations are $\$0.20$, $\$0.05$, and $\$0.22$ under the continuous logit and $\$0.06$, $\$0.03$, and $\$0.10$ under the CCNL. Thus, CS estimates under the continuous logit specification suffer from greater uncertainty, which is due to the larger interval estimates for the network variable parameters (i.e., the parameters related to travel time, variance, and cost) under the continuous logit specification. It should be noted that even though CS is negative under each tolling policy simulation, total economic welfare is not, since one must also consider toll revenues. After accounting for toll revenues, the mean of total economic welfare change at the traveler level is estimated

to be \$0.33, \$0.17, and \$0.40 for simulations 1, 2, and 3, respectively, under the continuous logit model; and \$0.36, \$0.22, and \$0.41 for the three simulations under the CCNL model. Of course, these are only three examples, and drawing firm conclusions at this stage may be unwise; but the example does illustrate how CS is computed for the CCNL, while illustrating a key advantage of the Bayesian approach (i.e., obtaining the distribution of CS rather than simply a point estimate).

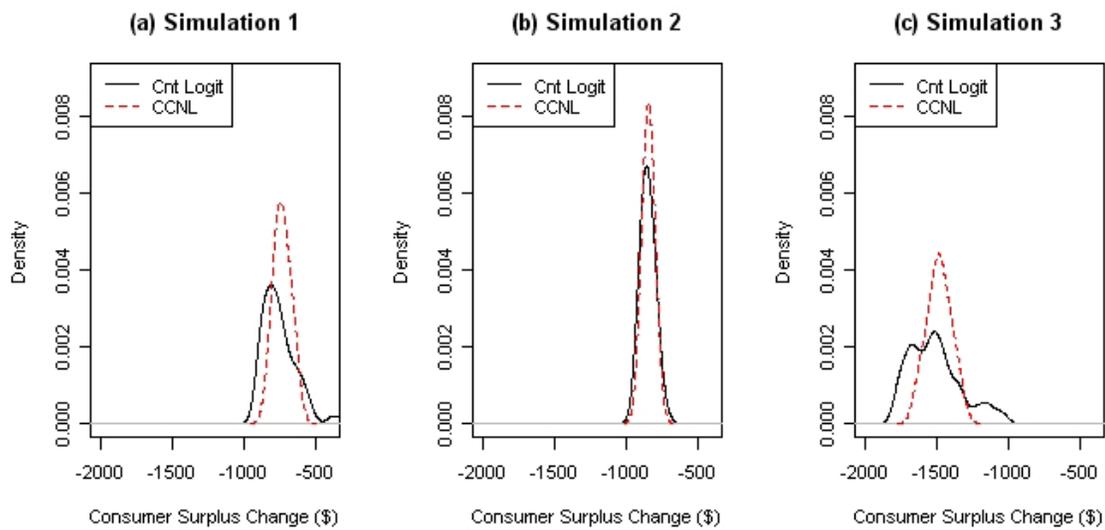


Figure 5.6: Distribution of Consumer Surplus Change for Three Tolling Policy Simulations under Continuous Logit and CCNL Specifications

Another interesting evaluation can be made with these simulations with regard to how the models predict departure time choice changes at an aggregate level. Figure 5.7 shows the aggregate (over all individuals) departure time choice distributions for each of the simulations along with the status quo (or baseline scenario). As expected, the predictive distributions are wider (and less peaked) under each of the simulations as compared to the status quo scenario for both models. In addition, there appears to be some evidence of grouping near the left- and right-side shoulders of the AM peak period, particularly under simulation 3. This seems very reasonable since these shoulders represent times right before and after tolling is assumed to begin.

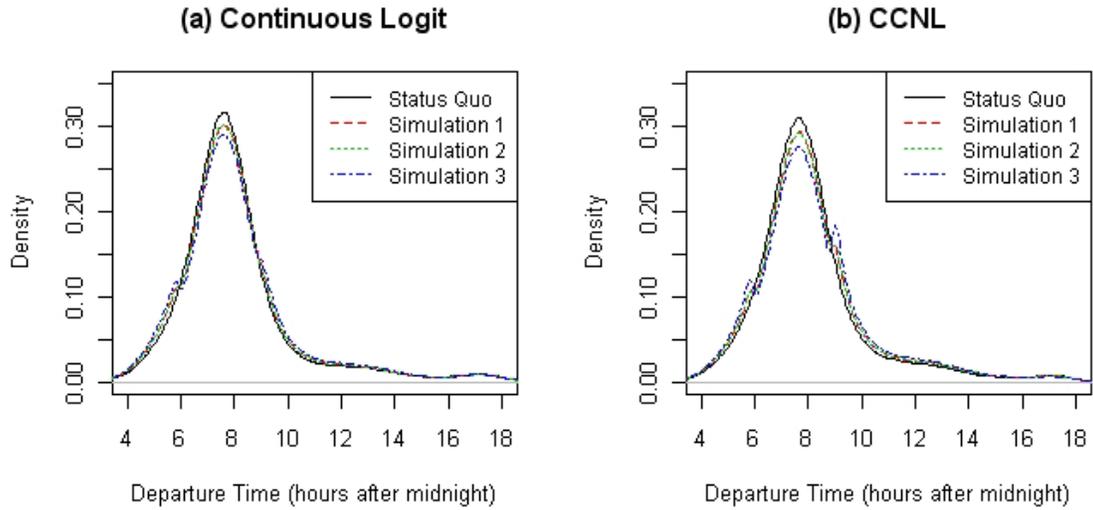


Figure 5.7: Distribution of Travelers' Departure Time Choices for Four Simulations under Continuous Logit and CCNL

Figure 5.8 shows the same plots as Figure 5.7, but with the densities grouped by simulation rather than by model. Figure 5.8 suggests that the two models predict very similar results under the status quo scenario, which is not surprising. But for each of the toll policy simulations, the aggregate predictive densities are slightly different. This is particularly evident with simulation 3, where two new peaks in the distribution emerge just before and just after the AM peak period. Under the CCNL specification, the height of these peaks appears to be a bit more pronounced than under the continuous logit specification. Since the correlations under the CCNL essentially allow areas of higher utility to draw probability away from areas of lower utility, it is not surprising to see such behavior. In fact, this behavior seems very reasonable, since given the choice between driving to work at 5:55 with no tolls or at 6:05 with tolls, most people would likely choose the earlier time.

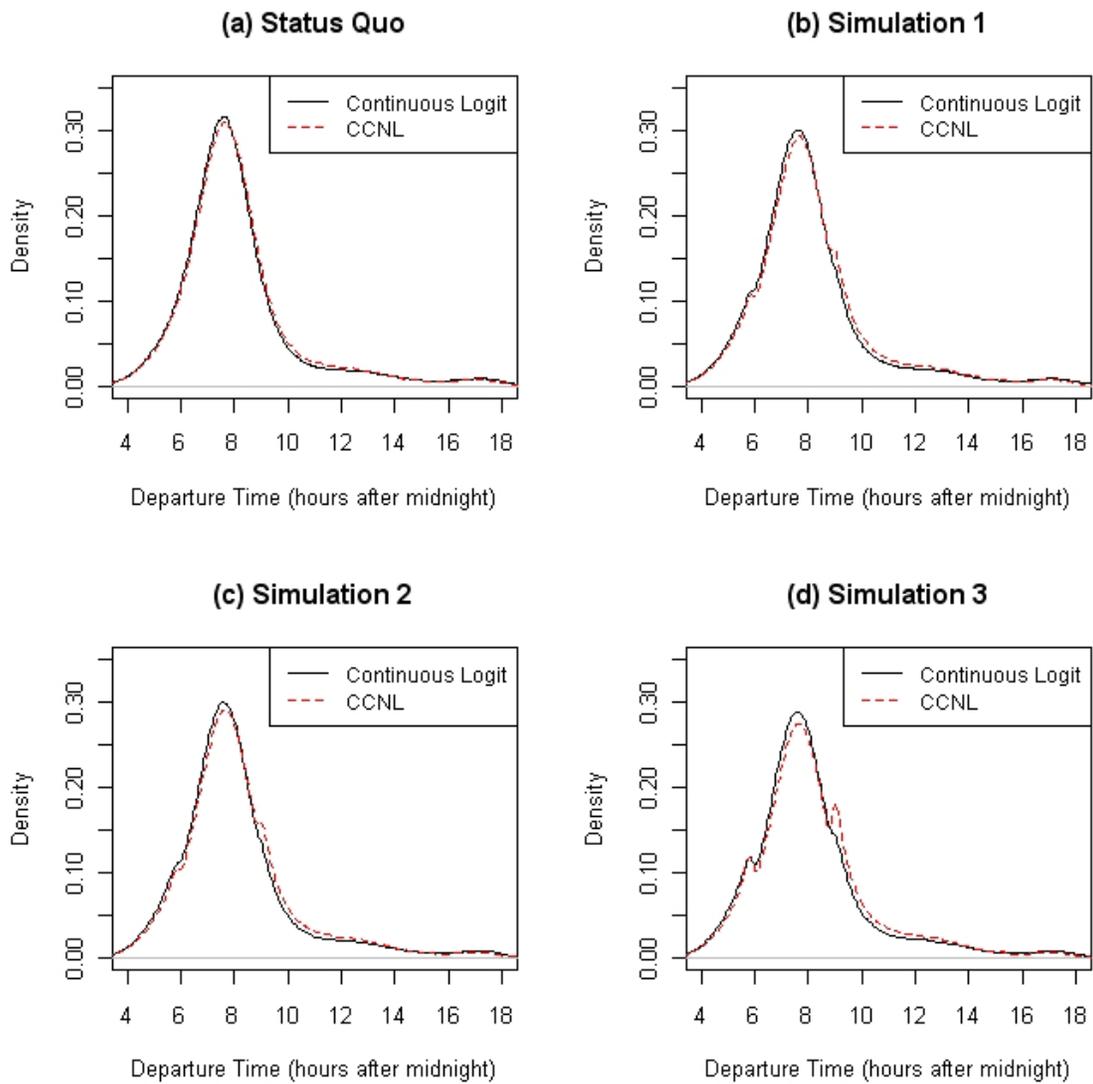


Figure 5.8: Distribution of Travelers' Departure Time Choices for Continuous Logit and CCNL by Simulation Exercise

To further investigate these predictive distributions, Table 5.4 presents the proportion of individuals choosing each of five TOD periods for each simulation and both models. The TOD periods include before 5 am, 5 to 6 am, 6 to 9 am (the AM peak), 9 to 10 am, and after 10 am. These were specifically chosen to examine the shoulder periods one hour before and one hour after the tolled AM peak period. While the status quo simulation proportions are similar between the two models, the CCNL predicts larger percentage

drops in AM peak period travelers in each of the three toll policy simulations relative to the status quo. More importantly, however, are the predicted changes relative to the peak change. These are computed as the difference in TOD-specific shares between the toll policy simulations and status quo simulation, divided by the status quo simulation's TOD share. As shown in Table 5.4, the continuous logit predicts larger share increases in the "before 5 am" TOD period than in the "5 to 6 am" period for each simulation, while the CCNL predicts the opposite. In contrast to the continuous logit, the ability of the CCNL to capture correlations across alternatives allows this very reasonable behavior to emerge. This behavior is also evident in comparison of the "after 10 am" and "9 to 10 am" periods, where "9 to 10 am" share increases under the CCNL are relatively larger than "after 10 am" share increases when compared to the continuous logit. Of course, these results seem very reasonable, since one would expect travelers to shift toward shoulder peak periods more so than to other periods.

Table 5.4: Predicted Departure Time Proportions for Five TOD Periods and Four Simulations

	Proportion		Proportional Increase from Base	
	Cnt. Logit	CCNL	Cnt. Logit	CCNL
Status Quo				
Before 5 am	0.031	0.030	n/a	n/a
5 to 6 am	0.074	0.070	n/a	n/a
6 to 9 am (Peak)	0.700	0.691	n/a	n/a
9 to 10 am	0.079	0.087	n/a	n/a
After 10 am	0.116	0.123	n/a	n/a
Simulation 1				
Before 5 am	0.036	0.033	0.155	0.101
5 to 6 am	0.084	0.079	0.135	0.135
6 to 9 am (Peak)	0.671	0.652	-0.041	-0.056
9 to 10 am	0.086	0.100	0.086	0.148
After 10 am	0.123	0.136	0.060	0.107
Simulation 2				
Before 5 am	0.036	0.033	0.166	0.101
5 to 6 am	0.084	0.080	0.136	0.136
6 to 9 am (Peak)	0.670	0.651	-0.043	-0.058
9 to 10 am	0.085	0.101	0.081	0.168
After 10 am	0.124	0.136	0.072	0.107
Simulation 3				
Before 5 am	0.040	0.036	0.284	0.220
5 to 6 am	0.092	0.090	0.243	0.279
6 to 9 am (Peak)	0.643	0.613	-0.082	-0.113
9 to 10 am	0.093	0.114	0.182	0.313
After 10 am	0.132	0.148	0.137	0.204

Figure 5.9 further details these shifts in departure time choices by displaying the portion of travelers shifting from peak periods to combined shoulder periods (i.e., “5 to 6 am” and “9 to 10 am”) and to combined off-peak periods (i.e., “before 5 am” and “after 10 am”) for both models. In other words, of those travelers no longer choosing peak periods under the simulations’ tolls, Figure 5.9 shows the share of those travelers shifting to peak shoulders versus off-peak periods. In some sense, this assumes that only peak-period travelers shift their departure times in the face of tolls, and either shift to peak shoulders

or off-peak periods. One would expect most peak-period travelers to shift to similar departure time alternatives (i.e., peak shoulders) in the face of peak-period tolls. In comparison to the continuous logit's predictions, the CCNL generally predicts larger shifts to peak shoulders (though shifts under simulation 1 are nearly identical), as expected. Of course, this comes from the CCNL's ability to capture correlations across departure time choices.

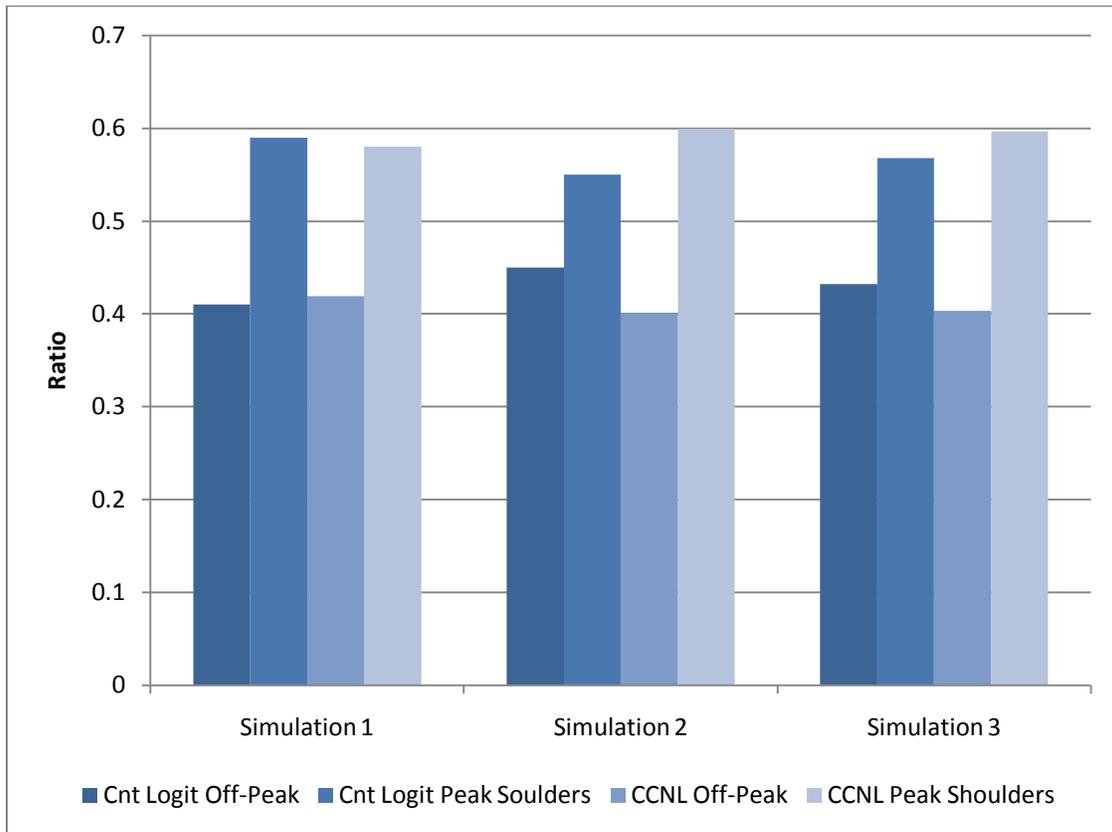


Figure 5.9: AM Peak Period Arrival Time Shifts to Shoulder and Off-Peak Periods for Three Tolling Policy Simulations under Continuous Logit and CCNL Models

While the added ability to capture correlations across choice alternatives under the CCNL model (in contrast to the continuous logit) appears to have only minor implications on individual-level predictive densities (as illustrated in Figure 5.2 and Figure 5.3), out-of-

sample predictions are better (and statistically significant) for the CCNL. In addition, the empirical results suggest the CCNL offers reasonable choice behavior in the departure time choice context while providing added model flexibility (and a more defensible behavioral basis) over the continuous logit specification.

5.4 Chapter Summary

A number of choice contexts may be best handled using continuous response variables (e.g., departure time, location, activity duration, and vehicle usage). This chapter has presented empirical results of a new GEV model for continuous response, the CCNL. The model represents a generalization of the discrete cross-nested logit for continuous response, much like the continuous logit represents a generalization of the MNL. And, like any in the GEV class, the model conforms to random utility theory, offering a strong basis by which to estimate the economic welfare implications for evaluation of policy alternatives, as demonstrated in Section 5.3.2.

Empirical results suggest that the CCNL performs better than the continuous logit model, in terms of out-of-sample prediction of departure times, and the CCNL offers more flexible choice behavior to emerge (as illustrated through three tolling policy simulations), along with welfare estimates. However, the CCNL model is much more computationally burdensome in estimation. Here, generating draws from the CCNL's posterior distribution took on the order of 30 times longer than those for the continuous logit. Moreover, the numerical integration procedure for generating likelihood values suffered from more error in the CCNL context than in the continuous logit setting. Reducing such error is likely to result in even greater estimation time. In the next chapter, empirical results of the BVMNP model (developed in Chapter 3) are presented.

CHAPTER 6: EMPIRICAL RESULTS OF DISCRETE TOD MODELS

In this chapter, empirical results of the bivariate multinomial probit (BVMNP) model specifications are presented. The models were estimated using Bayesian estimation methods (as discussed in Chapter 3) on work-tour arrival- and return-time choice data. For consistency in model estimates, the same data used in Chapter 5's continuous model estimations is used here. Section 6.1 discusses this data in more detail as it relates to bivariate choice of tour timing.

6.1 Work-Tour Scheduling Data

As mentioned above, the sample used here represents the same sample used in the estimation of the continuous logit and CCNL. In other words, the same random sample of 997 observations is used in this analysis. Instead of only a single dimension of tour timing investigated for the previous models, two timing dimensions are considered: arrival time at work and return time from work. These were chosen to provide simplicity in defining choice alternatives that are not possible. With these choice definitions, alternatives that are not possible are those where the return time alternative is earlier than the arrival time alternative. Figure 6.1a shows the distributions of home-to-work arrival times for the full sample (of 17,820) and the 6% sample used in model estimation, and Figure 6.1b shows the distribution of work-to-home return times. Only minor differences between the samples are evident, though the peak of the distribution for return times is slightly lower for the 6% sample as compared to the full sample. This is likely due to work-to-home timing dimension being more variable (i.e., having higher variance) than the home-to-work timing dimension.

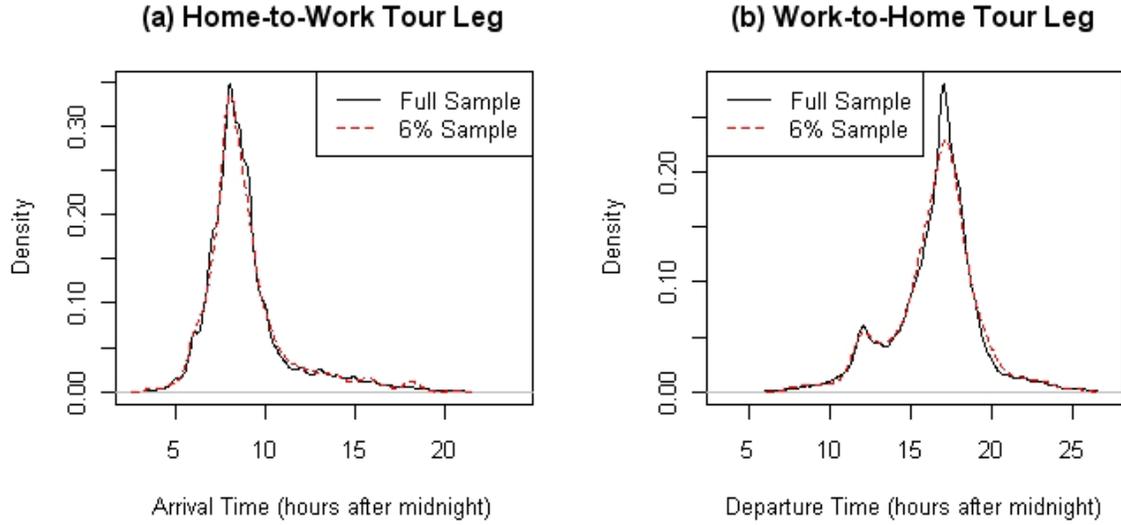


Figure 6.1: Densities of Home-to-Work Arrival Times and Work-to-Home Departure Times

As discussed in Chapter 3, 30-minute time intervals serve as the choice alternatives here. Since there are very few individuals choosing times very early in the day and very late in the day (for both arrival and return choice dimensions), boundary alternatives were created, essentially grouping many 30-minute alternatives into a single alternative. Boundaries were set here to achieve about 10 observations within each boundary category. For work arrival times, this resulted in an early morning period from midnight to 5 am (9 observations) and an evening period from 6 pm to midnight (12 observations). For work return times, an early morning period from midnight to 10 am (15 observations) and an evening period from 11 pm to midnight (16 observations) were created. As the utility equations for the BVMNP model rely on cyclical functions of time (arrival time or return time), assumptions regarding the time implied by each boundary alternative were needed. Assumed arrival times for work arrival boundary alternatives are 4:30 am and 7:00 pm, and assumed return times for work return boundary alternatives are 8:30 am and 11:30 pm. These assumed times represent average arrival/return times recorded for the observations within each category. To accommodate that these boundary intervals may

exhibit very different properties than non-boundary alternatives, variances and correlations for these boundary alternatives were adjusted. These adjustments are detailed in Section 6.2 below.

Explanatory variables considered in this analysis are identical to those evaluated in the previous time-of-day (TOD) choice models. For descriptive statistics on these variables, readers may refer to Table 5.1 in the previous chapter. Like the earlier models, monetary travel costs are computed as follows: for automobiles, they are \$0.15/mile (the operating cost) plus any applicable tolls, divided by the number of vehicle occupants for automobile modes; for transit, they are the transit fare (by TOD); and for walk and bike modes, they are zero. In the following section, model estimation details are presented.

6.2 Model Estimation Details

While Chapter 3 discusses the Bayesian estimation procedure in detail, some additional details are needed here. First, to ensure the time-varying covariates of travel time, travel time variance, and travel cost have the expected (negative) effect on utility functions, the priors on parameters related to time, variance, and cost variables were chosen to be truncated normal (i.e., truncated from above at zero). Second, since very few individuals in the sample choose very early or very late arrival and return time periods, “boundary” alternatives were created (as discussed above). Of course, this simplifies the model in the sense that fewer alternative-specific utilities need to be drawn in model estimation. However, it creates problems in that one could not reasonably expect the error structures for these alternatives to be the same as those of other alternatives. Thus, several modifications of the covariance matrix are specified here.

First, each of the four boundary alternatives’ error terms is allowed to have its own variance parameter, while the variance parameters for other alternatives were set to one for identification purposes (as discussed in Section 3.3). Independent gamma priors are used for these parameters with shape and scale parameters of 2 and 1, respectively.

Second, for the CAR covariance specification, the λ terms (which control the correlation across alternatives) are assumed to be inversely related to each alternative's interval size. More specifically, these parameters are assumed to vary across each pair of alternatives as follows:

$$\lambda_{CAR,a,pq} = \lambda_{CAR,a} \left(\frac{\text{size}_p}{\text{size}_{\text{def}}} \right)^{-\tau_{CAR,ar}} \left(\frac{\text{size}_q}{\text{size}_{\text{def}}} \right)^{-\tau_{CAR,ar}} \quad (6.1)$$

$$\lambda_{CAR,r,pq} = \lambda_{CAR,r} \left(\frac{\text{size}_p}{\text{size}_{\text{def}}} \right)^{-\tau_{CAR,ar}} \left(\frac{\text{size}_q}{\text{size}_{\text{def}}} \right)^{-\tau_{CAR,ar}} \quad (6.2)$$

$$\lambda_{CAR,d,pq} = \lambda_{CAR,d} \left(\frac{\text{size}_p}{\text{size}_{\text{def}}} \right)^{-\tau_{CAR,d}} \left(\frac{\text{size}_q}{\text{size}_{\text{def}}} \right)^{-\tau_{CAR,d}} \quad (6.3)$$

Here, size_{def} is 30 minutes (the default interval size), and size_p and size_q are the interval sizes of alternatives p and q (measured in minutes). $\lambda_{CAR,a}$, $\lambda_{CAR,r}$, and $\lambda_{CAR,d}$ are the same as described in Chapter 3, and $\tau_{CAR,ar}$ and $\tau_{CAR,d}$ are two new (non-negative) parameters to be estimated. The assumption of non-negativity presumes that correlations between boundary alternatives and other alternatives are smaller than those across non-boundary alternatives. It should be clear that the size terms only come into play when one or both of the alternatives are boundary alternatives. Moreover, these size terms reduce to the original model specification if $\tau_{CAR,ar}$ or $\tau_{CAR,d}$ are zero. Finally, it is worth noting that $\tau_{CAR,ar}$ affects both the correlations across arrival alternatives and those across departure alternatives. Thus, it is assumed that the way in which the interval size affects correlation patterns across arrival times is the same as the way it affects correlation patterns across return time alternatives. Independent gamma priors are assumed for $\tau_{CAR,ar}$ and $\tau_{CAR,d}$ with prior shape and scale parameters of 1 each.

For the AR1 covariance specification, actual correlations between alternative utilities are scaled by the size of the alternatives. This is similar to the above concept, but realized in a slightly different way. The correlations between any two alternatives are formulated as follows:

$$Cor_{AR1}(p_a, q_a) = \lambda_{AR1,a}^{|t_{p_a} - t_{q_a}|} \left(\frac{size_p}{size_{def}} \right)^{-\tau_{AR1,ar}} \left(\frac{size_q}{size_{def}} \right)^{-\tau_{AR1,ar}} \quad (6.1)$$

$$Cor_{AR1}(p_r, q_r) = \lambda_{AR1,r}^{|t_{p_r} - t_{q_r}|} \left(\frac{size_p}{size_{def}} \right)^{-\tau_{AR1,ar}} \left(\frac{size_q}{size_{def}} \right)^{-\tau_{AR1,ar}} \quad (6.1)$$

$$Cor_{AR1}(p_a, q_r) = \lambda_{AR1,d}^{|t_{q_r} - t_{p_a} - (\mu_{AR1,1} + \mu_{AR1,2} t_{p_a}) + 1|} \left(\frac{size_p}{size_{def}} \right)^{-\tau_{AR1,d}} \left(\frac{size_q}{size_{def}} \right)^{-\tau_{AR1,d}} \quad (6.3)$$

Thus, correlations are assumed to be directly related to these size terms, unlike the CAR model where actual parameters of the model were re-parameterized. In both cases, however, the effect should be similar in that correlations will be smaller between boundary alternatives and non-boundary alternatives than they will be between two non-boundary alternatives, unless, of course, the τ parameters are estimated to be zero, in which case alternative size will have no effect on correlations. Like the CAR specification, prior distributions for $\tau_{AR1,ar}$ and $\tau_{AR1,d}$ are assumed to be independent gamma distributions with shape and scale parameters of 1 each.

The draws of the covariance parameters within the Gibbs sampler remain largely unchanged, except there are six additional parameters. The MH proposal density is a multivariate normal distribution, and parameters are redrawn if the resulting Σ_{full} or Σ_{part} do not emerge to be positive definite. The MH acceptance probability also is updated to include the prior densities of the additional parameters.

Initial values of all parameters in the utility function were set to zero as were all correlation parameters. Initial values of the boundary alternative variances were set to one (consistent with the fixed variances of the non-boundary alternatives), and initial values of the baseline duration parameters were set to 6 and 0 for part-time workers and those undertaking additional travel tours, and to 9 and 0 for full-time workers with no additional travel tours (consistent with average work-activity durations for these groups). The model was run to achieve about 500,000 draws from the parameters' posterior

distribution for both covariance specifications. To reduce autocorrelation across draws, only every 200th draw from the last 100,000 is used for inference. Thus, the posterior distribution here is characterized by 500 MCMC draws. Unlike the models of Chapter 5, it was clear that even after 500,000 draws, the models had not converged. This is largely a result of the slow-converging MH process used to draw random utilities in the first step of the Gibbs sampler. In Figure 6.2 and Figure 6.3, non-convergence of posterior draws for the BVMNP models is illustrated for four selected parameters, with parameter values on the y-axis and iteration number on the x-axis. Despite non-convergence, parameter draws did appear to offer very reasonable conclusions, and so are used here (even though convergence was not reached).

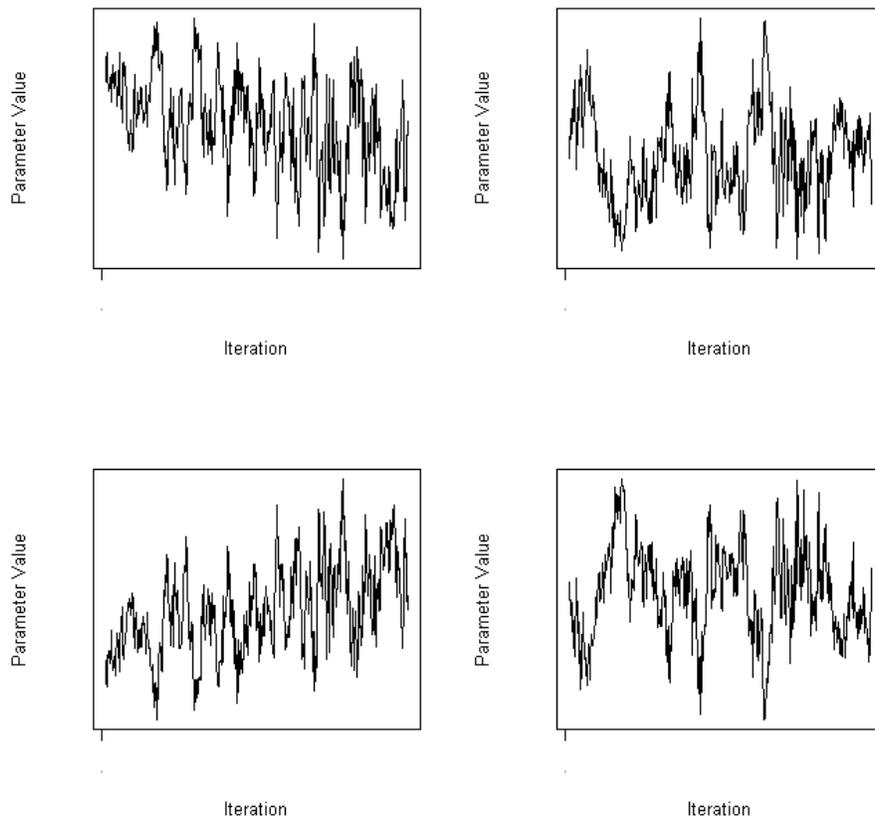


Figure 6.2: Trace Plot of Parameter Draws versus Iteration Number for Four Selected CAR Model Parameters (Every 200th Draw from Last 100,000)

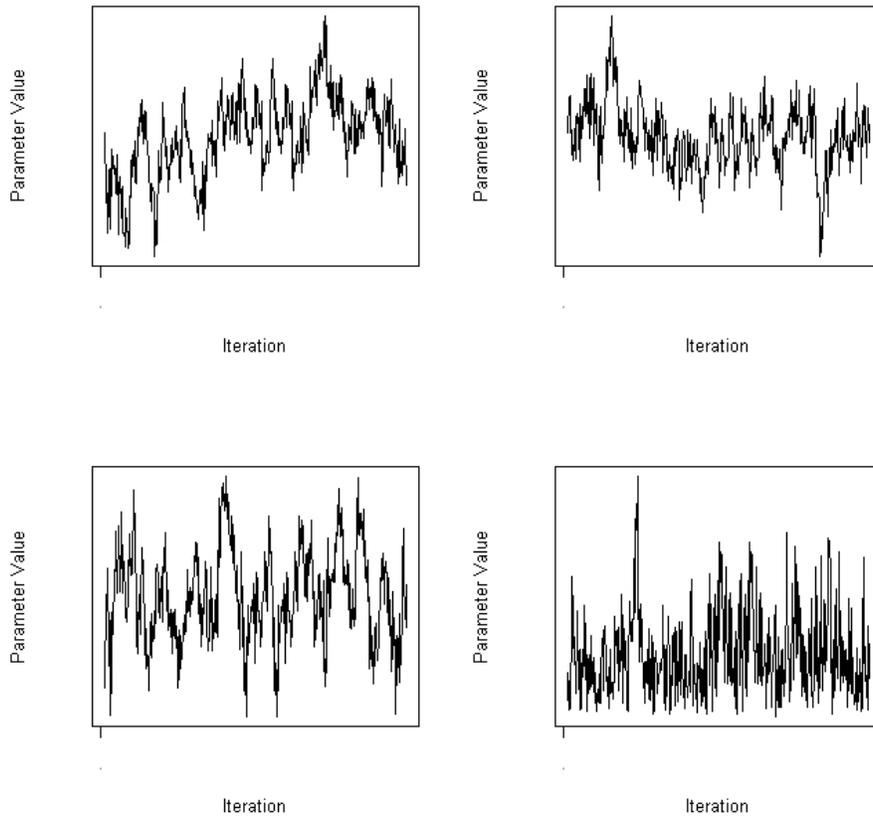


Figure 6.3: Trace Plot of Parameter Draws versus Iteration Number for Four Selected AR1 Model Parameters (Every 200th Draw from Last 100,000)

Like the model of Chapter 5, no formal variable selection methods were used. Instead, a similar set of covariates from the models of Chapter 5 were used. Of course, in the context of the BVMNP model, there are two parameters for each covariate in the utility equations rather than one, since the BVMNP model uses two separate utility functions for its two choice dimensions. To facilitate convergence, several covariate interaction terms were removed from the analysis. All of the same individual/tour-specific variables remain, but some of the interactions with higher-order cyclical functions were removed. Empirical results of the BVMNP are discussed in the following section.

6.3 Empirical Findings

Table 6.1, Table 6.2, and Table 6.3 show estimates of arrival time utility parameters, return time utility parameters, and covariance parameters for the two BVMNP models. Since parameters on time-varying covariates were constrained to be negative (through prior specification), those parameters all have negative signs (as shown at the end of Table 6.1 and Table 6.2). Moreover, implied values of travel time (VOTTs) for the two models are very similar: median arrival VOTTs for CAR and AR1 specifications are \$3.92/hour and \$5.41/hour, and return VOTTs are \$5.06/hour and \$4.74/hour, respectively. While arrival VOTTs are generally consistent with the estimates of such values for the continuous logit and CCNL models and return VOTTs have similar magnitudes, estimates are not consistent with expectations, since VOTTs are often estimated to be much higher (e.g., around \$8 to \$25 per hour [Brownstone and Small 2005]). Of course, these VOTT estimates are context-specific (for activity timing choices), and people may not trade off time and money in this time choice decisions (due to ignorance or scheduling constraints) like they do in route or mode choice contexts. Implied values of reliability (VORs) for the CAR and AR1 models are \$0.063 and \$0.047 per squared minute (about \$15.03 and \$13.02 per hour of standard deviation), respectively, on the home-to-work journey, and \$0.003 and \$0.004 per squared minute (about \$3.10 and \$3.75 per hour of standard deviation), respectively, on the work-to-home journey. While the actual estimates are not exactly the same across the two models (possibly due to non-convergent results), the two models are consistent in that reliability is much more important for the home-to-work journey than the work-to-home journey. Thus, it appears that commuters are more sensitive to travel time reliability in their home-to-work journeys than their work-to-home journeys, but less sensitive to travel time in their home-to-work journeys, relative to work-to-home. This seems rather intuitive, since workers are usually constrained in their working hours. Consider a worker that needs to arrive for work at 9 am. If this worker is late (or late on a regular basis), that may affect his or her job security. Thus, there is some incentive for leaving a buffer period to ensure arrival at or before work is scheduled to begin. In other words, it may be more

acceptable to the individual to arrive 10 minutes early than to arrive 10 minutes late, and so that individual may depart from home in order to arrive nearly 10 minutes early, on average.⁵⁴ On the other hand, most workers presumably do not need to be home from work at any particular time, though they may wish to arrive home as quickly as possible. Thus, it would make sense for average travel time to be more important for the work-to-home journey than the home-to-work journey. Thus, the implied VOTTs and VORs appear reasonable, relative to one another, but VOTTs are low in comparison to other studies.

Of course, the effect of these variables also depends on the variability in them. For instance, the average standard deviation of average travel time across TODs for sample individuals is about 4.4 minutes, while the average standard deviation of travel time variances across TODs is 22.2 squared minutes. Thus, if average travel times and travel time variances were to have the same effect on utilities, parameter coefficients on travel time should be about five times greater in magnitude than coefficients on variance. It turns out that conclusions remain the same here. Coefficients on times and variances for arrival time choice are similar in magnitude, suggesting reliability is more practically significant for the home-to-work journey, but coefficients on travel times for return time choice are on the order of 20 to 30 times greater than those on variances, meaning travel times are more practically significant for the work-to-home journey. One should again note, however, that these VOTT and VOR estimates are context-specific (for activity scheduling), and may not be valid for other choice contexts (e.g., mode or route choice) or under different network variable imputation assumptions.

⁵⁴ Small et al. (1999) estimated that the marginal costs of early arrival increase from about 0.028/min at 5 minutes early up to about \$0.128/min at 15 minutes early. The marginal cost of late arrival, however, was estimated to be 2.5 to 11 times greater at \$0.31/min.

Table 6.1: BVMNP Parameter Estimation Results for Arrival Time Variables

Arrival Time Variables		CAR Specification		AR1 Specification	
Variable	Interaction Term	Mean Estimate	95% Interval	Mean Estimate	95% Interval
Constant	$\sin(2\pi t/24)$	0.6730	(-0.1897, 1.6178)	0.2760	(-0.3667, 0.8177)
	$\sin(4\pi t/24)$	-1.0107	(-1.6965, -0.3341)	-1.0812	(-1.5824, -0.7299)
	$\cos(2\pi t/24)$	-3.9473	(-5.0264, -2.9246)	-1.3839	(-2.5713, -0.3558)
	$\cos(4\pi t/24)$	-1.3027	(-1.9116, -0.4098)	-0.3193	(-1.0473, 0.2298)
Male Indicator	$\sin(2\pi t/24)$	0.5664	(0.2417, 0.9221)	0.2574	(0.0203, 0.4377)
	$\sin(4\pi t/24)$	0.5806	(0.3191, 0.8805)	0.3170	(0.1493, 0.4936)
	$\cos(2\pi t/24)$	-0.1422	(-0.8963, 0.9342)	0.2555	(-0.1428, 0.6539)
	$\cos(4\pi t/24)$	0.0573	(-0.3515, 0.4605)	0.1481	(-0.0989, 0.3872)
Age	$\sin(2\pi t/24)$	0.0005	(-0.0106, 0.0132)	0.0013	(-0.0095, 0.011)
	$\sin(4\pi t/24)$	0.0017	(-0.009, 0.0138)	0.0015	(-0.0058, 0.0097)
	$\sin(6\pi t/24)$	-0.0092	(-0.0135, -0.0052)	-0.0085	(-0.0109, -0.0059)
	$\cos(2\pi t/24)$	0.0340	(0.0134, 0.0615)	0.0114	(-0.0134, 0.0324)
	$\cos(4\pi t/24)$	0.0227	(0.0102, 0.0354)	0.0105	(-0.0026, 0.0201)
	$\cos(6\pi t/24)$	0.0175	(0.0126, 0.022)	0.0127	(0.0089, 0.0157)
Part-Time Worker Indicator	$\sin(2\pi t/24)$	-1.9855	(-2.6685, -1.0937)	-1.5943	(-2.2234, -0.9971)
	$\sin(4\pi t/24)$	-0.7703	(-1.46, -0.0306)	-0.5840	(-1.1752, 0.0051)
	$\cos(2\pi t/24)$	-0.1136	(-1.7462, 1.6577)	-1.6182	(-3.3973, 0.056)
	$\cos(4\pi t/24)$	-0.1822	(-0.9828, 0.6103)	-0.7873	(-1.7212, 0.1405)
High-Income HH Indicator	$\sin(2\pi t/24)$	-0.1031	(-0.3988, 0.2918)	-0.0403	(-0.276, 0.1824)
	$\sin(4\pi t/24)$	-0.1972	(-0.4967, 0.1118)	-0.1293	(-0.3049, 0.0485)
	$\cos(2\pi t/24)$	-0.5877	(-1.2117, 0.1857)	-0.2214	(-0.7011, 0.1801)
	$\cos(4\pi t/24)$	-0.2208	(-0.4683, 0.1058)	-0.0285	(-0.2683, 0.2006)
Household Size	$\sin(2\pi t/24)$	-0.0752	(-0.2283, 0.0629)	-0.0317	(-0.116, 0.0408)
	$\sin(4\pi t/24)$	-0.0011	(-0.1285, 0.1108)	0.0180	(-0.0422, 0.0786)
	$\cos(2\pi t/24)$	0.3685	(0.0878, 0.7311)	0.0820	(-0.0573, 0.2134)
	$\cos(4\pi t/24)$	0.0741	(-0.1274, 0.2787)	-0.0322	(-0.1161, 0.0469)
Number of Other Tours	$\sin(2\pi t/24)$	-0.4858	(-0.6803, -0.2642)	-0.3450	(-0.4888, -0.1933)
	$\sin(4\pi t/24)$	-0.1215	(-0.2905, 0.0815)	-0.0845	(-0.2107, 0.0446)
	$\cos(2\pi t/24)$	0.4128	(0.0194, 0.8854)	0.1616	(-0.1414, 0.5264)
	$\cos(4\pi t/24)$	0.1641	(-0.0426, 0.367)	0.0863	(-0.0787, 0.31)
Free-Flow Distance	$\sin(2\pi t/24)$	0.0277	(0.0147, 0.0424)	0.0192	(0.0057, 0.0296)
	$\sin(4\pi t/24)$	0.0112	(0.0025, 0.0215)	0.0047	(-0.0059, 0.012)
	$\cos(2\pi t/24)$	0.0415	(0.013, 0.0658)	0.0372	(0.0188, 0.0615)
	$\cos(4\pi t/24)$	0.0307	(0.0149, 0.0462)	0.0245	(0.0139, 0.0379)
CBD Destination Indicator	$\sin(2\pi t/24)$	-0.3770	(-0.8408, 0.1825)	-0.2380	(-0.5849, 0.1159)
	$\sin(4\pi t/24)$	-0.3229	(-0.6963, 0.047)	-0.1935	(-0.5072, 0.1121)
	$\cos(2\pi t/24)$	1.3178	(0.0678, 2.5269)	0.7014	(-0.1569, 1.6148)
	$\cos(4\pi t/24)$	0.7145	(0.0349, 1.4205)	0.4703	(-0.0477, 0.8976)
Average Travel Time		-0.0047	(-0.0122, -0.0001)	-0.0039	(-0.0097, -0.0004)
Travel Time Variance		-0.0035	(-0.0053, -0.001)	-0.0022	(-0.0038, -0.0005)
Travel Cost		-0.0723	(-0.2152, -0.0018)	-0.0468	(-0.1187, -0.0008)

Table 6.2: BVMNP Parameter Estimation Results for Return Time Variables

Return Time Variables		CAR Specification		AR1 Specification	
Variable	Interaction Term	Mean Estimate	95% Interval	Mean Estimate	95% Interval
Constant	$\sin(2\pi t/24)$	-1.8787	(-3.7698, -0.6809)	-2.1569	(-2.6371, -1.6381)
	$\sin(4\pi t/24)$	-0.3494	(-1.1085, 0.2225)	-0.5066	(-0.9239, -0.062)
	$\cos(2\pi t/24)$	0.2008	(-0.5113, 0.9734)	-0.1751	(-0.6624, 0.3258)
	$\cos(4\pi t/24)$	-0.5984	(-1.096, 0.0905)	-0.0291	(-0.3726, 0.279)
Male Indicator	$\sin(2\pi t/24)$	0.4327	(-0.1097, 0.9983)	0.1220	(-0.1078, 0.3368)
	$\sin(4\pi t/24)$	0.1635	(-0.1388, 0.4256)	0.0261	(-0.1191, 0.1571)
	$\cos(2\pi t/24)$	0.4822	(0.1167, 0.7888)	0.2035	(-0.0155, 0.4049)
	$\cos(4\pi t/24)$	-0.1461	(-0.4282, 0.1229)	-0.0609	(-0.2202, 0.1105)
Age	$\sin(2\pi t/24)$	0.0208	(-0.0087, 0.0499)	0.0379	(0.0291, 0.0502)
	$\sin(4\pi t/24)$	0.0296	(0.0135, 0.0482)	0.0316	(0.0236, 0.0398)
	$\sin(6\pi t/24)$	0.0159	(0.0123, 0.0192)	0.0132	(0.01, 0.0165)
	$\cos(2\pi t/24)$	-0.0011	(-0.0164, 0.0128)	0.0110	(0.0008, 0.0198)
	$\cos(4\pi t/24)$	0.0079	(-0.0052, 0.0212)	-0.0021	(-0.0089, 0.0055)
	$\cos(6\pi t/24)$	-0.0041	(-0.0099, 0.0005)	-0.0044	(-0.0072, -0.0019)
Part-Time Worker Indicator	$\sin(2\pi t/24)$	1.0923	(-1.2257, 2.9789)	1.2840	(0.3705, 2.0983)
	$\sin(4\pi t/24)$	0.7050	(-0.0488, 1.4918)	0.7089	(0.1941, 1.154)
	$\cos(2\pi t/24)$	1.7009	(0.945, 2.721)	1.2802	(0.6692, 1.79)
	$\cos(4\pi t/24)$	-0.2847	(-1.2995, 0.8873)	-0.3608	(-0.9345, 0.2627)
High-Income HH Indicator	$\sin(2\pi t/24)$	-0.4165	(-1.1156, 0.4435)	-0.4598	(-0.7878, -0.1984)
	$\sin(4\pi t/24)$	-0.2732	(-0.6867, 0.0234)	-0.2629	(-0.4333, -0.0868)
	$\cos(2\pi t/24)$	-0.4033	(-0.9891, 0.0025)	-0.3832	(-0.6045, -0.1852)
	$\cos(4\pi t/24)$	0.0658	(-0.361, 0.3225)	0.0920	(-0.0559, 0.2249)
Household Size	$\sin(2\pi t/24)$	-0.0223	(-0.264, 0.3485)	0.0444	(-0.0324, 0.1121)
	$\sin(4\pi t/24)$	0.0432	(-0.0687, 0.1946)	0.0465	(-0.0164, 0.1042)
	$\cos(2\pi t/24)$	0.0833	(-0.0217, 0.2072)	0.0806	(0.0095, 0.1543)
	$\cos(4\pi t/24)$	0.1013	(-0.0547, 0.2231)	0.0510	(0.0013, 0.1096)
Number of Other Tours	$\sin(2\pi t/24)$	0.5508	(0.1755, 0.943)	0.1766	(-0.0053, 0.3967)
	$\sin(4\pi t/24)$	0.1097	(-0.1019, 0.2833)	0.0224	(-0.0954, 0.1762)
	$\cos(2\pi t/24)$	-0.3056	(-0.5042, -0.0974)	-0.2784	(-0.4348, -0.0851)
	$\cos(4\pi t/24)$	0.1561	(-0.0331, 0.3487)	0.1429	(0.0412, 0.2542)
Free-Flow Distance	$\sin(2\pi t/24)$	-0.0005	(-0.0296, 0.0211)	-0.0091	(-0.0303, 0.0114)
	$\sin(4\pi t/24)$	0.0053	(-0.0067, 0.0148)	0.0022	(-0.0061, 0.0101)
	$\cos(2\pi t/24)$	-0.0021	(-0.0181, 0.0121)	-0.0034	(-0.013, 0.0038)
	$\cos(4\pi t/24)$	0.0023	(-0.0084, 0.014)	0.0045	(-0.0042, 0.0149)
CBD Destination Indicator	$\sin(2\pi t/24)$	-1.5180	(-2.5096, -0.4648)	-2.3921	(-3.0222, -1.5625)
	$\sin(4\pi t/24)$	-0.6283	(-1.208, -0.1105)	-0.7920	(-1.2791, -0.3899)
	$\cos(2\pi t/24)$	-0.5795	(-1.3496, 0.1055)	-0.8624	(-1.514, -0.361)
	$\cos(4\pi t/24)$	0.1104	(-0.3065, 0.4727)	0.5756	(0.2502, 0.9174)
Average Travel Time		-0.0132	(-0.0276, -0.0019)	-0.0119	(-0.0177, -0.0055)
Travel Time Variance		-0.0005	(-0.0021, 0)	-0.0006	(-0.0014, 0)
Travel Cost		-0.1338	(-0.2296, -0.0187)	-0.1362	(-0.2303, -0.044)

Table 6.3: BVMNP Parameter Estimation Results for Covariance Parameters

Covariance Parameters	CAR Specification		AR1 Specification	
	Mean Estimate	95% Interval	Mean Estimate	95% Interval
Sigma (First Arrival Alt)	2.4683	(1.6973, 3.4494)	1.2665	(0.9355, 1.6774)
Sigma (Last Arrival Alt)	3.6735	(2.2952, 5.6019)	1.9052	(0.3248, 4.6787)
Sigma (First Return Alt)	4.0875	(2.7775, 5.4247)	3.1108	(2.2882, 3.9976)
Sigma (Last Return Alt)	1.0011	(0.5939, 1.6616)	1.8897	(0.2903, 5.2299)
Lambda (Arrival)	0.3157	(0.2173, 0.3939)	0.6039	(0.5581, 0.6403)
Lambda (Return)	0.3755	(0.31, 0.4298)	0.1499	(0.0841, 0.2117)
Lambda (Duration)	0.0926	(0.0564, 0.1381)	0.5566	(0.5224, 0.5982)
Lambda (First/Last Arr/Ret Alts)	1.4142	(0.2127, 3.4442)	0.4493	(0.0067, 1.319)
Lambda (First/Last Dur Alts)	0.3066	(0.0067, 0.9576)	0.0637	(0.0041, 0.214)
Gamma1 (Full)	8.6026	(8.0641, 9.0473)	9.5766	(9.1953, 10.0273)
Gamma1 (Part)	6.1580	(4.5539, 8.0324)	10.3693	(9.7848, 10.9077)
Gamma2 (Full)	0.0581	(0.0029, 0.1207)	-0.0416	(-0.0906, -0.0047)
Gamma2 (Part)	0.3972	(0.1834, 0.6228)	-0.1257	(-0.1902, -0.0555)
Observations	997		997	

Parameters relating to the covariance matrix (Table 6.3) do not seem as consistent across the two models, but these two sets of parameters do not have the exact same meaning (i.e., the correlation parameters, λ_a , λ_r , and λ_d , have different meanings across the two models). For the CAR specification, the largest correlations are estimated to be between return time alternatives, and the lowest are estimated for the duration component, while for the AR1 specification, the highest correlations are estimated for the arrival time component and the lowest for return time alternatives. Since it is expected that workers would be more constrained in their arrival times (due to normal work start times), the AR1 estimates (with high arrival time correlations) appear more reasonable. The low duration correlation parameter estimates of the CAR model are also rather unexpected. Of course, the duration component is also very much related to baseline durations implied by the estimates of the μ 's. With the CAR specification, baseline durations are predicted to start at about 9.0 hours for full-time workers and 6.3 hours for part-time workers. For both worker types, baseline durations are predicted to increase with increasing arrival time, which is counterintuitive and a possible explanation for the low duration-specific correlation parameter estimate of the CAR specification. It can be expected that as arrival time moves later in the day, the duration of the work activity becomes shorter,

since there are fewer hours left in the day. In fact, just such behavior is realized with the estimates of the AR1 specification, suggesting the AR1 specification to be more reasonable. Variances of boundary alternatives (i.e., $\sigma_{a,first}^2$, $\sigma_{a,last}^2$, $\sigma_{r,first}^2$, and $\sigma_{r,last}^2$) are estimated to be greater than one (the fixed variance of non-boundary alternatives) for both models, with CAR estimates generally greater than AR1 estimates. This result is certainly expected, since there is less information in the models for the boundary alternatives (e.g., assumed arrival/return times for these alternatives are approximations). In the following section, effects of individuals' attributes are examined.

6.3.1 Individual-Specific Covariate Effects

As shown in Table 6.1 and Table 6.2, the majority of individual-specific parameters are very similar between the two specifications, with only a couple of exceptions, which is not surprising since the only difference between the models is their covariance structures. To fully understand the effect of these variables, predictive density plots are presented for average individuals (i.e., covariates evaluated at the average over the sample) with the exception of one covariate, where the values of that covariate are varied to understand its effect on arrival and return time choices. Figure 6.4 and Figure 6.5 show arrival time predictive densities for each of the 8 individual-specific covariates of the models. Not surprisingly, effects of these covariates are quite similar across the two BVMNP models. Moreover, the effects are similar to the effects found for departure time profiles in Chapter 5's CCNL and continuous logit models. One notable difference is the effect of travel distance (Figure 6.5c). For continuous logit and CCNL models, those individuals with greater travel distances were predicted to depart earlier, all else being equal. For the BVMNP models, travel distance appears to have only minor effects on arrival time profiles. This is largely because the BVMNP models consider arrival times, rather than departure times, and workers would be expected to be more constrained in their arrival time choice due to regular work hours. Thus, for models of departure time, workers traveling longer distances must depart earlier in order to arrive on time. However, the commute distance plays little role in determining when they must arrive for work.

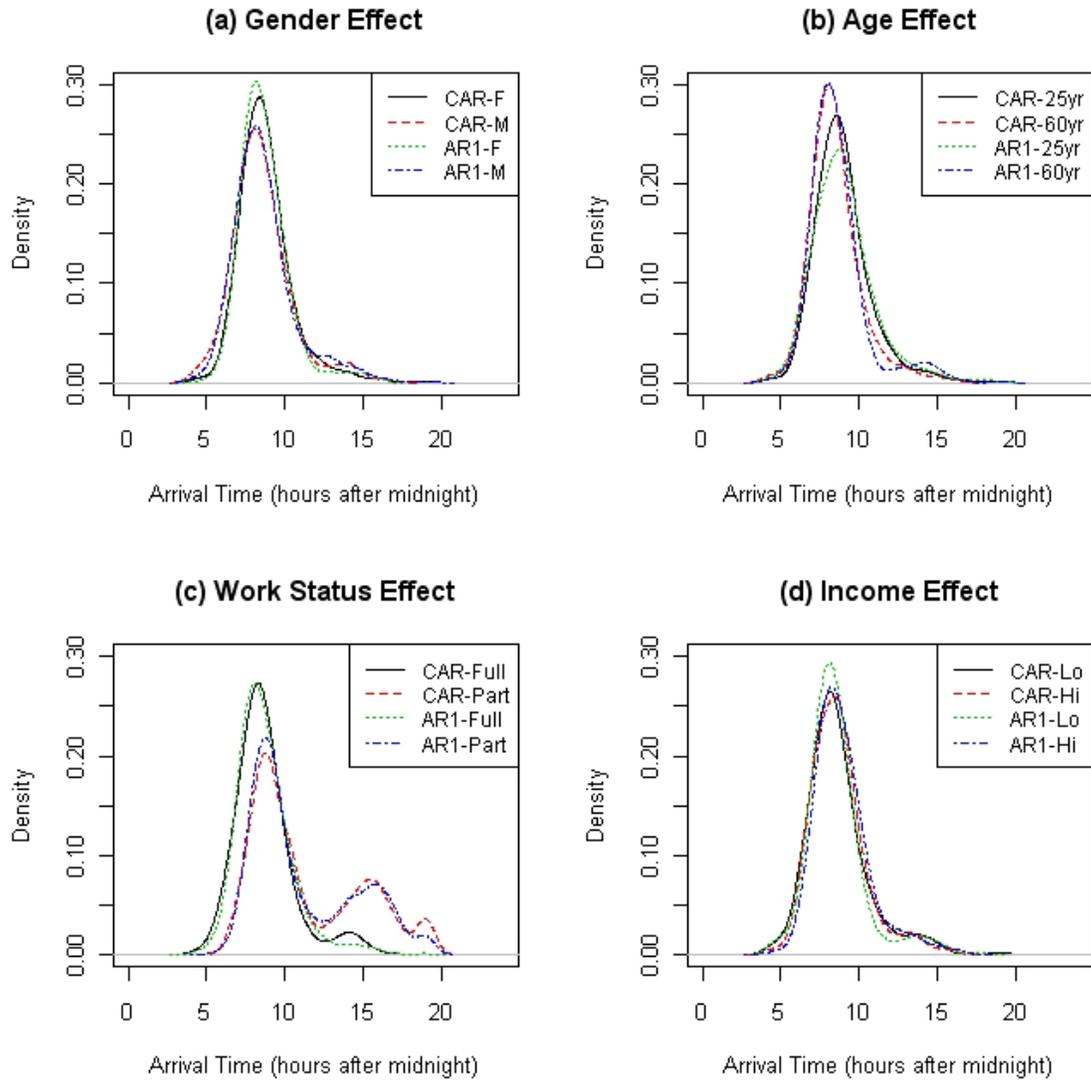


Figure 6.4: Gender, Age, Work Status, and Income Effects on Arrival Time Profiles for BVMNP Models

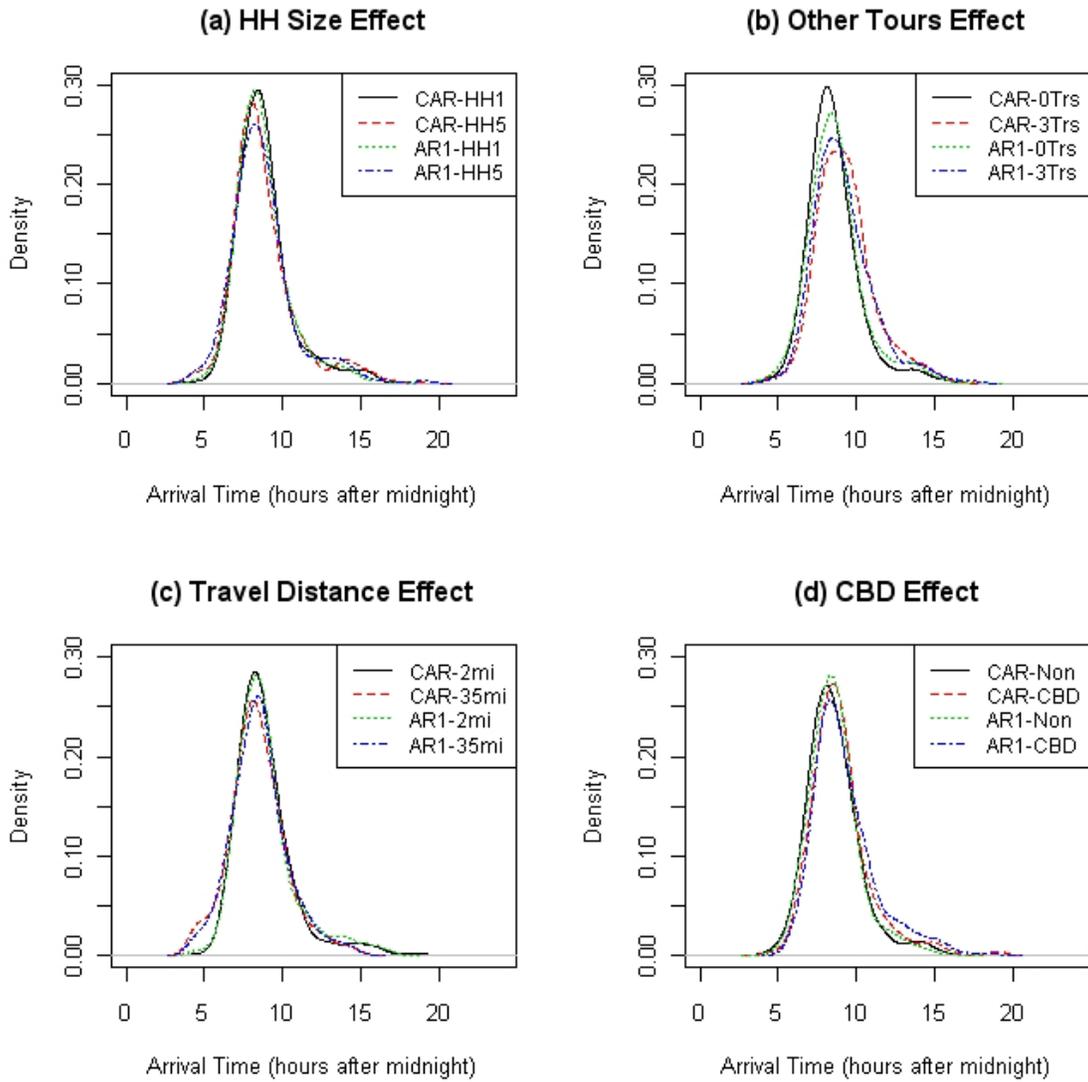


Figure 6.5: Household Size, Other Tours, Travel Distance, and CBD Effects on Arrival Time Profiles for BVMNP Models

In Figure 6.6 and Figure 6.7, return time predictive densities are plotted for each of the 8 individual-specific covariates of the model. While males are only predicted to return slightly later than females (Figure 6.6a), it is interesting to note that they are also predicted to arrive slightly earlier. Thus, work durations are slightly longer for males than females. Maybe men are more likely to work overtime than women, or females are more likely to have other responsibilities, such as dropping children off at and picking

them up from school. Older individuals are predicted to return earlier than younger ones (Figure 6.6b), which is not so surprising given that they are also predicted to arrive earlier. Interestingly, part-time workers' return time profiles (Figure 6.6c) seem to mimic their arrival time profiles (Figure 6.4c) with return times shifted to later hours, of course. This is very reasonable considering such workers may have very different work scheduling constraints, as compared to full-time workers.

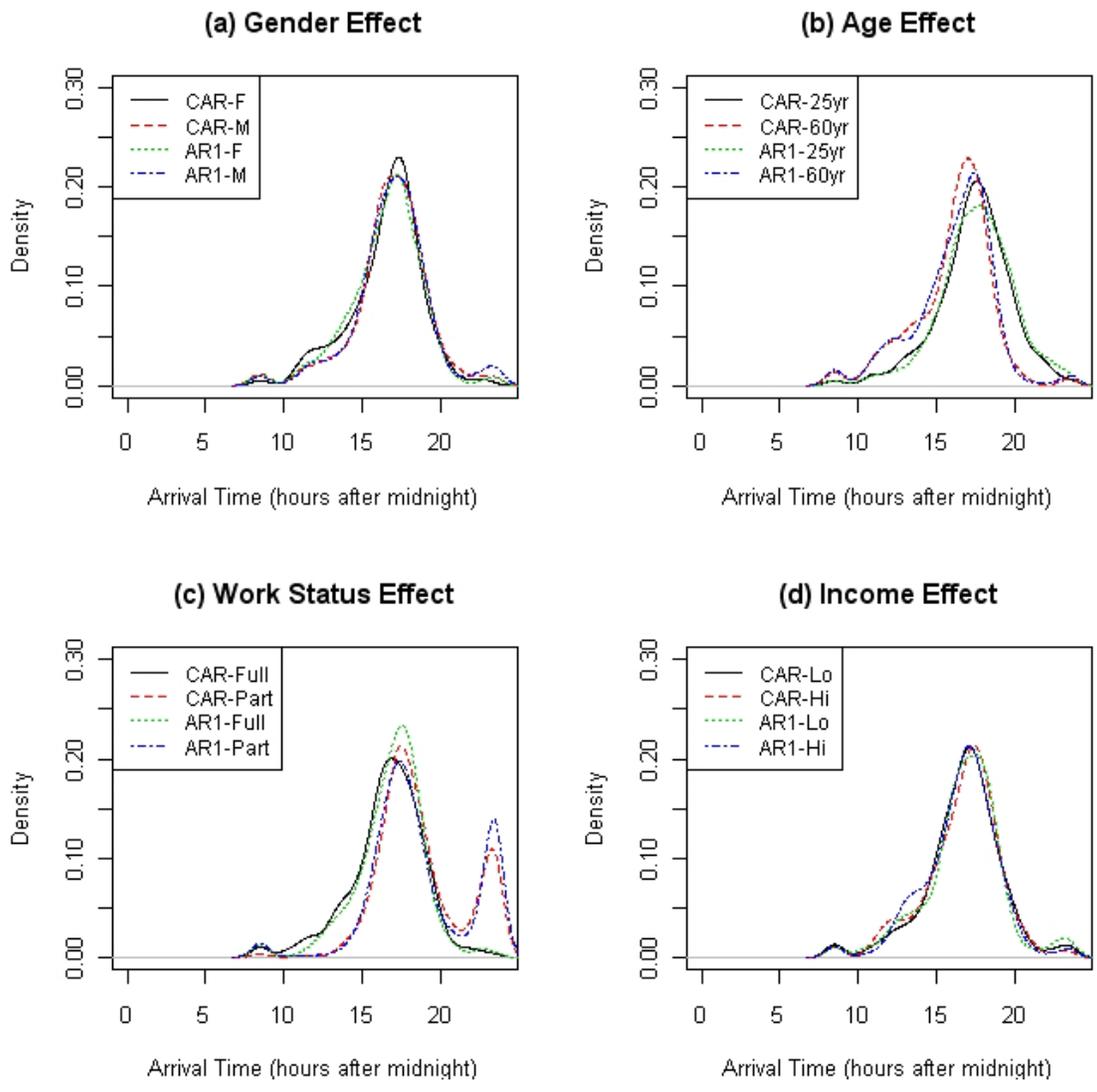


Figure 6.6: Gender, Age, Work Status, and Income Effects on Return Time Profiles for BVMNP Models

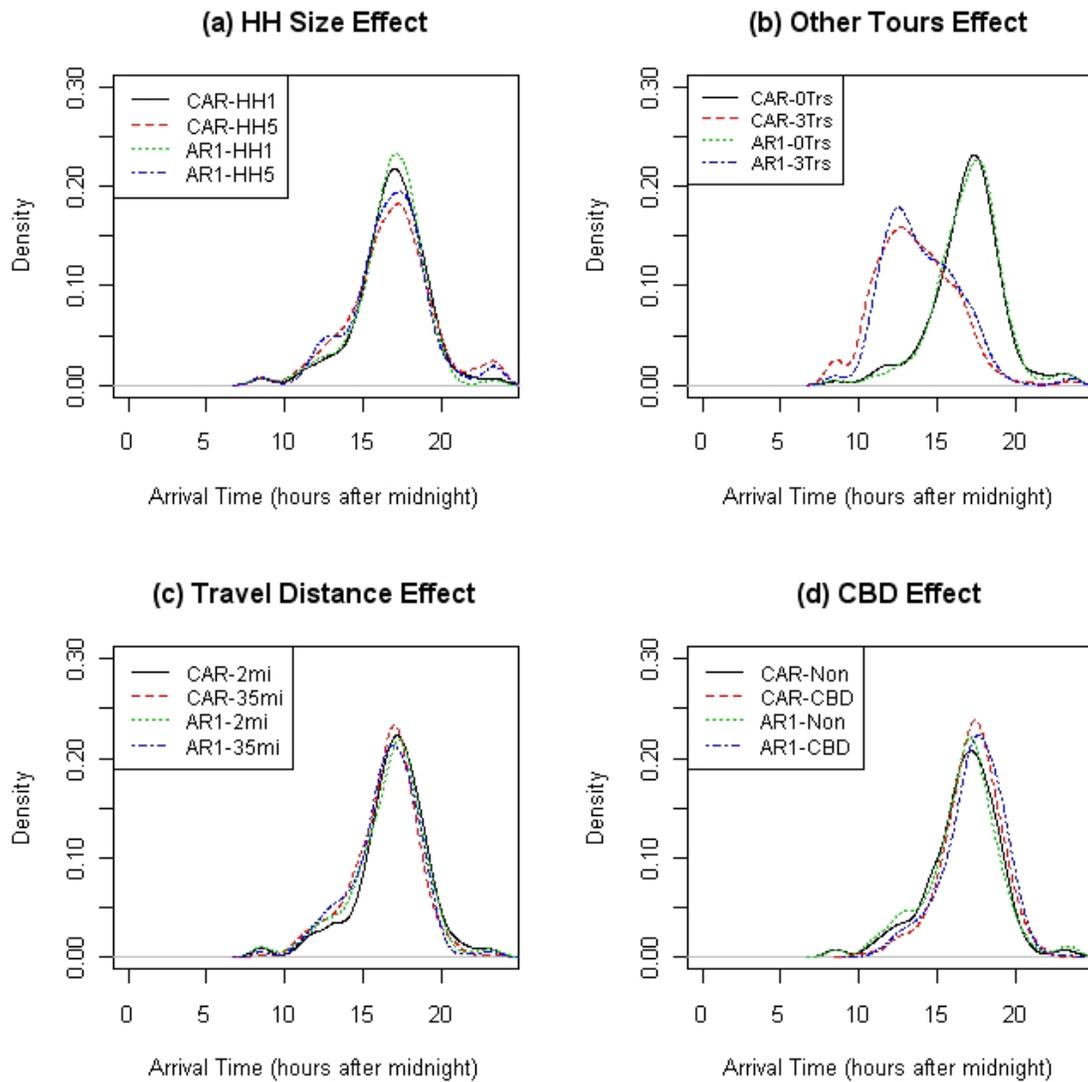


Figure 6.7: Household Size, Other Tours, Travel Distance, and CBD Effects on Return Time Profiles for BVMNP Models

Both income (Figure 6.4d and Figure 6.6d) and household size (Figure 6.5a and Figure 6.7a) appear to play only small roles in both arrival and return time profiles of workers. Of course, one would not necessarily expect such variables to have important consequences here. However, the presence of additional tours undertaken in the day for a worker has very important effects on a worker's return time (Figure 6.7b). In particular, such individuals are predicted to return from work much earlier as the number of such

additional tours increases. Since these workers obviously have other scheduling considerations for the day, this seems very reasonable. Like the effect of travel distance on arrival times, its effect on return times does not appear to be all that substantial (Figure 6.7c), as expected. Finally, the effect of traveling to work in the central business district (CBD) is to push return times later in the day (Figure 6.7d). As arrival times for such individuals also appear to be a bit later in the day than others (Figure 6.5d), this seems reasonable. Maybe those traveling to the CBD are more often in particular lines of work where there is a preference for slightly later work schedules. In summary, the effects of each of these covariates appear to be reasonable and consistent with prior expectations.

6.3.2 Out-of-Sample Predictive Performance

To better appreciate the predictive ability of the BVMNP models, out-of-sample prediction (using the same 20% sample as in Chapter 5) was performed for them along with a simple joint-choice multinomial logit (MNL) model. In the case of the MNL, each choice alternative represents the arrival and return time alternative pair, unlike the BVMNP models, which represent arrival times and return times as distinct choices. The MNL model was estimated using BIOGEME software and employed 50 randomly chosen alternatives from the set of all 621 joint choice alternatives along with the chosen alternative⁵⁵. Since the MNL model represents arrival and return time choice jointly, an additional utility component related to alternatives' implied durations and the duration taken to the power of two was included in the model, similar to the specification of Popuri et al. (2008). For consistency with the duration components of the BVMNP models, the sample was segmented by full-time workers with no additional tours for the day and all other workers (i.e., part-time workers and those with additional tours). Since the MNL model is estimated using classical techniques, the predictive likelihood is simply a fixed value. For the BVMNP models (like any MNP), predictive densities and

⁵⁵ McFadden (1978) shows that one can use a simple random sample of alternatives for MNL estimation and still obtain consistent parameter estimates.

likelihoods are difficult to compute due to open form likelihood expressions. Instead, using random parameter draws from the posterior, utilities are drawn from their corresponding distribution for each individual, taking the maximum utilities to be the predicted value. The probability of accurate prediction is then averaged over all individuals. Figure 6.8a shows aggregate out-of-sample predictive densities for arrival times under the BVMNP and MNL models, Figure 6.8b shows aggregate densities for return times, and Figure 6.8c shows the implied activity or tour duration profiles. In Figure 6.8d, the predictive accuracy of the MNL (a point estimate) is plotted against the distribution of predictive accuracies for the BVMNPs.

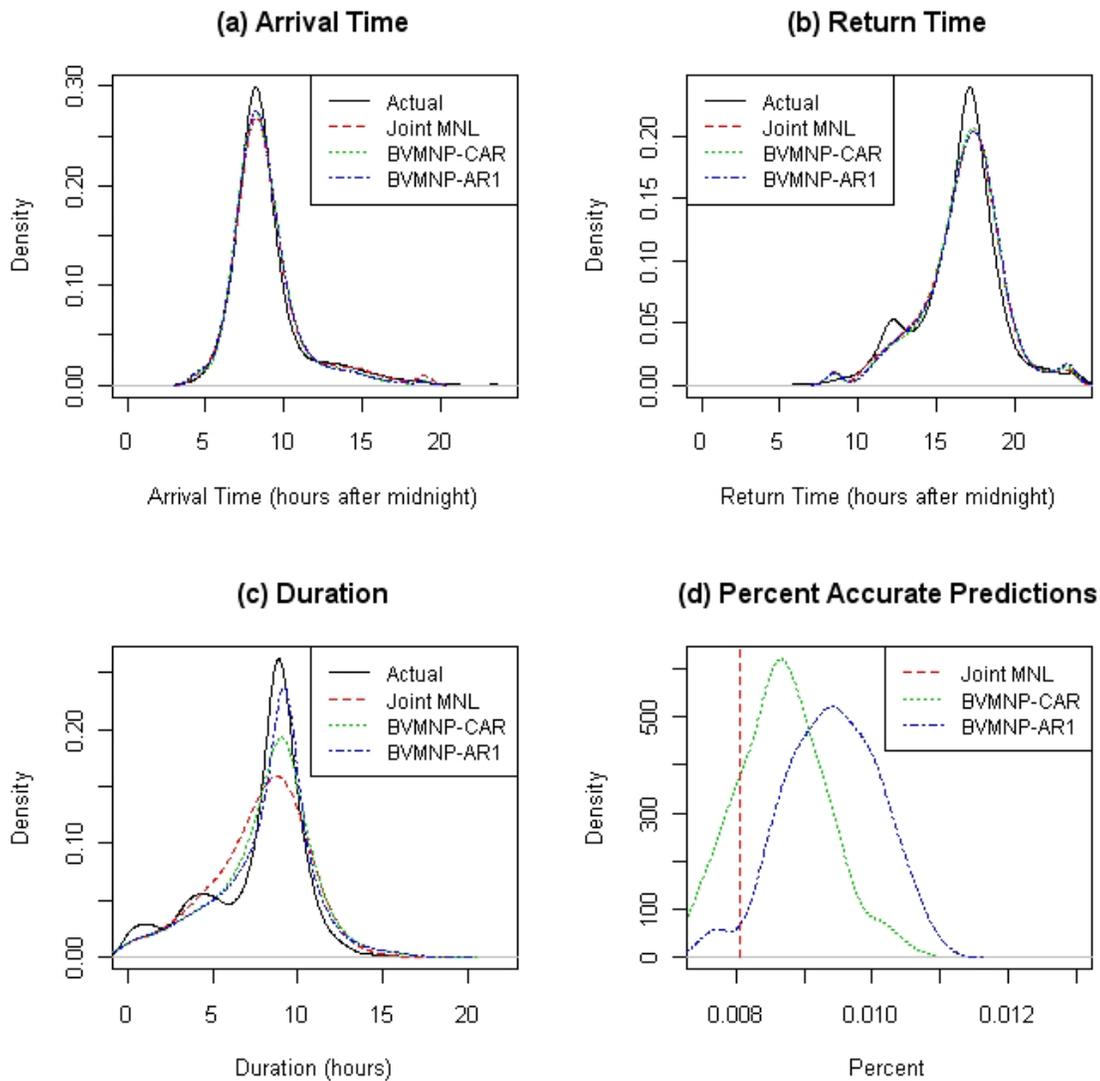


Figure 6.8: Out-of-Sample Predictions for MNL versus BVMNP Models

As shown in Figure 6.8a and Figure 6.8b, aggregate out-of-sample predictive distributions of arrival and return times are nearly identical for the three models and closely match estimated densities of the actual data. This is not surprising since each of the models has a number of utility components specifically related to arrival and return time choice. However, predictive duration profiles differ slightly between each (Figure 6.8c), with neither the MNL nor CAR specifications matching the actual duration profiles so well. This is likely because the duration-related components of each model are not as

rigorously specified. Nonetheless, it does appear the BVMNP models (particularly the AR1 specification) perform better here. Moreover, predictive accuracies of the BVMNP models are clearly superior to the MNL (Figure 6.8d). In fact, the CAR BVMNP specification beats the MNL about 81% of the time and the AR1 BVMNP specification beats the MNL about 97% of the time. Alternatively, one could measure model performance by the amount of error in predictions (i.e., how close predicted arrival and return times and durations are to actual ones). Table 6.4 shows mean squared errors (MSEs)⁵⁶ in predicted arrival times, return times, and durations for each of the models. According to Table 6.4, the MSE in arrivals is nearly the same for both BVMNP specifications, while it is more than 1 squared hour greater for the MNL. For returns, the MSE is very similar across the models, but for durations, the MNL out-performs both BVMNP specifications by a rather wide margin. This is a bit surprising considering the predictive duration profiles for the two BVMNP models appear to more closely match actual duration densities (Figure 6.8c) than does the MNL. Of course, since BVMNP parameter estimates do not necessarily represent posterior draws (due to non-convergence), it is difficult to draw firm conclusions.

Table 6.4: MSEs of MNL and BVMNP Model Predictions

Measure	MNL	BVMNP (CAR)	BVMNP (AR1)
Arrival Time MSEs (hrs ²)	10.3	9.1	9.1
Return Time MSEs (hrs ²)	12.2	12.2	12.5
Duration MSEs (hrs ²)	13.5	14.9	14.8
Total	36.0	36.2	36.3

Overall, it appears that both BVMNP specifications are superior in terms of model fit to the MNL, particularly for the AR1 model. In the following section, toll policy simulations are examined.

⁵⁶ The error for a single observation is computed as the difference between predicted time and actual (observed) time for that observation. MSE is the summation of squared errors over all observations, divided by the total number of observations.

6.3.3 Economic Welfare Demonstration

Similar to Chapter 5, an economic welfare exercise is demonstrated here. Consumer surplus (CS) for the MNL is computed using typical logsums. For the BVMNP models, CS can be estimated by recognizing that CS (in a random utility maximization model) is measured by expected maximum utility achieved over all alternatives. Since it is not difficult to obtain utility draws in the BVMNP context, one can simply obtain a number of utility draws from the posterior distribution of utilities for each individual, take the maximum utility for each draw, and average these over draws. The average for an individual then serves as an estimate of the expected maximum utility for that individual. Like in Chapter 5, CS can be divided by the cost coefficients from the model to obtain a measure in monetary units.

With the sample of 821 automobile mode tours used in model estimation⁵⁷, the same set of highly idealized scenarios from Chapter 5 is used again here. The first is the base (or status quo) scenario, where travelers are assumed to face travel conditions found in the actual data. In the first and second tolling policy simulation, \$0.15/mile tolls are assessed on all roads during peak periods (6 to 9 am and 3:30 to 6:30 pm), and in the last tolling simulation, \$0.30/mile tolls are assessed on all roads during peak periods. In the first and third tolling policy simulations, it is assumed the tolls reduce peak period travel time delay by 50%, while in the second simulation, tolls are assumed to reduce such delay by only 10%.

Since the MNL model was not estimated using Bayesian methods, CS changes represent point estimates, rather than the distributional estimates from the BVMNP models. For BVMNP models, CS is computed for each of 1000 random parameter draws from the respective posterior distributions. The same random parameter draws are used in each of the policy simulations.

⁵⁷ There are actually 997 tours in the estimation sample, but only 821 are tours where the automobile mode is chosen.

Figure 6.9a shows the distribution of CS change (measured as the difference in CS between tolling policy simulation and status quo simulation) for the three tolling policy simulations under the CAR model specification, and Figure 6.9b shows the distribution of CS change under the AR1 model specification. Under both model specifications, CS change under simulation 1 is estimated to be greatest, not surprisingly since its combination of tolls and delay reductions should offer the greatest value to travelers. In addition, both models predict the CS change under tolling policy simulation 3 to be about twice as negative as simulation 1, which makes good sense since delay reductions are identical for the two simulations, but tolls are twice as large in simulation 3. Based on the differences between simulations 1 and 2 under both model specifications, it appears that the peak period delay reduction does have a significant effect on CS change. Tolls under these simulations are identical, but peak travel delay is reduced by 40% more in simulation 1 as compared to simulation 2.

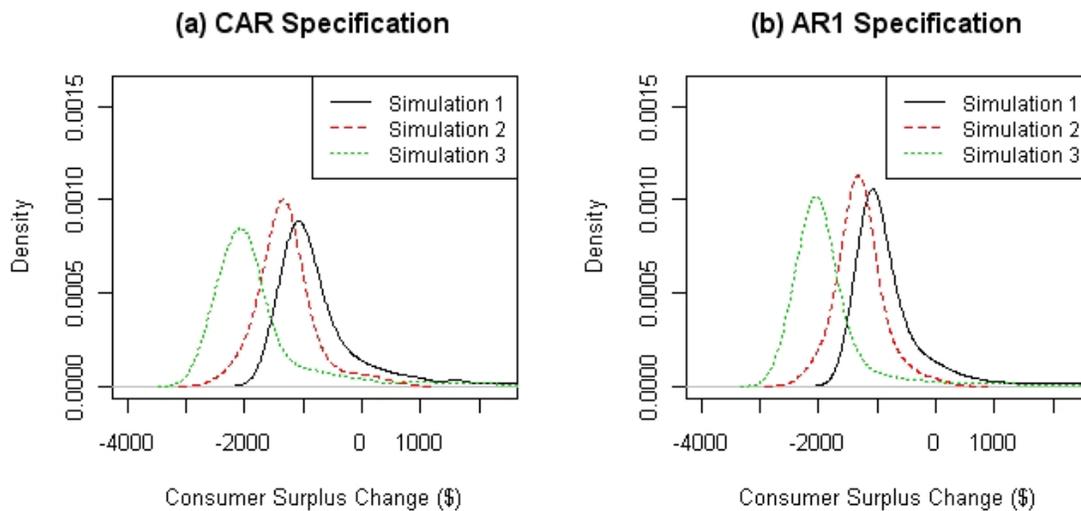


Figure 6.9: Consumer Surplus Change Distribution for CAR and AR1 Specifications under Three Tolling Policy Simulations

Figure 6.10 shows the same plots as Figure 6.9, but grouped by simulation. In addition, the MNL's point estimates of CS change are displayed in Figure 6.10. For each tolling policy simulation, the magnitude of CS change across the three models is similar, though in each case, MNL estimates are lower than both BVMNP model estimates. Mean CS change estimates per traveler were $-\$0.82$, $-\$0.94$, and $-\$1.40$ for simulation 1 under the CAR, AR1, and MNL models, respectively, $-\$1.57$, $-\$1.57$, and $-\$1.76$ for simulation 2, and $-\$2.18$, $-\$2.25$, and $-\$3.01$ for simulation 3. For each simulation, the mean CS change for the MNL model is much lower than those for the CAR and AR1 specifications. This could be because the BVMNP models recognize the similarities between alternatives near peak periods (but not within them) and peak periods (via error term correlations). In other words, peak shoulder periods may not be viewed as poorly by travelers under the BVMNP specifications compared to the MNL, where correlations do not exist. The distribution of CS change under the CAR specification is wider and less peaked than that of the AR1 specification. In fact, the standard deviation of CS change per traveler for the CAR specification ranges from $\$0.60$ to $\$1.15$, while CS change standard deviations of the AR1 specification are only in the range of $\$0.47$ to $\$0.89$. This is probably because the bounds on parameters related to network attribute variables in the CAR specification were estimated to be wider than in the AR1 specification. While CS change is very important, it is also important to understand the effects of these tolls on travelers' scheduling choices.

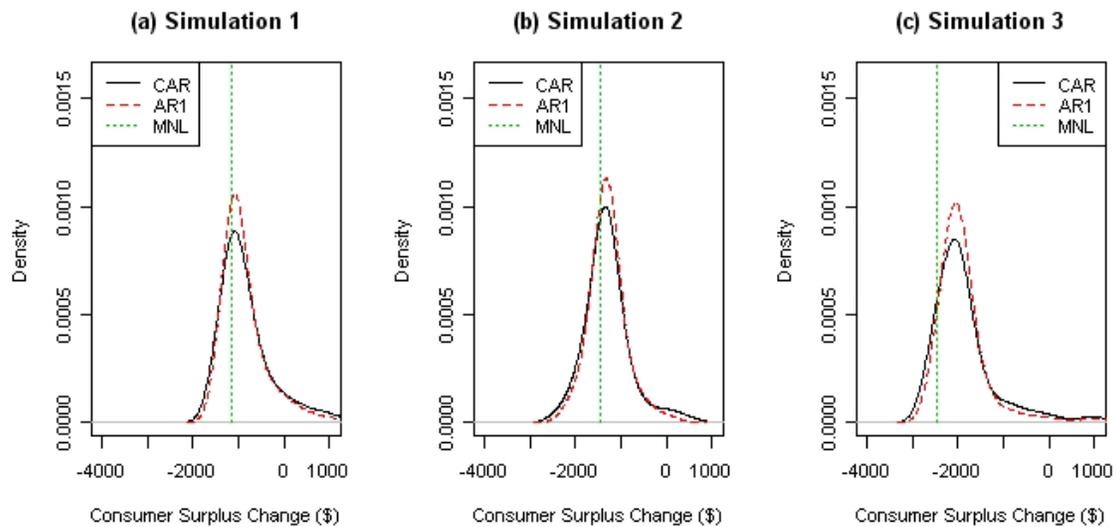


Figure 6.10: Consumer Surplus Change Distribution for Three Tolling Policy Simulations under CAR, AR1, and MNL Model Specifications

To appreciate the tolls' effects on traveler behavior, Table 6.5 presents the proportions of workers choosing arrival times during each of five TOD periods: before 5 am, 5 to 6 am, 6 to 9 am (the AM peak), 9 to 10 am, and after 10 am. For each scenario, predictions for the two BVMNP models are similar within each TOD period. And, while MNL predictions are fairly different in each scenario, similar changes in TOD shares emerge in comparison to both BVMNP models. One other interesting observation that can be made here is that the proportion of travelers within each TOD period is almost identical across tolling policies 1 and 2 for each model specification. The only difference between these two simulations is that peak period travel delay reduction is 40% greater for simulation 1, which affects the imputed average travel times and travel time variances. Thus, for a 40% decrease in peak period travel delay reduction (which is fairly substantial), there is almost no change in predicted arrival times, suggesting that the practical effects of travel time and its variance in the models may not be so large. Of course, this could also be a result of the imputation methods for these variables. Travel times and variances are not predicted to vary a tremendous amount even under status quo conditions, so perhaps

better methods for imputing these variables are needed to fully appreciate their effects on workers' TOD choices.

Table 6.5: Predicted Arrival Time Proportions for Three TOD Periods and Four Simulations under CAR, AR1, and MNL Model Specifications

	Proportion			Proportional Increase from Base		
	CAR	AR1	MNL	CAR	AR1	MNL
Status Quo						
Before 5 am	0.015	0.016	0.011	n/a	n/a	n/a
5 to 6 am	0.020	0.024	0.029	n/a	n/a	n/a
6 to 9 am	0.630	0.624	0.612	n/a	n/a	n/a
9 to 10 am	0.167	0.167	0.159	n/a	n/a	n/a
After 10 am	0.167	0.170	0.189	n/a	n/a	n/a
Simulation 1						
Before 5 am	0.017	0.017	0.013	0.106	0.097	0.146
5 to 6 am	0.024	0.029	0.032	0.191	0.211	0.098
6 to 9 am	0.589	0.586	0.579	-0.066	-0.061	-0.054
9 to 10 am	0.187	0.187	0.172	0.118	0.119	0.086
After 10 am	0.183	0.182	0.204	0.098	0.070	0.078
Simulation 2						
Before 5 am	0.018	0.018	0.013	0.167	0.133	0.186
5 to 6 am	0.025	0.029	0.033	0.224	0.233	0.123
6 to 9 am	0.589	0.586	0.579	-0.066	-0.061	-0.054
9 to 10 am	0.186	0.186	0.172	0.115	0.115	0.084
After 10 am	0.182	0.181	0.203	0.092	0.066	0.074
Simulation 3						
Before 5 am	0.019	0.019	0.014	0.222	0.201	0.308
5 to 6 am	0.028	0.034	0.036	0.390	0.415	0.235
6 to 9 am	0.551	0.551	0.547	-0.125	-0.118	-0.106
9 to 10 am	0.205	0.206	0.185	0.227	0.236	0.165
After 10 am	0.197	0.191	0.217	0.178	0.123	0.150

Table 6.5 also reports the proportional change in TOD shares from the status quo simulation for each TOD period. This is computed as the difference in the proportions for a TOD period from status quo scenario to tolling policy, divided by the share predicted under the status quo scenario for that TOD period. One would expect peak

shoulder periods to enjoy a relatively large increase in shares, as compared to off-peak periods (i.e., before 5 am and after 10 am). Under the MNL specification, the opposite result emerges in comparing the “Before 5 am” period to the “5 to 6 am” period, while both MNP specifications exhibit share changes more in line with expectations. Of course, this is due to the correlations offered under the MNP specifications. A similar result is found when comparing the “After 10 am” period to the “9 to 10 am” period, where the MNL model predicts small share increases during the shoulder period relative to predictions of the MNP models under each tolling simulation.

Figure 6.11 further details these share changes by illustrating the share of travelers shifting from the AM peak period to combined shoulder periods (i.e., “5 to 6 am” and “9 to 10 am”) and to combined off-peak periods (i.e., “before 5 am” and “after 10 am”). In other words, of those workers shifting away from the AM peak in the face of tolls, Figure 6.11 shows the share of those workers shifting to the shoulder periods and off-peak periods. Again, one would expect larger shifts toward shoulder periods than off-peak periods, since the shoulder periods are more similar to the AM peak than off-peak periods. The MNL specification predicts approximately equal shoulder and off-peak period shifts for each simulation, while under both BVMNP models, peak-period travelers are predicted to shift more toward shoulder periods, with the greatest shoulder period shifts under simulation 3, not surprisingly. In addition, the AR1 specification predicts the largest shoulder period shifts, which also is not surprising, since the AR1 correlation parameters specific to arrival time alternatives were estimated to be so high.

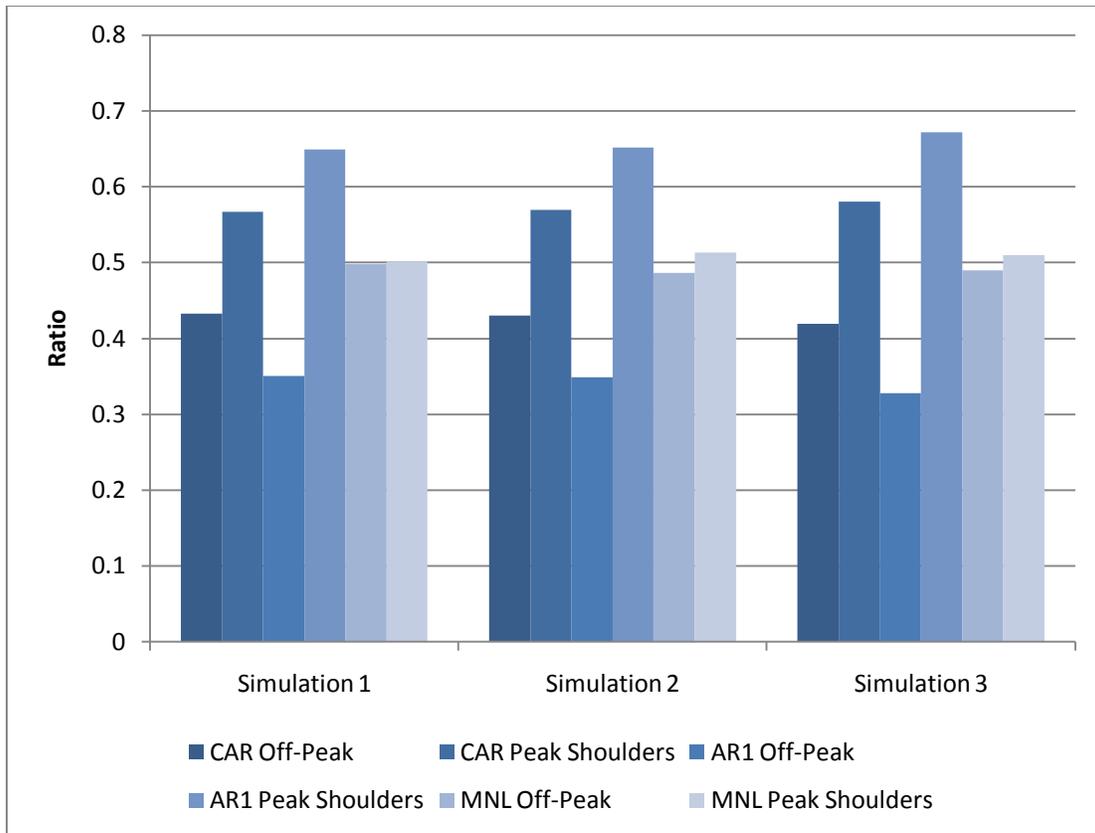


Figure 6.11: AM Peak Period Arrival Time Shifts to Shoulder and Off-Peak Periods for Three Tolling Policy Simulations under CAR, AR1, and MNL Models

Table 6.6 shows aggregate TOD period prediction proportions for return time choices under the four simulations and three models. Like arrival time choices, all three models predict similar return time choices for the status quo scenario. However, the AR1 specification predicts much fewer PM peak period travelers under each tolling policy simulation than either of the other two models, while the MNL predicts the most peak-period travelers. This is somewhat strange since travel cost parameters across the three models are similar. This could be resulting from the AR1's strong duration specific correlation parameter, which is translating shifts in arrival times to return time choices. Similar to the results for arrival times, return times are predicted to be quite similar under each model for simulations 1 and 2 (i.e., comparing simulation 1 results to simulation 2

results for each model). Again, this is an indication that average travel times and travel time variances may have little practical significance on return time choice.

Table 6.6: Predicted Return Time Proportions for Three TOD Periods and Four Simulations under CAR, AR1, and MNL Model Specifications

Status Quo	Proportion			Proportional Increase from Base		
	CAR	AR1	MNL	CAR	AR1	MNL
Before 2:30 pm	0.165	0.172	0.183	n/a	n/a	n/a
2:30 to 3:30 pm	0.085	0.084	0.089	n/a	n/a	n/a
3:30 to 6:30 pm	0.548	0.538	0.531	n/a	n/a	n/a
6:30 to 7:30 pm	0.108	0.110	0.104	n/a	n/a	n/a
After 7:30 pm	0.094	0.096	0.093	n/a	n/a	n/a
Simulation 1						
Before 2:30 pm	0.182	0.199	0.191	0.104	0.158	0.044
2:30 to 3:30 pm	0.103	0.105	0.096	0.205	0.245	0.076
3:30 to 6:30 pm	0.483	0.451	0.498	-0.119	-0.161	-0.061
6:30 to 7:30 pm	0.127	0.133	0.113	0.178	0.209	0.082
After 7:30 pm	0.105	0.112	0.102	0.119	0.167	0.098
Simulation 2						
Before 2:30 pm	0.184	0.201	0.192	0.112	0.172	0.052
2:30 to 3:30 pm	0.102	0.104	0.096	0.193	0.228	0.074
3:30 to 6:30 pm	0.479	0.446	0.496	-0.125	-0.171	-0.066
6:30 to 7:30 pm	0.128	0.134	0.114	0.188	0.221	0.091
After 7:30 pm	0.107	0.115	0.102	0.139	0.197	0.099
Simulation 3						
Before 2:30 pm	0.197	0.221	0.199	0.190	0.289	0.090
2:30 to 3:30 pm	0.118	0.120	0.102	0.379	0.427	0.145
3:30 to 6:30 pm	0.423	0.379	0.466	-0.227	-0.295	-0.123
6:30 to 7:30 pm	0.147	0.152	0.122	0.358	0.381	0.173
After 7:30 pm	0.116	0.127	0.111	0.233	0.327	0.189

Like Table 6.5, Table 6.6 also shows the proportional change in TOD shares from the status quo simulation for each TOD period. These results are further detailed in Figure 6.12, which shows the share of PM peak period travelers shifting to the combined shoulder periods (i.e., “2:30 to 3:30 pm” and “6:30 to 7:30 pm”) and the combined off-

peak periods (i.e., “before 2:30 pm” and “after 7:30 pm”). The MNL model does not predict very large shifts to shoulder periods, as compared to off-peak periods. Similar results emerge for the AR1 specification (though shifts to shoulder periods are larger than those of the MNL), which is likely due to the rather low estimate of the model’s return correlation parameter, but could also be resulting from the relatively high share reductions during peak periods. Since workers may be more affected by work duration under this model, shifts in arrival times could be playing a large role here. Under the CAR specification, one sees the largest share increases during the peak shoulders. Of course, it is not surprising that the CAR specification predicts the largest share of travelers to shift to peak shoulders, since the return-specific correlation parameter estimate of the CAR specification is relatively high.

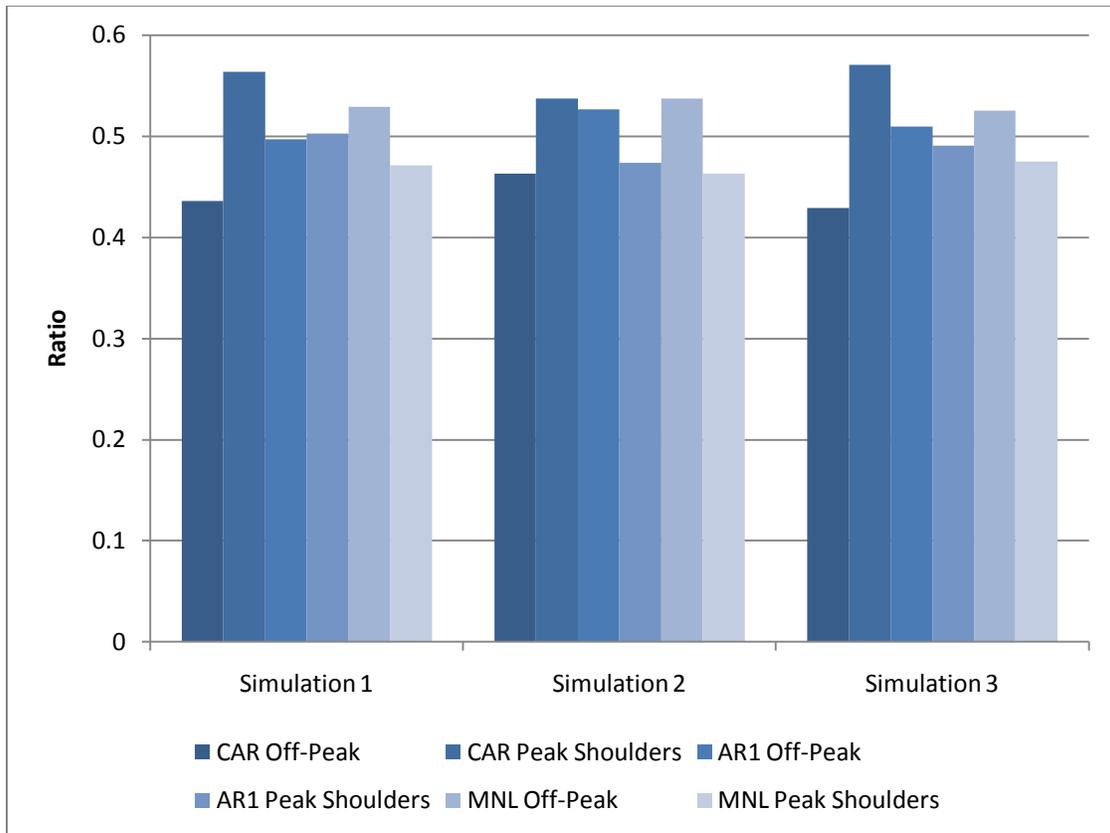


Figure 6.12: PM Peak Period Return Time Shifts to Shoulder and Off-Peak Periods for Three Tolling Policy Simulations under CAR, AR1, and MNL Models

6.4 Chapter Summary

This chapter presents the empirical results of the BVMNP models. All of the utility-specific parameter estimates appear reasonable and generally consistent across the two models, as well as with Chapter 5's continuous choice model results. Despite the differences in covariance structures between the two BVMNP models (CAR and AR1), their predictive abilities appear very similar. While not completely unexpected, it seems strange that the CAR and AR1 specifications could have such different covariance parameter estimates and yet produce similar predictions. Nonetheless, it is clear that both BVMNP specifications offer added flexibility over a simple MNL model (since they can capture correlations across arrival and return time alternatives), and predictive

performance of the BVMNPs appear to be at least as good as the MNL, and possibly much better – depending on the measure of performance used. If one considers that the BVMNP model estimates here are not yet converged, it does not seem unreasonable to say that the BVMNP models offer added model performance capabilities relative to the MNL. In addition, there are numerous opportunities for extending the BVMNP to capture even more variability in the data through refinement of the covariance structure.

Results of three tolling policy simulations demonstrate how consumer surplus can be estimated for the BVMNP models. They also reveal how BVMNP arrival and return correlations are realized, while highlighting the independence of irrelevant alternatives assumption of the MNL. The following chapter summarizes Chapter 5's and Chapter 6's findings, and offers a review of the models' specifications, their benefits and limitations, and some potential extensions.

CHAPTER 7: CONCLUSION

7.1 Summary

The need for better methods of modeling travelers' time-of-day (TOD) choices is clear. A key weakness of existing *continuous* models of departure time and duration is that they do not offer a defensibly or econometrically rigorous connection to microeconomic theories of behavior (e.g., Wang 1996, Bhat and Steed 2002, Komma and Srinivasan 2008, and Gadda et al. 2009). Current travel demand model systems rely heavily on random utility maximization for other travel choices (such as destination and mode), often integrating such choices in a behaviorally consistent fashion. Moreover, utility models offer a basis for calculating consumer surplus change, which is useful for policy and project evaluation (including, for example, environmental justice concerns). In addition to behavioral and welfare considerations, existing continuous methods do not seem capable of consistently incorporating multivariate features of tour timing.

Of course, existing *discrete* methods also suffer from a number of weaknesses. For instance, the only discrete choice models offering a two-dimensional choice framework across a 24-hour period with relatively small time intervals fail to allow for correlations across alternatives close in time (e.g., Ettema and Timmermans 2003, Vovsha and Bradley 2004, Abou Zeid et al. 2006, and Popuri et al. 2008). Certainly, one cannot expect such alternatives to exhibit independence in their error terms. Discrete choice models that do account for such correlations fail to recognize the two-dimensional choice context of the tour, do not consider the full 24-hour day period, and/or consider rather coarse TOD alternatives (e.g., Small 1987, Chin 1990, Bhat 1998a, Steed and Bhat 2000, and de Jong et al. 2003).

This dissertation established two new methods (the CCNL and the BVMNP) for examining travelers' timing decisions, and turned to Bayesian methods for these new models' estimation and application. This work first generalized the continuous logit to

release its independence of irrelevant alternatives assumption. The continuous cross-nested logit (CCNL) model develops random utility theory in a continuous choice context, while offering the ability to capture correlations across alternatives close in time (or space or some other dimension). Of course, the model was estimated for only one timing dimension here, neglecting the return time dimension of a travel tour. While the model specification can be formulated to accommodate two timing dimensions without great difficulty, estimating the model in two dimensions presently appears computationally prohibitive.

Parameter estimates of this new model (using San Francisco Bay Area data) were generally consistent with expectations and with continuous logit parameter estimates. For instance, results suggest that males, older individuals, and those from households with many members are more inclined to depart earlier in the day, all else being equal, while high income individuals, part-time workers, and those working in the central business district are more likely to depart later in the day. Not surprisingly, those with longer travel distances are found to depart earlier, presumably to arrive at work on time. Those with additional tours during the day are found to depart significantly later, probably because these individuals have more scheduling constraints than others. While values of travel time (VOTTs) were estimated to be quite low (median values of \$1.43 and \$2.28 per hour for continuous logit and CCNL models, respectively), this is probably at least partially due to the higher (and more practically important) estimated value of reliability (VOR), at least in the case of the CCNL. It could also be that workers do not trade off time and money in their departure time decisions (like they do for mode or route choice decisions), for various reasons (ignorance or other scheduling considerations) as noted earlier.

While empirical results of the CCNL suggest that predictive densities for specific individuals can appear very similar between it and the continuous logit, the CCNL performs better than the continuous logit model in terms of out-of-sample prediction of

departure times, while allowing more flexible choice behavior to emerge. Consumer surplus changes were estimated for several simple road pricing scenarios and found to be very similar in both models, though departure time predictions exhibited some differences. The CCNL's predictions appear more reasonable in the simulation examples, since they concentrate time-of-day shifts to peak-period shoulders, just before and after times where tolls were employed. Of course, the CCNL model is more computationally burdensome to estimate. Here, generating draws from the CCNL's posterior distribution took on the order of 30 times longer than those of the continuous logit. Moreover, the numerical integration procedure for generating likelihood values suffered from more error in the CCNL context than in the continuous logit setting. Reducing such error is likely to result in longer estimation times, though the benefits of the model may be more apparent.

Like the CCNL, the bivariate multinomial probit (BVMNP) model is built on random utility theory and allows for correlations across alternatives. Moreover, it incorporates two dimensions of choice (outbound and return times, in the case of trip/tour scheduling). While the choice context is discrete, the level of temporal resolution used here (30 minutes) is fairly reasonable, especially considering the rounding error prevalent in reported departure and return times (Stopher et al. 2008). In addition, no existing bi-dimensional tour timing models have recognized any sort of correlation across alternatives (see, e.g., Ettema and Timmermans 2003, Vovsha and Bradley 2004, Abou Zeid et al. 2006, and Popuri et al. 2008).

Two BVMNP covariance structures were formulated and empirical results offer similar inference regarding variables' effects. In addition, variables' effects on arrival time choice were found to be consistent with effects on departure time choice for continuous logit and CCNL models. The only exception here was the effect of travel distance (as expected), which was estimated to have very little effect on arrival times. Return time choice was found to be most influenced by an individual's age (with older individuals

returning earlier in the day), worker status (with part-time workers returning later in the day), and the number of other travel tours undertaken for the individual (with those having other travel tours departing much earlier, on average), each of which seems reasonable. Like the continuous models of departure time, VOTTs for the home-to-work journey were estimated to be quite low, but VOTTs for the work-to-home journey were estimated to be much higher (and more reasonable). VORs were estimated to be much higher for the home-to-work journey as compared to the work-to-home journey, which seems reasonable, since the cost of late arrival at work should be higher than late arrival returning home.

Empirical evidence suggests that the predictive ability of the two BVMNP models is better than a relatively straightforward multinomial logit (MNL) model of all paired timing choices jointly. In addition, the BVMNP models offer more reasonable scheduling predictions under various tolling policy simulations. While the MNL predicted very small changes in the number of travelers choosing peak periods for the work-to-home journey, both BVMNP models predicted much more substantial peak-period travel reductions. In addition, the BVMNP models consistently predicted relatively large share increases for peak-period shoulders (consistent with expectations), whereas MNL predictions varied greatly over the three simulations, often predicting much larger share increases for very different times of day (rather than peak shoulders). Unfortunately, the BVMNP parameter estimates did not converge, which highlights the only real weakness of Bayesian methods and casts some doubt on the validity of results.

In addition to models of travelers' scheduling choices, this dissertation also developed a new method for imputing travel time variance by TOD. The method built on existing methods for imputing average travel time by TOD using regression models. These models allowed a measure of reliability to be incorporated in the scheduling choice models. While empirical results suggest that reliability (as well as travel time) may have

little practical significance on timing decisions, it was found to be at least as important as average travel time.

The models developed here were estimated using Bayesian techniques. Such methods seem particularly advantageous in estimating the BVMNP model, where conditional posterior distributions were derived for latent utility variables, covariance components, and utility function parameters separately. Such methods allow these utility variables and parameters to be drawn from standard distributions. While covariance components could not be drawn from standard distributions, and standard methods for drawing utility variables were not used due to computational instability, Metropolis-Hastings (MH) algorithms were developed using standard Bayesian methods, allowing for relative ease in drawing these parameters. R code was developed to draw from each of these models' posterior distributions. In addition, Bayesian methods provide great flexibility in model specification, and other (possibly more appropriate) specifications can be examined with relative ease.

Of course, complex-model estimation is not the only advantage of Bayesian methods. As noted in Chapter 1, one key area of concern in TOD modeling is the ability to capture heterogeneity across travelers, particularly variations in value-of-time (Vovsha et al. 2005). While this was not sought here, incorporating such heterogeneity in a Bayesian framework is not difficult, and can be accomplished through hierarchical modeling. On the other hand, classical methods must rely on maximum simulated likelihood estimation (MSLE) methods to estimate such models. In this context, Huber and Train (2001) point out that MSLE techniques can have difficulty locating the global maximum (particularly when many local maxima exist), computational difficulties can arise with MSLE if the dimension of the random parameters is large, and statistical identification issues are more likely under MSLE (since that method does not allow for inclusion of prior information). In addition, Bayesian estimation provides draws from the multivariate posterior distribution of all parameters. In terms of risk and uncertainty analysis (which is

particularly important for toll road analyses), posterior draws offer a natural setting for capturing such uncertainty in demand modeling and other systems.⁵⁸ The scenario analyses provided in this dissertation illustrate how uncertainty passes through the model into welfare estimates.

7.2 Opportunities for Future Research

This research sought methods to capture correlation structures inherent across time-of-day choice alternatives, while reflecting random utility theory in one- and two-dimensional continuous choice contexts. Although these goals were generally achieved (in separate model settings), several opportunities exist for model enhancements. This section discusses some of the limitations and extensions of this dissertation's models.

One limitation of this analysis stems from the regression models used to impute average travel times and travel time variances continuously over time. Clearly, it would be preferable to have such time-varying network data available, rather than imputing it. While traffic monitoring equipment (such as loop detectors, roadside receivers, or cellular towers) may be able to provide such information in some areas on some roadways, it is unlikely one could obtain such data for all roadways in a region. Of course, other methods for imputing such variables may be preferred, including dynamic network simulations (assuming good network data and reasonable demand profiles by TOD are available). This would indeed be favored over the methods used here, since network simulations need not rely on traveler estimates. Given the limitations of the data available for this dissertation research, the methods used to obtain network variables appear reasonable. However, the travel time variance computed may be, to some extent, reflecting the degree of misreporting and rounding error contained in the data, rather than actual, day to day, recurring or non-recurring travel time variability. This is certainly a concern and could be inflating travel time variance estimates to some extent. On the

⁵⁸ Capturing uncertainty in model inputs is also important for risk analyses (see, e.g., Lemp and Kockelman 2009).

other hand, the models control for day-of-week and other variables, which many travelers do not perceive, so uncertainty estimates may be lower than those perceived by travelers.

As is typical practice, the model specifications pursued in this dissertation do not recognize heterogeneity across preference parameters and values of travel time (and its reliability), though the Bayesian methods applied here quickly lend themselves to such added flexibility. More controlled data sets (i.e., those with better covariate estimates) also would be useful for evaluating such heterogeneity. (As noted above, the methods for imputing average travel times and travel time variances were not ideal, and the cost variable only varied by TOD period.) In the case of the CCNL, mean estimates for the parameters relating to these variables were negative, but with large 95% intervals that contained zero. In the case of the BVMNP, these variables were restricted to be negative through prior specification. Intuitively, these variables should have negative effects on utility, but such results did not emerge definitively. Several factors may be at work here. For instance, the methods for imputing these variables are imperfect, and may be biased or incorrect, typically resulting in greater uncertainty in parameter estimates. Alternatively, the models may not be controlling for enough other variables and the time-varying variables may be playing the role of proxy for some uncontrolled variables. Maybe travelers do not have much flexibility in their work start times, which leads them to be unaffected by travel times and variances during other times of day. Of course, most travelers probably do not have a precise understanding of how travel times vary over the day. In order to satisfactorily assess heterogeneity, highly controlled data sets are really desired.

The models estimated in this dissertation are computationally burdensome for most computers (model estimation can take days to obtain satisfactory convergence) and so necessitated the use of a relatively small sample in model estimation (997 observations, or 6% of the original sample). While working with smaller samples may not create

biases in parameter estimation, it certainly leads to wider distributions for parameter estimates. Longer run times with larger samples would address this issue.

One also can imagine a number of ways to formulate the utility function used in the CCNL and BVMNP models. The formulation used here was chosen for its familiarity, for the fact that it offers continuity in utility profiles across days, and because it allows for multi-modality in TOD choice. However, other functional forms (such as quadratic interactions [rather than the cyclical interactions used in this dissertation] or linear shift variables) may allow better models to emerge.

A number of opportunities also exist for the BVMNP model. For instance, the covariance structures of alternative-specific utilities are very specific. One major reason for choosing the AR1 correlation structure is to allow easy interpretation of the covariance parameters. On the other hand, the main reason for choosing the CAR correlation structure is to offer a clear relationship between utility error terms. However, any number of correlation structures could be imagined. For instance, one could turn to an AR2-type covariance specification rather than AR1 or a simultaneous autoregressive (SAR) structure rather than CAR. Maybe most critical are the parameters relating to correlations between arrival and return time alternatives. These correlations are controlled by just three parameters in both the CAR and AR1 specifications here: one directly related to correlations and two related to some baseline durations for which correlations should be highest. Since the current utility specifications do not recognize duration explicitly, the baseline duration parameters are of particular importance. Here, these parameters only vary across two groups of individuals (full-time workers with no additional travel tours and part-time workers and/or those with other tours). Moreover, there may be a number of factors that influence each worker's chosen activity duration. One may allow for these baseline durations to be a function of individual-specific variables, though this will lead to distinct covariance matrices for each individual, adding

to computational burden in model estimation. Of course, there may be other ways to formulate covariance parameters, as functions of individual-specific variables.

Other model extensions also exist. The set of individual-specific variables used in the models could be expanded to include any number of other effects (e.g., number of household vehicles, presence and number of children in the household [and their school start and end times], whether the individual has flexible work hours, occupation, and additional origin- and destination-specific variables). The CCNL could be extended to accommodate the two-dimensional timing choice of a travel tour, though this would be computationally burdensome. The time interval size of alternatives in the BVMNP model could be reduced, but again, this would add computational difficulty. Obviously, many opportunities for extending such work exist. There are many applications that await such random utility and random profit-maximizing specifications in a continuous choice context, and different data environments will apply.

7.3 Concluding Remarks

As activity-based travel demand modeling and DTA techniques advance, TOD modeling remains a key weakness of model systems. Moving toward continuous-time models (such as the continuous logit or CCNL), or at least toward discrete choice methods that recognize rather small time intervals and allow for thoughtful correlations across proximate alternatives, will enhance the temporal resolution of such models. Almost no continuous models have a solid behavioral basis, while the continuous logit and CCNL provide direct measures of utility and allow for correlations in unobserved heterogeneity. All these attributes are important for prediction, project evaluation, policy analysis, and welfare calculations, and can be used to link TOD models to other travel choice dimensions (such as destination and mode). Of course, the CCNL is not just limited to the context of TOD choice. A number of other transportation-related choices can best be handled in a continuous choice setting (e.g., location or destination choice and vehicle

usage). And countless opportunities lie outside transportation, in modeling human response, firm choices, and biological processes, among others.

Until now, no two-dimensional TOD discrete choice models had incorporated correlations across nearby alternatives (adjacent or otherwise). Clearly, one would not expect independence across such similar alternatives. Thus, in both models developed here, a much needed element has been added to the set of existing models of travelers' timing decisions.

In addition, this dissertation has illustrated how one can impute reliability measures for use in activity scheduling models. While the methods of imputing these measures are imperfect, they appear to provide reasonable estimates. Moreover, the effect of these variables was found to be as or more practically significant than average travel time, which demonstrates how important travel reliability can be.

The work undertaken here represents a meaningful step in the wide field of behavioral modeling. Superior methods for handling individuals' continuous choices are needed, and CCNL and BVMNP models offer two promising techniques. However, more research and experimentation is needed to fully appreciate their relative merits and limitations.

APPENDIX A: R CODE FOR CCNL MODEL ESTIMATION

This appendix details R code used for the estimation of the continuous cross-nested logit model. This part of the code represents only the main estimation module. Code not presented here includes those modules for reading data, generating results figures, and obtaining simulation results.

```
#mfact used in factoring covariance matrix of parameters
mfact = 0.2

#s is the length between discrete times where likelihood evaluated
s = 0.05

ttot = seq(0,24,s)
dim11 = length(ttot)
simpfac2 = matrix(rep(48,length(ttot)),ncol=1)
simpfac2[1] = 17
simpfac2[length(ttot)] = 17
simpfac2[2] = 59
simpfac2[length(ttot)-1] = 59
simpfac2[3] = 43
simpfac2[length(ttot)-2] = 43
simpfac2[4] = 49
simpfac2[length(ttot)-3] = 49

G = matrix(rep(0,length(ttot)*count),ncol=length(ttot))
Grow = matrix(rep(0,length(ttot)*count),ncol=length(ttot))
GG = matrix(rep(0,count),ncol=1)
F = matrix(rep(0,count),ncol=1)

#code is set to run for sets of 5000 iterations (i.e., k)
#j represents the number of 5000 iteration sets to run

for(j in 1:20){

for(k in 1:5000){

#*****
```

```

# get proposal parameter values
#*****
L = try(chol(Sigma_prop1),TRUE)
if(length(L) == 1){
  L=sechol(Sigma_prop1)
}
L = t(L)

#ensure valid h & rho values
#hprop is h proposal
#rowprop is rho proposal
#index500 is the size of groups in main computations (memory problems with large h
and whole sample)
#ncovtot is the total number of parameters
test_index = 0
while(test_index==0){
  tmp1 = rnorm(ncovtot,0,1)
  dim(tmp1) = c(ncovtot,1)
  hpropt = h + L[ncovtot-1,]%*%tmp1
  hpropt = s*round(hpropt/s)
  hprop = hpropt[1,1]
  rowpropt = row + L[ncovtot,]%*%tmp1
  rowprop = rowpropt[1,1]
  if(hprop <= 0.8) {
    index500 = count
  } else if(hprop <= 1.7) {
    index500 = 500
  } else if(hprop <= 2.5) {
    index500 = 334
  } else if(hprop <= 3.4) {
    index500 = 250
  } else if(hprop <= 4.3) {
    index500 = 200
  } else if(hprop <= 5.3) {
    index500 = 170
  } else if(hprop <= 6.4) {
    index500 = 145
  } else if(hprop <= 7.4) {
    index500 = 125
  } else {index500 = 115}
  if(hprop>=0.25 & rowprop>=1 & hprop<=8.5 & rowprop<=10.0) test_index = 1
}

```

```

#individual-specific interaction term parameters
#theta is current parameter vector
#ncovar is number of individual-specific variables
thetaprop3 = theta + L[1:(ncovtot-2),]%%tmp1
for(i in 1:ncovar){
  if(i == 1){
    id2 = 1
  } else {
    id2 = sum(id1[1:(i-1)]) + 1
  }
  id3 = sum(id1[1:i])
  temp5 = thetaprop3[id2:id3]
  if(id1[i] == 8){
    thetaprop[i,1:8] = temp5
  } else if(id1[i] == 6){
    thetaprop[i,1:3] = temp5[1:3]
    thetaprop[i,5:7] = temp5[4:6]
  } else if(id1[i] == 4){
    thetaprop[i,1:2] = temp5[1:2]
    thetaprop[i,5:6] = temp5[3:4]
  }
}

#network variables parameters
thetaprop2 = rep(0,3)
for(i in 1:3){
  thetaprop2[i] = (-1)*exp(thetaprop3[ncovtot-5+i])
}

#constant added to utility functions to prevent overflow
util_const = -20

#*****
# get new proposal covariance matrix (Sigma_prop1)
#*****
if(j >= 2){
  if(k == 5000){
    count1_id = 0
    count2_id = theta_store[1,1]
    for(ii in 2:5000){
      if(count2_id == theta_store[ii,1]){
        www1 = 0
      } else {

```

```

        count1_id = count1_id + 1
        count2_id = theta_store[ii,1]
    }
}
acceptrate = count1_id / 4999
if(acceptrate >= 0) {mfact = 0.5*mfact + 0.5*mfact / (0.25/acceptrate)} else { mfact =
0.25}
}
if(floor(k/20)==k/20){
  for(ii in 1:ncovtot){
    for(jj in 1:ncovtot){
      Sigma_prop1[ii,jj] = mfact*cov(theta_store[,ii],theta_store[,jj])
    }
  }
}
} else if(k >= 1000){
  if(floor(k/20)==k/20){
    for(ii in 1:ncovtot){
      for(jj in 1:ncovtot){
        Sigma_prop1[ii,jj] = mfact*cov(theta_store[1:k,ii],theta_store[1:k,jj])
      }
    }
  }
}
}
}
}

```

```

#*****

```

```

# set up arrays based on new parameter values of h & row

```

```

#*****

```

```

dim1 = floor(24/s + 1.01)

```

```

dim2 = floor(2*hprop/s + 1.01)

```

```

t1 = rep(0,dim1*dim2)

```

```

dim(t1) = c(dim1,dim2)

```

```

t2 = matrix(rep(0,dim1*dim2),ncol=dim1*dim2)

```

```

t3 = matrix(rep(0,dim1*dim2),ncol=dim1*dim2)

```

```

trev = seq((0-hprop),(24+hprop),s)

```

```

dummy4 = matrix(rep(1,length(trev)),ncol=length(trev))

```

```

dummy1 = matrix(rep(1,dim1*dim2),ncol=dim1*dim2)

```

```

dummy2a = matrix(rep(1,index500),ncol=1)

```

```

#tv_cov is network variables

```

```

#i_unavail is a dummy network variable for unavailable alternatives (for transit only)

```

```

#cyc_fxn is the collection of cyclical functions interacted with ind-specific vars

```

```

tv_cov2 = rep(0,(length(trev))*count*3)

```

```

dim(tv_cov2) = c(count,length(trev),3)
i_unavail2 = rep(1,(length(trev))*count)
dim(i_unavail2) = c(count,length(trev))
cyc_fxn2 = matrix(rep(0,length(trev)*8),ncol=length(trev))
startindex = 1 + floor(hprop/s + 0.01)
endindex = length(trev) - floor(hprop/s + 0.01)
tv_cov2[,startindex:endindex,] = tv_cov
i_unavail2[,startindex:endindex] = i_unavail
cyc_fxn2[,startindex:endindex] = cyc_fxn

startindex2 = dim11 - startindex + 2
endindex2 = length(trev) - endindex
tv_cov2[,1:(startindex-1),] = tv_cov[,startindex2:dim11,]
tv_cov2[, (endindex+1):(length(trev)),] = tv_cov[,1:endindex2,]
i_unavail2[,1:(startindex-1)] = i_unavail[,startindex2:dim11]
i_unavail2[, (endindex+1):(length(trev))] = i_unavail[,1:endindex2]
cyc_fxn2[,1:(startindex-1)] = cyc_fxn[,startindex2:dim11]
cyc_fxn2[, (endindex+1):(length(trev))] = cyc_fxn[,1:endindex2]

for(i in 1:dim2){
  t1[,i] = ttot - hprop + (i-1)*s
  for(ii in 1:dim1){
    index9 = (ii-1)*dim2 + i
    t3[index9] = t1[ii,i]
    t2[index9] = ttot[ii]
  }
}
simpfac1 = matrix(rep(48,dim2),ncol=1)
simpfac1[1] = 17
simpfac1[dim2] = 17
simpfac1[2] = 59
simpfac1[dim2-1] = 59
simpfac1[3] = 43
simpfac1[dim2-2] = 43
simpfac1[4] = 49
simpfac1[dim2-3] = 49

g = matrix(rep(0,dim1*dim2*index500),ncol=(dim1*dim2))
g2 = matrix(rep(0,dim2*count),ncol=dim2)
f = matrix(rep(0,dim2*count),ncol=dim2)

#####
# Main computations section

```

```

#####

# compute alpha
height1 = 1 / hprop
alpha = (hprop - abs(t3 - t2) + 0.0000000000001) / (hprop*hprop)

# compute utilities and g (a count x dim1*dim2 matrix)
v_g = exp(rowprop*((util_const)-
20000*i_unavail2+thetaprop2[1]*tv_cov2[,1]+thetaprop2[2]*tv_cov2[,2]+thetaprop2[3
]*tv_cov2[,3]+X%*%thetaprop%*%cyc_fxn2))

temp3 = round(Y/s) + 1 - (dim2 - 1) / 2
index11 = ceiling(count/index500 - 0.0001)
for(i in 1:index11){
  # get correct indices for individuals
  startt = (i-1)*index500+1
  if(i == index11){
    endv = count - (index11-1)*index500
    endt = count
    dummy2a = matrix(rep(1,endv),ncol=1)
  } else {
    endv = index500
    endt = i*index500
  }

  for(ii in 1:dim1){
    index8 = (ii-1)*dim2 + 1
    index9 = ii*dim2
    index19 = dim2 + ii - 1
    g[1:endv,index8:index9] = (dummy2a%*%((alpha[index8:index9])^rowprop)) *
(v_g[startt:endt,ii:index19])
  }

# locate and store alternative specific g elements for each individual
index7 = dim1*dim2 + 0.5
for(ii in startt:endt){
  index10 = ii - index500*floor((ii-0.1)/index500)
  for(jj in 1:dim2){
    index8 = jj - 1
    index9 = (temp3[ii] + index8)*dim2 - index8
    if(index9 >= 0.5 & index9 <= index7){
      g2[ii,jj] = g[index10,index9]
    } else if(index9 <= 0.5){

```

```

    index9 = index9 + dim1*dim2 - dim2
    g2[ii,jj] = g[index10,index9]
  } else if(index9 >= index7){
    index9 = index9 - dim1*dim2 + dim2
    g2[ii,jj] = g[index10,index9]
  } else {
    g2[ii,jj] = 0
  }
}
}
}

# integrate over nests
for(ii in 1:dim1){
  index8 = (ii-1)*dim2 + 1
  index9 = ii*dim2
  G[startt:endt,ii] = (s / 48) * g[1:endv,index8:index9] %%% simpfac1
}
}

# compute generating function
Grow = G^(1/rowprop)
GG = (s / 48) * (Grow %%% simpfac2)

#*****
# compute densities for individuals
temp1 = floor((Y - hprop)/s + 1.01)
temp2 = floor((Y + hprop)/s + 1.01)
temp4 = matrix(rep(1,count),ncol=1)
temp5 = matrix(rep(dim2,count),ncol=1)
Grow2 = g2[1,]
G22 = g2[1,]
for(i in 1:count){
  if(temp1[i] <= 0.5){
    index1 = 1 - temp1[i]
    Grow2[(1+index1):dim2] = Grow[i,1:temp2[i]]
    Grow2[1:index1] = Grow[i,(dim1-index1+1):dim1]
    G22[(1+index1):dim2] = G[i,1:temp2[i]]
    G22[1:index1] = G[i,(dim1-index1+1):dim1]
  } else if(temp2[i] >= (24/s+1.5)){
    index1 = temp2[i] - floor(24/s + 1.01)
    Grow2[1:(dim2-index1)] = Grow[i,temp1[i]:dim1]
    Grow2[(dim2-index1+1):dim2] = Grow[i,1:index1]
    G22[1:(dim2-index1)] = G[i,temp1[i]:dim1]
  }
}

```

```

    G22[(dim2-index1+1):dim2] = G[i,1:index1]
  } else {
    Grow2 = Grow[i,temp1[i]:temp2[i]]
    G22 = G[i,temp1[i]:temp2[i]]
  }
  f[i,] = (Grow2 / GG[i]) * (g2[i,] / G22)
  F[i] = (s / 48) * (f[i,] %*% simpfac1)
}

# compute log likelihood
ll_prop = log(F)
loglik_prop = sum(ll_prop)

#prior log likelihoods
prior_prop1 = (-1/2)*t(thetaprop3-thetabar)%*%(solve(Sigma_bar))%*%(thetaprop3-
thetabar)
prior_prop2 = log(dgamma(hprop-0.5,eta_h[1],eta_h[2]))
prior_prop3 = log(dgamma(rowprop-1.0,eta_row[1],eta_row[2]))
loglik_prop2 = loglik_prop
loglik_prop = loglik_prop + prior_prop1 + prior_prop2 + prior_prop3

#MH proposal acceptance step
ratio1 = loglik_prop[1] - loglik[1]
alpha1 = try(if(ratio1 >= 0) {1} else {exp(ratio1)}, TRUE)
if(is(alpha1,"numeric")){
  temp11 = runif(1,0,1)
  if(temp11 <= alpha1){
    loglik = loglik_prop
    loglik2 = loglik_prop2
    theta = thetaprop3
    h = hprop
    row = rowprop
  }
}

#Store new theta values for iteration k
tmp1 = ncovtot-2
for(i in 1:tmp1){
  theta_store[k,i] = theta[i]
}
theta_store[k,tmp1+1] = h
theta_store[k,tmp1+2] = row
loglik_store[k] = loglik2

```

}

APPENDIX B: R CODE FOR BVMNP MODEL ESTIMATION

This appendix details R code used for the estimation of the bivariate multinomial probit (BVMNP) model. This part of the code represents only the main estimation module. Code not presented here includes those modules for reading data, generating results figures, and obtaining simulation results. In addition, only the code used for the AR1 specification is detailed here, though it is almost identical to that used for estimation of the CAR specification.

```
mfact = 0.2
factf = 0.05
factp = 0.05

for(j in 141:160){

for(k in 1:5000){

#*****
#*****
# Step 1: Draw Utilities
#*****
#*****

utils2 = utils
nkern_prop = rep(0,count)
nkern_act = rep(0,count)

Lfull = t(chol(factf*Sigmafull))
Lpart = t(chol(factp*Sigmapart))

delta1 = utils - V
delta2 = delta1

nkern_act[1:nfull] = diag(delta1[1:nfull,]%*%Sigmafulli%*%t(delta1[1:nfull,]))
nkern_act[(nfull+1):count] =
diag(delta1[(nfull+1):count,]%*%Sigmaparti%*%t(delta1[(nfull+1):count,]))

for(i in 1:count){
```

```

if(i <= nfull){
  L11 = Lfull
  Sigmai11 = Sigmafulli
} else {
  L11 = Lpart
  Sigmai11 = Sigmaparti
}
check11 = 1
while(check11 == 1){
  rand11 = rnorm(dim1,0,1)
  utils2[i,] = utils[i,] + L11%%rand11
  maxua = max(utils2[i,1:dim2])
  maxur = max(utils2[i,(dim2+1):dim1])
  actua = utils2[i,Y2[i,1]]
  actur = utils2[i,Y2[i,2]]
  if(maxua == actua & maxur == actur){
    check11 = 0
    delta2[i,] = utils2[i,] - V[i,]
    nkern_prop[i] = (t(delta2[i,]))%%Sigmai11%%delta2[i,]
  }
}
ratio11 = (-0.5)*(nkern_prop[i] - nkern_act[i])
alpha11 = try(if(ratio11 >= 0) {1} else {exp(ratio11)}, TRUE)
if(is(alpha11,"numeric")){
  temp11 = runif(1,0,1)
  if(temp11 <= alpha11){
    utils[i,] = utils2[i,]
    utilmax[i,1] = maxua
    utilmax[i,2] = maxur
  }
}
}
}

#####
#####
# Step 2: Draw Elements of Covariance Matrices
#####
#####

#####
# Step 2.1: Obtain satisfactory draw for covariance matrix elements (one that
#           results in Sigma's being pos def, parameters positive, & rhos < 1)

```

```

#####
L = try(chol(Sigma_prop1),TRUE)
if(length(L) == 1){
  L=sechol(Sigma_prop1)
}
L = t(L)
cov_param_prop = cov_param
check = 1
while(check == 1){
  rnd_tmp1 = rnorm(ncovsig-2,0,1)
  cov_param_prop[3:ncovsig] = cov_param[3:ncovsig] + L%%rnd_tmp1

# Are parameters postive? If not, redraw.
if(min(cov_param_prop[1:(ncovsig-4)]) < 0){
  check = 1
} else {
  sig2a_prop = cov_param_prop[1]
  sig2r_prop = cov_param_prop[2]
  sig2a1_prop = cov_param_prop[3]
  sig2a2_prop = cov_param_prop[4]
  sig2r1_prop = cov_param_prop[5]
  sig2r2_prop = cov_param_prop[6]
  rhoa_prop = cov_param_prop[7]
  rhor_prop = cov_param_prop[8]
  rhod_prop = cov_param_prop[9]
  rhoar2_prop = cov_param_prop[10]
  rhod2_prop = cov_param_prop[11]
  gfull1_prop = cov_param_prop[12]
  gpart1_prop = cov_param_prop[13]
  gfull2_prop = cov_param_prop[14]
  gpart2_prop = cov_param_prop[15]
  for(i in 1:dim2){
    if(i >= 2 & i <= dim2){
      Wa_prop[i,i] = 1
      Wr_prop[i,i] = 1
    } else if(i == 1){
      Wa_prop[i,i] = sig2a1_prop
      Wr_prop[i,i] = sig2r1_prop
    } else {
      Wa_prop[i,i] = sig2a2_prop
      Wr_prop[i,i] = sig2r2_prop
    }
  }
}
}

```

```

for(i in 1:dim2){
  for(ii in 1:dim2){
    if(i == ii){
      www = 1
    } else {
      tmp1 = (abs(ta[i]-ta[ii]))
      tmp2 = (abs(tr[i]-tr[ii]))
      tmp3 = ((ta3[i] / s)^((-1)*rhoar2_prop)) * ((ta3[ii] / s)^((-1)*rhoar2_prop))
      tmp4 = ((tr3[i] / s)^((-1)*rhoar2_prop)) * ((tr3[ii] / s)^((-1)*rhoar2_prop))
      tmp5 = sqrt(Wa_prop[i,i] * Wa_prop[ii,ii])
      tmp6 = sqrt(Wr_prop[i,i] * Wr_prop[ii,ii])
      Wa_prop[i,ii] = tmp5 * tmp3 * (rhoa_prop^tmp1)
      Wr_prop[i,ii] = tmp6 * tmp4 * (rhor_prop^tmp2)
    }

    tmp1 = (abs(tr[ii]-ta[i]-(gfull1_prop + gfull2_prop * ta[i])) + 1)
    tmp2 = (abs(tr[ii]-ta[i]-(gpart1_prop + gpart2_prop * ta[i])) + 1)
    tmp3 = ((ta3[i] / s)^((-1)*rhod2_prop)) * ((tr3[ii] / s)^((-1)*rhod2_prop))
    tmp5 = sqrt(Wa_prop[i,i] * Wr_prop[ii,ii])

    if(ta2[i] <= tr2[ii]){
      Caf_prop[i,ii] = tmp5 * tmp3 * (rhod_prop^tmp1)
      Cap_prop[i,ii] = tmp5 * tmp3 * (rhod_prop^tmp2)
      Crf_prop[ii,i] = tmp5 * tmp3 * (rhod_prop^tmp1)
      Crp_prop[ii,i] = tmp5 * tmp3 * (rhod_prop^tmp2)
    } else {
      Caf_prop[i,ii] = 0
      Cap_prop[i,ii] = 0
      Crf_prop[ii,i] = 0
      Crp_prop[ii,i] = 0
    }
  }
}

```

```

#Proposed covariance matrices for full & part
Sigmaf_prop[1:dim2,1:dim2] = Wa_prop
Sigmaf_prop[1:dim2,(dim2+1):dim1] = Caf_prop
Sigmaf_prop[(dim2+1):dim1,1:dim2] = Crf_prop
Sigmaf_prop[(dim2+1):dim1,(dim2+1):dim1] = Wr_prop
Sigmap_prop[1:dim2,1:dim2] = Wa_prop
Sigmap_prop[1:dim2,(dim2+1):dim1] = Cap_prop
Sigmap_prop[(dim2+1):dim1,1:dim2] = Crp_prop
Sigmap_prop[(dim2+1):dim1,(dim2+1):dim1] = Wr_prop

```

```

tmpf1 = eigen(Sigmaf_prop,only.values=TRUE)
tmpp1 = eigen(Sigmap_prop,only.values=TRUE)
if(is.complex(tmpf1$values)){
  tmpf2 = -10
} else {
  tmpf2 = min(tmpf1$values)
}
if(is.complex(tmpp1$values)){
  tmpp2 = -10
} else {
  tmpp2 = min(tmpp1$values)
}

# Are Sigmaf_prop and Sigmap_prop positive definite? If not, redraw.
if(min(tmpf2,tmpp2) < 0){
  check = 1
} else {
  check = 0
  Sigmaf_propi = solve(Sigmaf_prop)
  Sigmap_propi = solve(Sigmap_prop)
}
}
}

#if we get to here, each parameter is positive and the corresponding covariance
# matrix is positive definite

#update Sigma_prop1
if(j >= 2){
  if(floor(k/1000) == k/1000){
    count1_id = 0
    count2_id = theta_store[k-999,ncovtot2]
    for(ii in (k-998):(k-1)){
      if(count2_id == theta_store[ii,ncovtot2]){
        www1 = 0
      } else {
        count1_id = count1_id + 1
        count2_id = theta_store[ii,ncovtot2]
      }
    }
  }
  acceptrate = count1_id / 998
  if(acceptrate >= 0) {mfact = 0.5*mfact + 0.5*mfact / (0.3/acceptrate)} else {mfact =
0.3}

```

```

}
if(floor(k/5)==k/5){
  for(ii in 1:(ncovsig-2)){
    for(jj in 1:(ncovsig-2)){
      Sigma_prop1[ii,jj] =
mfact*cov(theta_store[,ii+ncovtot+2],theta_store[,jj+ncovtot+2])
    }
  }
}
} else if(k >= 2000){
  if(floor(k/5)==k/5){
    for(ii in 1:(ncovsig-2)){
      for(jj in 1:(ncovsig-2)){
        Sigma_prop1[ii,jj] =
mfact*cov(theta_store[1:k,ii+ncovtot+2],theta_store[1:k,jj+ncovtot+2])
      }
    }
  }
}

#####
# Step 2.2: Compute log posteriors under proposal and current
#####
delta1 = utils - V

# for proposal
nkernf1 = sum(diag(delta1[1:nfull,]%**Sigmaf_propi**t(delta1[1:nfull,])))
nkernp1 =
sum(diag(delta1[(nfull+1):count,]%**Sigmap_propi**t(delta1[(nfull+1):count,])))
log_nkern1 = (-0.5)*(nkernf1 + nkernp1)
nkernf2 = (det(Sigmaf_propi))
nkernp2 = (det(Sigmap_propi))
log_nkern2 = (0.5*nfull)*log(nkernf2) + (0.5*npart)*log(nkernp2)
ll_prop = log_nkern1 + log_nkern2
prior_prop1 = log(dgamma(sig2a_prop,eta_sig[1],eta_sig[2])) +
log(dgamma(sig2r_prop,eta_sig[1],eta_sig[2])) +
log(dgamma(sig2a1_prop,eta_sig[1],eta_sig[2])) +
log(dgamma(sig2a2_prop,eta_sig[1],eta_sig[2])) +
log(dgamma(sig2r1_prop,eta_sig[1],eta_sig[2])) +
log(dgamma(sig2r2_prop,eta_sig[1],eta_sig[2]))
prior_prop2 = log(dbeta(rhoa_prop,eta_rho[1],eta_rho[2])) +
log(dbeta(rhor_prop,eta_rho[1],eta_rho[2])) +
log(dbeta(rhod_prop,eta_rho[1],eta_rho[2])) +

```

```

log(dgamma(rhoar2_prop,eta_rho2[1],eta_rho2[2])) +
log(dgamma(rhod2_prop,eta_rho2[1],eta_rho2[2]))
  prior_prop3 = log(dnorm(gfull1_prop,gf_par[1],gf_par[2])) +
log(dnorm(gpart1_prop,gp_par[1],gp_par[2])) +
log(dnorm(gfull2_prop,g2_par[1],g2_par[2])) +
log(dnorm(gpart2_prop,g2_par[1],g2_par[2]))
  lpost_prop = ll_prop + prior_prop1 + prior_prop2 + prior_prop3

#for current
  nkernf1 = sum(diag(delta1[1:nfull,]%*%Sigmafulli%*%t(delta1[1:nfull,])))
  nkernp1 =
sum(diag(delta1[(nfull+1):count,]%*%Sigmaparti%*%t(delta1[(nfull+1):count,])))
  log_nkern1 = (-0.5)*(nkernf1 + nkernp1)
  nkernf2 = (det(Sigmafulli))
  nkernp2 = (det(Sigmaparti))
  log_nkern2 = (0.5*nfull)*log(nkernf2) + (0.5*npart)*log(nkernp2)
  ll_cur = log_nkern1 + log_nkern2
  lpost_cur = ll_cur + prior1 + prior2 + prior3

#*****
# Step 2.3: Accept or reject proposal
#*****
ratio1 = lpost_prop[1] - lpost_cur[1]
alpha1 = try(if(ratio1 >= 0) {1} else {exp(ratio1)}, TRUE)
if(is(alpha1,"numeric")){
  temp11 = runif(1,0,1)
  if(temp11 <= alpha1){
    #if proposal accepted, set values to proposal values
    lpost_cur = lpost_prop
    cov_param = cov_param_prop
    sigma2a = sig2a_prop
    sigma2r = sig2r_prop
    sigma2a1 = sig2a1_prop
    sigma2r1 = sig2r1_prop
    sigma2a2 = sig2a2_prop
    sigma2r2 = sig2r2_prop
    rhoa = rhoa_prop
    rhor = rhor_prop
    rhod = rhod_prop
    rhoar2 = rhoar2_prop
    rhod2 = rhod2_prop
    gammafull1 = gfull1_prop
    gammapart1 = gpart1_prop

```

```

gammafull2 = gfull2_prop
gammapart2 = gpart2_prop
prior1 = prior_prop1
prior2 = prior_prop2
prior3 = prior_prop3
Sigmafull = Sigmaf_prop
Sigmapart = Sigmap_prop
Sigmafulli = Sigmaf_propi
Sigmaparti = Sigmap_propi
Cafull = Caf_prop
Capart = Cap_prop
Crfull = Crf_prop
Crpart = Crp_prop
Wa = Wa_prop
Wr = Wr_prop
}
}
ii = floor(k/5)
if(ii == k/5){
  iii = (j-1)*1000 + ii
  lpost_store[iii] = lpost_cur
}

#####
#####
# Step 3: Draw Betas
#####
#####

#preliminaries
utilsf = matrix(t(utils[1:nfull,]),nfull*dim1)
utilsp = matrix(t(utils[(nfull+1):count,]),npart*dim1)
i = nfull*dim1
ii = i + 1
iii = count*dim1

#compute variance
#full-time
test10 = Sigmafulli %% matrix(X3[1:i,],dim1)
test11 = matrix(test10,nfull*dim1)
test12 = t(X3[1:i,]) %% test11
#part-time

```

```

test10 = Sigmaparti %%% matrix(X3[ii:iii,],dim1)
test11 = matrix(test10,npart*dim1)
test13 = t(X3[ii:iii,]) %%% test11
b_par13 = solve(test12 + test13 + Sigma_bari)

#compute mean
#full-time
test10 = Sigmafulli %%% matrix(utilsf,dim1)
test11 = matrix(test10,nfull*dim1)
test12 = t(X3[1:i,]) %%% test11
#part-time
test10 = Sigmaparti %%% matrix(utilsp,dim1)
test11 = matrix(test10,npart*dim1)
test13 = t(X3[ii:iii,]) %%% test11
b_par14 = b_par13 %%% (test12 + test13 + priorbar)

#draw beta
L_beta = t(chol(b_par13))
check2 = 1
counter1 = 0

#this loop ensures network variable parameters are negative (prior is truncated normal)
while(check2 == 1){
  temp10 = rnorm(ncovtot,0,1)
  beta = b_par14 + L_beta%%temp10
  temp20 = max(beta[(ncovtot-6):ncovtot])
  if(temp20 <= 0){
    check2 = 0
  } else {
    counter1 = counter1+1
  }
  if(counter1 >= 100){
    if(beta[ncovtot-6] >= 0){beta[ncovtot-6] = -0.0000001}
    if(beta[ncovtot-5] >= 0){beta[ncovtot-5] = -0.0000001}
    if(beta[ncovtot-4] >= 0){beta[ncovtot-4] = -0.0000001}
    if(beta[ncovtot-3] >= 0){beta[ncovtot-3] = -0.0000001}
    if(beta[ncovtot-2] >= 0){beta[ncovtot-2] = -0.0000001}
    if(beta[ncovtot-1] >= 0){beta[ncovtot-1] = -0.0000001}
    if(beta[ncovtot] >= 0){beta[ncovtot] = -200}
    check2 = 0
  }
}
}

```

```

#recompute systematic utilities
for(i in 1:count){
  ii = (i-1)*dim1 + 1
  iii = i*dim1
  V[i,] = X3[ii:iii,]%*%beta
}

#*****
#*****
# Step 4: Store Draws
#*****
#*****

theta_store[k,1:ncovtot] = beta
theta_store[k,(ncovtot+1):ncovtot2] = cov_param

} # for k

} # for j

```

REFERENCES

- Abbe, E., M. Bierlaire, and T. Toledo (2007) Normalization and Correlation of Cross-Nested Logit Models. *Transportation Research Part B*, 41, 795-808.
- Abkowitz, M.D. (1981) An Analysis of the Commuter Departure Time Decision. *Transportation*, 10 (3), 283-297.
- Abou Zeid, M., T.F. Rossi, and B. Gardner (2006) Modeling Time of Day Choice in the Context of Tour and Activity Based Models. *Transportation Research Record*, 1981, 42-49.
- Akaike, H. (1974) A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, 19, 716-723.
- Albert, J.H. and S. Chib (1993) Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, 88, 669-679.
- Bain, R. and M. Wilkins (2002) Infrastructure Finance: Traffic Risk in Start-Up Toll Facilities. Standard & Poor's, McGraw-Hill International (UK) Ltd., September 2002.
- Bain, R. and J.W. Plantagie (2003) Traffic Forecasting Risk: Study Update 2003. Standard & Poor's, McGraw-Hill International (UK) Ltd., November 2003.
- Bain, R. and J.W. Plantagie (2004) Traffic Forecasting Risk: Study Update 2004. Standard & Poor's, McGraw-Hill International (UK) Ltd., October 2004.
- Bain, R. and L. Polakovic (2005) Traffic Forecasting Risk Study Update 2005: Through Ramp-Up and Beyond. Standard & Poor's, McGraw-Hill International (UK) Ltd., August 2005.
- Bain, R., K. Forsgren, and P.B. Calder (2006) Credit FAQ: Assessing the Credit Quality of Highly Leveraged Deep-Future Toll-Road Concessions, Standard & Poor's, McGraw-Hill International (UK) Ltd., February 2006.
- Bates, J., J. Polak, P. Jones, and A. Cook (2001) The Valuation of Reliability for Personal Travel. *Transportation Research Part E*, 37, 191-229.

- Ben-Akiva, M. and T. Watanatada (1981) Application of a Continuous Spatial Choice Logit Model. In *Structural Analysis of Discrete Choice Data with Econometric Applications* (C.F. Manski and D. McFadden, eds.), MIT Press, Cambridge, MA, 320-343.
- 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*, 19 (2), 119-154.
- Ben-Akiva, M. and M. Bierlaire (1999) Discrete Choice Methods and Their Applications to Short-Term Travel Decisions. In *Handbook of Transportation Science* (R. Hall ed.), Kluwer, 5-34.
- 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*, 30 (3), 189-207.
- Bhat, C.R. (1998a) Analysis of Travel Mode and Departure Time Choice for Urban Shopping Trips. *Transportation Research Part B*, 32 (6), 361-371.
- Bhat, C.R. (1998b) Accommodating Flexible Substitution Patterns in Multi-Dimensional Choice Modeling: Formulation and Application to Travel Mode and Departure Time Choice. *Transportation Research Part B*, 32 (7), 455-466.
- Bhat, C.R. and F.S. Koppelman (1999) Activity-Based Modeling for Travel Demand. *Handbook of Transportation Science*, R.W. Hall (ed.), Kluwer Academic Publisher.
- Bhat, C.R. and J.L. Steed (2002) A Continuous-Time Model of Departure Time Choice for Urban Shopping Trips. *Transportation Research Part B*, 36 (3), 207-224.
- Bhat, C.R. and R. Sardesai (2006) The Impact of Stop-Making and Travel Time Reliability on Commute Mode Choice. *Transportation Research Part B*, 40 (9), 709-730.
- Bhat, C.R. and A.R. Pinjari (2008) Duration Modeling. In *Handbook of Transport Modeling, 2nd edition*, (D.A. Hensher and K.J. Button, eds.), Elsevier Science, 105-132.

- Bierlaire, M. (2006) A Theoretical Analysis of the Cross-Nested Logit Model. *Annals of Operations Research*, 144, 287-300.
- Bowman, J.L., M.A. Bradley, and J. Gibb (2006) The Sacramento Activity-Based Travel Demand Model: Estimation and Validation Results. Proceedings of the European Transport Conference, Strasbourg, France, September 2006.
- Brownstone, D. and K.A. Small (2005) Valuing Time and Reliability: Assessing the Evidence from Road Pricing Demonstrations. *Transportation Research Part A*, 39, 279-293.
- Cambridge Systematics, Inc. (2005) *Forecasting Person Travel by Time of Day*. Final Report prepared for the Federal Highway Administration (FHWA).
- Chib, S. and E. Greenberg (1998) Analysis of Multivariate Probit Models. *Biometrika*, 85 (2), 347-361.
- Chin, A.T.H. (1990) Influences on Commuter Trip Departure Time Decisions in Singapore. *Transportation Research Part A*, 24 (5), 321-333.
- Chipman, H., E. George, J. Lemp, and R. McCulloch (2009) Bayesian Flexible Modeling of Trip Durations. Working paper, University of Texas at Austin.
- Cox, D.R. (1972) Regression Models and Life Tables. *Journal of the Royal Statistical Society, Series B*, 34, 187-220.
- Cressie, N. (1995) Bayesian Smoothing of Rates in Small Geographic Areas, *Journal of Regional Science*, 35 (4), 659-673.
- Daly, A.J. and S. Zachary (1979) Improved Multiple Choice Models. In *Identifying and Measuring the Determinants of Mode Choice* (D. Hensher and Q. Dalvi, eds.), Teakfield, London, 335-357.
- 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*, 39 (3), 245-268.
- de Jong, G., A. Daly, M. Pieters, and T. van der Hoorn (2007) The Logsum as an Evaluation Measure: Review of the Literature and New Results. *Transportation Research Part A*, 41, 874-889.

- DYNASMART-P (2009) Dynamic Network Assignment-Simulation Model for Advanced Roadway Telematics. Website accessed July 20, 2009 at <http://mctrans.ce.ufl.edu/featured/dynasmart/>.
- Ettema, D., A. Borgers, and H.J.P. Timmermans (1995) Competing Risk Hazard Model of Activity Choice, Timing, Sequencing, and Duration. *Transportation Research Record*, 1493, 101-109.
- Ettema, D. and H. Timmermans (2003) Modeling Departure Time Choice in the Context of Activity Scheduling Behavior. *Transportation Research Record*, 1831, 39-46.
- Gadda, S., K.M. Kockelman, and P. Damien (2009) Continuous Departure Time Models: A Bayesian Approach. Proceedings of the 88th Annual Meeting of the Transportation Research Board, January 2009, Washington, D.C.
- Gamerman, D. (1997) Efficient Sampling from the Posterior Distribution in Generalized Linear Mixed Models. *Statistics and Computing*, 7, 57-68.
- Gamerman, D. and H.F. Lopes (2006) *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd Edition, Chapman & Hall/CRC, Boca Raton.
- Gelfand, A.E. and A.F.M. Smith (1990) Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85, 398-409.
- Gelman, A. and D. Rubin (1992) Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7, 457-511.
- Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin (2004) *Bayesian Data Analysis*, 2nd Edition, Chapman & Hall/CRC, Boca Raton.
- George, C., W. Streeter, and S. Trommer (2003) Bliss, Heartburn, and Toll Road Forecasts. Project Finance Special Report, Fitch Ratings, November 2003.
- Geweke, J. (1991) Efficient Simulation from the Multivariate Normal and Student t -Distributions Subject to Linear Constraints and the Evaluation of Constraint Probabilities. Proceedings of the 23rd Symposium on the Interface between Computer Science and Statistics, 571-578.

- Geweke, J. (1992) Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In *Bayesian Statistics 4* (J.M. Bernardo, J.O. Berger, A.P. Dawid, and A.F.M. Smith, eds.), Oxford University Press, Oxford, 169-193.
- Geweke, J., M. Keane, and D. Runkle (1994) Alternative Computational Approaches to Inference in the Multinomial Probit Model. *The Review of Economics and Statistics*, 76 (4), 609-632.
- Golob, T.F. and A.C. Regan (2002) Trucking Industry Adoption of Information Technology: A Multivariate Discrete Choice Model. *Transportation Research Part C*, 10, 205-228.
- Good, I.J. (1958) Significance Tests in Parallel and in Series. *Journal of the American Statistical Association*, 53, 799-813.
- Hastings, W.K. (1970) Monte Carlo Sampling Methods Using Markov Chains and their Applications. *Biometrika*, 57, 97-109.
- Heckman, J. and B. Singer (1984) A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica*, 52 (2), 271-320.
- Hendrickson, C. and E. Plank (1984) The Flexibility of Departure Times for Work Trips. *Transportation Research Part A*, 18 (1), 25-36.
- Hensher, D.A. (2001) The Valuation of Commuter Travel Time Savings for Car Drivers: Evaluating Alternative Model Specification. *Transportation*, 28, 101-118.
- Hobeika, A. (2005) TRANSIMS Fundamentals. Travel Model Improvement Program, http://tmip.fhwa.dot.gov/resources/clearinghouse/docs/transims_fundamentals/, Accessed July 2009.
- Holden, L., R. Hauge, and M. Holden (2009) Adaptive Independent Metropolis-Hastings. *The Annals of Applied Probability*, 19 (1), 395-413.
- Huber, J. and K. Train (2001) On the Similarity of Classical and Bayesian Estimates of Individual Mean Partworths. *Marketing Letters*, 12, 259-269.

- Hunt, L.M., P.C. Boxall, and B. Boots (2007) Accommodating Complex Substitution Patterns in A Random Utility Model of Recreational Fishing. *Marine Resource Economics*, 22, 155-172.
- Jonnalagadda, N., J. Freedman, W.A. Davidson, and J.D. Hunt (2001) Development of Microsimulation Activity-Based Model for San Francisco: Destination and Mode Choice Models. *Transportation Research Record*, 1777, 25-35.
- Kass, R.E. and A.E. Raftery (1995) Bayes Factors. *Journal of the American Statistical Association*, 90 (430), 773-795.
- Kissling, W.D. and G. Carl (2008) Spatial Autocorrelation and the Selection of Simultaneous Autoregressive Models. *Global Ecology and Biogeography*, 17, 59-71.
- Kockelman, K.M. and J.D. Lemp (2009) The Financing of New Highways: Opportunities for Welfare Analysis and Credit-Based Congestion Pricing. Proceedings of the 88th Annual Meeting of the Transportation Research Board, Washington, D.C.
- Komma, A. and S. Srinivasan (2008) Modeling Home-to-Work Commute-Timing Decisions of Workers with Flexible Work Schedules. Proceedings of the 87th Annual Meeting of the Transportation Research Board, January 2008, Washington, D.C.
- Lam, T.C. and K.A. Small (2001) The Value of Time and Reliability: Measurement from a Value Pricing Experiment. *Transportation Research Part E*, 37 (3), 231-251.
- Lee, B. and H.J.P. Timmermans (2007) A Latent Class Accelerated Hazard Model of Activity Episode Durations. *Transportation Research Part B*, 41 (4), 426-447.
- Lemp, J.D. and K.M. Kockelman (2009) Understanding and Accommodating Risk and Uncertainty in Toll Road Projects: A Review of the Literature. Forthcoming in *Transportation Research Record*.
- Lichstein, J.W., T.R. Simons, S.A. Shriner, and K.E. Franzreb (2002) Spatial Autocorrelation and Autoregressive Models in Ecology. *Ecological Monographs*, 72 (3), 445-463.

- Lin, D.Y., N. Eluru, S.T. Waller, and C.R. Bhat (2008) Integration of Activity-Based Modeling and Dynamic Traffic Assignment, *Transportation Research Record*, 2076, 52-61.
- Marzano, V. and A. Papola (2008) On the Covariance Structure of the Cross-Nested Logit Model. *Transportation Research Part B*, 42, 83-98.
- McCafferty, D. and F.L. Hall (1982) The Use of Multinomial Logit Analysis to Model the Choice of Time to Travel. *Economic Geography*, 58 (3), 236-246.
- McCulloch, R. and P.E. Rossi (1994) An Exact Likelihood Analysis of the Multinomial Probit Model. *Journal of Econometrics*, 64, 207-240.
- McFadden, D. (1973) Conditional Logit Analysis of Qualitative Choice Behavior. In *Frontiers in Econometrics* (P. Zarembka ed.), Academic Press, New York.
- McFadden, D. (1976) The Mathematical Theory of Demand Models. In *Behavioral Travel Demand Models* (P.R. Stopher and A.H. Meyburg, eds.), Lexington Books, Lexington, MA, 305-314.
- McFadden, D. (1978) Modeling the Choice of Residential Location. In *Spatial Interaction Theory and Planning Models* (A. Karlquist, L. Lundquist, F. Snickbars, and J.W. Weibull, eds.), North-Holland, Amsterdam, 75-96.
- McFadden, D. (1989) A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration. *Econometrica*, 57 (5), 995-1026.
- McFadden, D. and K. Train (2000) Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, 15 (5), 447-470.
- Metropolis, N., A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller (1953) Equation of State Calculations by Fast Computing Machine. *Journal of Chemical Physics*, 21, 1087-1091.
- Niemeier, D.A. and J.G. Morita (1996) Duration of Trip-Making Activities by Men and Women: A Survival Analysis. *Transportation*, 23 (4), 353-371.

- Noland, R.B. and K.A. Small (1995) Travel Time Uncertainty, Departure Time Choice, and the Cost of Morning Commutes. *Transportation Research Record*, 1493, 150-158.
- Okola, A. (2003) Departure Time Choice for Recreational Activities by Elderly Non-Workers. *Transportation Research Record*, 1848, 86-93.
- Papola, A. (2004) Some Developments on the Cross-Nested Logit Model. *Transportation Research Part B*, 38, 833-851.
- Papola, A. and V. Marzano (2005) A Specification Procedure of the CNL Model Reproducing Any Homoskedastic Covariance Matrix. Proceedings of the European Transport Conference, Strasbourg, France.
- Parent, O. and J.P. LeSage (2008) Using the Variance Structure of the Conditional Autoregressive Spatial Specification to Model Knowledge Spillovers. *Journal of Applied Econometrics*, 23, 235-256.
- PB Consult (2005a) Transportation Models and Data Initiative: General Final Report, New York Best Practice Model (NYBPM). Prepared for the New York Metropolitan Transportation Council.
- PB Consult (2005b) The MORPC Travel Demand Model: Validation and Final Report. Prepared for the Mid-Ohio Regional Planning Commission as part of the MORPC Model Improvement Project.
- Popkowski Leszczyc, P.T.L. and H. Timmermans (2002) Unconditional and Conditional Competing Risk Models of Activity Duration and Activity Sequencing Decisions: An Empirical Comparison. *Journal of Geographical Systems*, 4 (2), 157-170.
- Popuri, Y., M. Ben-Akiva, and K. Prousaloglou (2008) Time of Day Modeling in a Tour-Based Context: The Tel-Aviv Experience. Proceedings of the 87th Annual Meeting of the Transportation Research Board, January 2008, Washington, D.C.
- Roberts, G.O. and A.F.M. Smith (1994) Simple Conditions for the Convergence of the Gibbs Sampler and Metropolis-Hastings Algorithms. *Stochastic Processes and their Applications*, 49, 207-216.

- Saleh, W. and S. Farrell (2005) Implications of Congestion Charging for Departure Time Choice: Work and Non-Work Schedule Flexibility. *Transportation Research Part A*, 39 (9), 773-791.
- Schofer, J.L. (2005) Summary Statement. Proceedings of the USDOT Expert Forum on Road Pricing and Travel Demand Modeling, Alexandria, VA.
- Schwarz, G. (1978) Estimating the Dimension of a Model. *The Annals of Statistics*, 6, 461-464.
- Small, K.A. (1982) The Scheduling of Consumer Activities: Work Trips. *The American Economic Review*, 72 (3), 467-479.
- Small, K.A. (1987) A Discrete Choice Model for Ordered Alternatives. *Econometrica*, 55 (2), 409-424.
- Small, K.A., R. Noland, X. Chu, and D. Lewis (1999) Valuation of Travel-Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation. NCHRP Report 431, Transportation Research Board, Washington, D.C.
- Small, K.A., C. Winston, and J. Yan (2005) Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability. *Econometrica*, 73 (4), 1367-1382.
- Smith, A.F.M. and G.O. Roberts (1993) Bayesian Computation Via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods. *Journal of the Royal Statistical Society B*, 55, 3-102.
- Smith, T.E. and J.P. LeSage (2004) A Bayesian Probit Model with Spatial Dependencies. In *Spatial and Spatiotemporal Econometrics* (J.P. LeSage and R.K. Pace, eds.), Elsevier, Amsterdam, 127-160.
- Spiegelhalter, D.J., N.G. Best, B.P. Carlin, A. van der Linde (2002) Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society, Series B*, 64 (4), 583-639.
- Srinivasan, K.K. and Z. Guo (2003) Analysis of Trip and Stop Duration for Shopping Activities: Simultaneous Hazard Duration Model System. *Transportation Research Record*, 1854, 1-11.

- Steed, J. and C.R. Bhat (2000) On Modeling the Departure Time Choice for Home-Based Social/Recreational and Shopping Trips. *Transportation Research Record*, 1706, 152-159.
- Stopher, P.R., R. Alsnih, C.G. Wilmot, C. Stecher, J. Pratt, J. Zmud, W. Mix, M. Freedman, K. Axhausen, M. Lee-Gosselin, A.E. Pisarski, and W. Brog (2008) Technical Appendix, NCHRP Report 571: Standardized Procedures for Personal Travel Surveys, National Cooperative Highway Research Program, Transportation Research Board, Washington, D.C.
- Texas Transportation Institute (TTI) and Cambridge Systematics (CS) Inc. (2006) Travel Time Reliability: Making it There on Time, All the Time, Report prepared for the Federal Highway Administration.
- Tierney, L. (1994) Markov Chains for Exploring Posterior Distributions. *Annals of Statistics*, 22, 1701-1762.
- Train, K. (2009) Bayesian Procedures. Chapter 12 of *Discrete Choice Methods with Simulation*, 2nd Edition, Cambridge University Press, 282-314.
- Transportation Research Board (2007) Metropolitan Travel Forecasting: Current Practice and Future Direction. TRB Special Report 288, Committee for Determination of the State of the Practice in Metropolitan Area Travel Forecasting, Washington, D.C.
- Tringides, C.A., X. Ye, and R.M. Pendyala (2004) Departure-Time Choice and Mode Choice for Nonwork Trips. *Transportation Research Record*, 1898, 1-9.
- Tseng, Y.Y. and E. Verhoef (2008) Value of Time by Time of Day: A Stated-Preference Study. *Transportation Research Part B*, 42 (7-8), 607-618.
- van Lint, J.W.C., H.J. van Zuylen, and H. Tu (2008) Travel Time Unreliability on Freeways: Why Measures Based on Variance Tell Only Half the Story. *Transportation Research Part A*, 42, 258-277.
- Vovsha, P. (1997) Application of Cross-Nested Logit Model to Mode Choice in Tel-Aviv, Israel, Metropolitan Area. *Transportation Research Record*, 1607, 6-15.

- Vovsha, P. and M. Bradley (2004) A Hybrid Discrete Choice Departure Time and Duration Model for Scheduling Travel Tours. *Transportation Research Record*, 1894, 46-56.
- Vovsha, P., W. Davidson, and R. Donnelly (2005) Making the State of the Art the State of the Practice: Advanced Modeling Techniques for Road Pricing. Proceedings of the USDOT Expert Forum on Road Pricing and Travel Demand Modeling, Alexandria, VA.
- Wagenmakers, E.J., M. Lee, T. Lodewyckx, and G.J. Iverson (2008) Bayesian Versus Frequentist Inference. Chapter 9 of *Bayesian Evaluation of Informative Hypotheses* (H. Hoijtink, I. Klugkist, and P.A. Boelen, eds.), Springer, 181-210.
- Wang, J.J. (1996) Timing Utility of Daily Activities and Its Impact on Travel. *Transportation Research Part A*, 30 (3), 189-206.
- Wen, C.H. and F.S. Koppelman (2001) The Generalized Nested Logit Model. *Transportation Research Part B*, 35, 627-641.
- Williams, H.C.W.L. (1977) On the Formation of Travel Demand Models and Economic Evaluation Measures of User Benefit. *Environment and Planning A*, 9, 285-344.
- Wooldridge, J.M. (2002) *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA and London, UK.
- Yee, J.L. and D.A. Niemeier (2000) Analysis of Activity Duration Using the Puget Sound Transportation Panel. *Transportation Research Part A*, 34 (8), 607-624.
- Zellner, A. and C.K. Min (1995) Gibbs Sampler Convergence Criteria. *Journal of the American Statistical Association*, 90, 921-927.
- Zhang, X., W.J. Boscardin, and T.R. Belin (2008) Bayesian Analysis of Multivariate Nominal Measures Using Multivariate Multinomial Probit Models. *Computational Statistics and Data Analysis*, 52, 3697-3708.
- Zhao, Y., K.M. Kockelman, and A. Karlstrom (2008) Welfare Calculations in Discrete Choice Settings: The Role of the Error Term Correlation. Proceedings of the 87th Annual Meeting of the Transportation Research Board, Washington, D.C.

Zhong, M. and J.D. Hunt (2005) Modeling Household Weekend Activity Durations in Calgary. Proceedings of the 2005 Annual Conference of the Transportation Association of Canada, Calgary, Alberta.

VITA

Jason Lemp was born August 14, 1982 to Kent and Gloria Lemp in St. Louis, Missouri. He is married to Sarah Carnes-Lemp and attended the University of Missouri in Columbia, Missouri, where he graduated cum laude, earning his B.S. in civil and environmental engineering. During that time he worked a research project, collecting and processing traffic data collected via mobile video cameras. Following graduation he worked briefly for the Michigan Department of Transportation in the Traffic Signals Unit, before coming to the University of Texas at Austin to pursue graduate studies in Transportation Engineering.

Jason started graduate studies at UT in the fall of 2005. During his four and a half years of graduate study, he developed an interest in travel demand modeling and traveler behavior, which involves rigorous statistical training and data analysis. He has contributed to several research projects during his time at UT, including projects sponsored by the Environmental Protection Agency (EPA), the National Cooperative Highway Research Program (NCHRP), and the Strategic Highway Research Program (SHRP). He was also a recipient of the Dwight D. Eisenhower Graduate Fellowship in 2007 and 2008. Jason's research interests include transportation economics, statistical methods for transportation data analysis, and risk assessment, and he is author of several published papers in these areas.

Permanent Address:

5127 Reynosa Drive

Saint Louis, MO 63128

This dissertation was typed by the author.