A RANDOM PARAMETERS FRAMEWORK FOR FREEWAY LANE USE MODELING

A Thesis in

Civil Engineering

by

Sai Sirandals

© 2012 Sai Sirandals

Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Science

August 2012
The thesis of Sai Sirandas was reviewed and approved* by the following:

Venky N. Shankar  
Associate Professor of Civil and Environmental Engineering  
Thesis Advisor

Swagata Banerjee  
Assistant Professor of Civil and Environmental Engineering

Prasenjit Basu  
Assistant Professor of Civil and Environmental Engineering

Peggy A. Johnson  
Professor of Civil and Environmental Engineering  
Head of the Department of Civil and Environmental Engineering

*Signatures are on file in the Graduate School
ABSTRACT

Lane changing is a behavioral parameter that is not often captured by conventional traffic data-collection means. By utilizing micro-simulation as a tool to model traffic flow, behavioral observations can be replicated and measured to the finest detail not available by typical field data-collection standards. This study attempts to use the VISSIM micro-simulation tool to obtain key insights into the influence of roadway geometrics and traffic flow metrics on the probability of lane-changing. For the purpose of this research, a 30-mile corridor of the heavily congested Interstate-5 in Washington State has been constructed and calibrated in the VISSIM software. The resulting simulation output yielded lane-change count data and various traffic flow metrics that were arranged in panel data format and analyzed in a random parameters modeling framework. A “heterogeneity in the means of random parameters” model of probability of high incoming-lane-change-ratio has been estimated; aggregate speed descriptors and speed class distributions were found to have a significant influence on the lane changing probability. Further, the observed heterogeneity in the means of random parameters due to cross-sectional effects such as lane type, lane position, and interchange effect has successfully been accounted for.

The estimated effects of traffic flow metrics and geometrics on lane-changing characteristics can not only be used to better understand the phenomenon of lane-changing, but can also assist in developing congestion-mitigation and evacuation strategies for complex transportation networks.
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ vi

LIST OF TABLES .......................................................................................................... vii

ACKNOWLEDGEMENTS ............................................................................................. viii

Chapter 1 INTRODUCTION ..................................................................................... 1

Chapter 2 RELATED WORKS AND RESEARCH QUESTIONS ........................... 3
  2.1 The Use of Micro-simulation ........................................................................ 3
  2.2 Statistical Modeling of Lane-Change Behavior ......................................... 4
  2.3 Panel Data and Random Parameters Framework .................................... 5

Chapter 3 ANALYSIS PROCESS AND EMPIRICAL SETTINGS ....................... 7
  3.1 Analysis Process .......................................................................................... 7
  3.2 Study Area ................................................................................................. 9
  3.3 VISSIM Network Construction ................................................................ 11
    3.3.1 Data Collection ............................................................................... 11
    3.3.2 Links and Connectors ..................................................................... 13
    3.3.3 Desired Speed Distributions ......................................................... 14
    3.3.4 Traffic Compositions and Vehicle Inputs ..................................... 15
    3.3.5 Traffic Routing ............................................................................. 15
    3.3.6 Driver Behavior Parameters ......................................................... 16
  3.4 VISSIM Calibration Results ....................................................................... 20
    3.4.1 Volumes ........................................................................................ 20
    3.4.2 Travel Times ............................................................................... 22
  3.5 Simulation Output and Descriptive Statistics ............................................ 23
    3.5.1 Data Collection Locations .............................................................. 23
    3.5.2 Descriptive Statistics ................................................................... 25

Chapter 4 STATISTICAL MODELING OF MICROSIMULATION OUTPUTS .... 32
  4.1 Random Parameters Model Specification ............................................... 32
    4.1.1 Notational Convention .................................................................. 32
    4.1.2 Model Specification ..................................................................... 32
  4.2 Random Parameters Approach for Modeling Incoming Lane-Change Ratio ................................................................. 35
  4.3 Model Estimation Results for Random Parameters Probit Model ............ 37
    4.3.1 Non Random Parameters .............................................................. 38
    4.3.2 Means of Random Parameters .................................................... 39
    4.3.3 Heterogeneity in the Means of Random Parameters ..................... 40
4.4 Comparing Estimates of Probit, Logit, Gompertz and Complementary Log Log Models ................................................................. 41
  4.4.1 Random and Non-Random Parameters .............................................. 43
  4.4.2 Maximum Simulated Likelihood ....................................................... 44
  4.4.3 Symmetry of Parameter Distributions ................................................. 44
4.5 Marginal Effects and Elasticities ............................................................ 46

Chapter 5 CONCLUSION AND DIRECTIONS FOR FUTURE RESEARCH ................................................................. 50
  5.1 Residuals Discussion ............................................................................ 50
  5.2 Discussion on Selection Effect ............................................................... 51
  5.3 Conclusions and Recommendations ..................................................... 53

References ........................................................................................................ 55
LIST OF FIGURES

Figure 3-1: Overview of this research .................................................................8
Figure 3-2: Study Area – Interstate 5 in Washington State ..........................10
Figure 3-3: Links and Connectors in VISSIM ..................................................13
Figure 3-4: Desired Speed Distribution in VISSIM .......................................14
Figure 3-5: Car-Following Parameters in VISSIM .........................................17
Figure 3-6: Lane-Changing Parameters in VISSIM .......................................18
Figure 3-7: Complex Interchange Segment ...............................................18
Figure 3-8: Finished 30 Mile network in VISSIM ........................................19
LIST OF TABLES

Table 3-1: Northbound SOV and HOV Volumes (veh/hr) ................................................. 21
Table 3-2: Southbound SOV and HOV Volumes (veh/hr) ................................................. 21
Table 3-3: Travel Times (min) .......................................................................................... 22
Table 3-4: Data Collection Locations ............................................................................. 24
Table 3-5: Descriptive Statistics ...................................................................................... 26
Table 4-1: Results of Random Parameters Probit Model for HICHNG ......................... 38
Table 4-2: Comparison of Model Estimates .................................................................... 42
Table 4-3: Comparison of Elasticities .............................................................................. 47
Table 4-4: Design Vectors having Elasticity of SPRANGE > 1 .................................. 48
Table 5-1: Comparison of Base Model Residuals to Full Model Residuals ............... 50
ACKNOWLEDGEMENTS

This thesis is dedicated to my parents Ramachandram Sirandas and Indira Vennem for their unconditional belief and support in all my endeavors.

I will always be indebted to my Advisor Dr. Venkataraman N. Shankar for his kindness, guidance, and emotional and financial support over my two-year stay at Penn State. I am thankful to Dr. Shankar for sharing his unparalleled knowledge of transportation and modeling concepts and helping me develop this thesis. I consider myself extremely fortunate to be his advisee and a part of his lab. Thanks Venky.

I wish to thank my committee members Dr. Swagata Banerjee Basu and Dr. Prasenjit Basu for their valuable inputs and thoughtful discussions. Thank you for taking the time to review and suggest edits to the previous drafts of my thesis. I acknowledge my lab-mate Ko-Yu Chen’s contributions to this research; I thank him for helping me to learn the VISSIM software, and collect data and calibrate the simulation model. A big thank-you to my lab-mate Daniel W. Kwon for mentoring me as a junior graduate student and a member of the lab.

I thank my uncle’s family in North Carolina - my uncle Ramulu Posham, my aunt Madhavi Posham and my cousins Raghuram Posham and Vinya Posham for providing me a home away from home and helping me forget that I am thousands of miles away from my home country. I owe a special thank-you to my aunt Madhavi for her motherly care and delicious food.
Finally, I would like to thank my roommates Mark A. Glowatz, Ramandeep Singh and Melvin Zelissen, and fellow graduate students Jonathan C. Crisafi, Cody M. Morris, and Jung Yeol Hung for all the good times in State College.
Chapter 1

INTRODUCTION

Lane-changing is a very common, yet critical phenomenon in traffic flow. It greatly affects the safety and mobility of the surrounding traffic and the overall network. In recent years, the worsening congestion on major interstate highways in the United States has created a need for adaptive and real time traffic management. In order to develop congestion-mitigation strategies and incident response traffic control, it is essential to closely understand the characteristics of lane-changing traffic flow.

Traffic micro-simulation models have proven to be powerful tools in transportation analyses. Micro-simulation platforms provide for a means to model and analyze complex transportation systems under various scenarios. These tools are primarily used in transportation research to replicate field conditions and measure various traffic-flow metrics at desired levels of resolution. In this study, the VISSIM (PTV AG, 2012) microscopic traffic simulation platform has been used to obtain insights on vehicular lane-changing behavior in a congested freeway setting.

The primary objective of this research is to obtain insights into freeway lane usage by modeling the influence of traffic flow metrics and roadway geometrics on lane-change probability. Data extracted from the VISSIM simulation output included traffic flow metrics such as vehicle speeds, throughput, travel times, speed class distributions and lane-change counts. This data was arranged in a panel format consisting of repeated time series observations. The lane-changing probability was then modeled in a random parameters modeling framework.
The uniqueness of this research is twofold. The usefulness of VISSIM software in simulating long (> 30 miles) corridors of congested freeway network is validated. Further, the applicability of random parameters modeling framework to model lane-change probabilities has been demonstrated.
Chapter 2

RELATED WORKS AND RESEARCH QUESTIONS

2.1 The Use of Micro-simulation


In order to accurately calibrate the VISSIM model to replicate field observations, the user needs to have a thorough knowledge of the various parameters in VISSIM that define the traffic control operations, traffic flow characteristics, and driver behavior (Choa et al. 2004). Often, many of these parameters require modifications in order to replicate field observations. In addition to the VISSIM user manual (PTV AG 2012), various publications that provide guidance on how to modify these parameters have been referenced in this study. Gomes et al. (2004) suggest modifications to some key lane-changing parameters (look back distance, and emergency stop distance) and car-following parameters (minimum headway, standstill distance and following thresholds) in order to model various link types such as basic freeway, weaving sections, merging
sections etc. Menneni et al. (2007) studied the effect of car-following parameters on the speed-flow relationships. Lownes and Machemehl (2006) provide useful insights on the influence of multiple car-following parameter combinations on the simulated capacity. Finally, the Oregon DOT VISSIM Protocol (ODOT 2011) provides a practical guide to constructing, calibrating and validating a VISSIM simulation model.

2.2 Statistical Modeling of Lane-Change Behavior

Prior research includes work in the areas of macroscopic as well as microscopic (predominantly in the former) lane-changing characteristics. Jin (2010) studied the aggregate properties of lane-changes and calibrated a relationship between lane-changing intensity and traffic density. Jin (2009) also demonstrated that lane-changing intensities are highly related to road geometry, location vehicle speeds etc. Chang and Kao (1991) investigated and predicted the influence of traffic flow variables such as density, flow rate and headway on the number of lane-changes and lane-change probabilities in a macroscopic setting. They specified a logistic transformation ($Y^*$) of the probability of lane-changing as

$$ Y^* = -5.45 + 0.0005(AF) + 0.037(ADR) + 0.45(AS) + \epsilon, $$

where

$AF$ is the average lane flow rate (vehicles/lane/hr),

$ADR$ is the average density ratio, and

$AS$ is the average lane speed (mile/hr).
Sheu and Ritchie (1999) proposed a method for stochastic modeling and real-time prediction of vehicular lane-changing behavior during incidents, using lane traffic counts as the sole input data. Goswami and Bham (2006) studied the microscopic lane-changing behavior characterized by individual behavior of drivers in terms of headways, speed profiles etc. over a half mile section.

2.3 Panel Data and Random Parameters Framework

In this study, data is gathered at different locations (groups) on the simulated freeway network over a period of time resulting in a panel or longitudinal data set. In recent years, panel data analysis has been widely favored and used in various areas of transportation including safety (Venkataraman et al. 2011), planning (Su 2010), operations (Karato et al. 2009) etc. The increasing interest in panel data analysis is due to its many advantages. The primary benefit of panel data is that having multiple observations at the same location allow for the modeler to capture the heterogeneity across individuals (or groups). Additionally, panel analyses also allow accounting for latent dynamic effects across cross-sections (Wooldridge 2003, Greene 2004). Another obvious advantage is that panel data sets due to their large size and rich nature, help to reduce the problem of collinearity among variables and may also give more precise modeling estimates.

The concept of heterogeneity across individuals (groups) was introduced into the classical panel model by the fixed effects and random effects models. In the fixed effects model, it is assumed that unobserved individual heterogeneity is correlated with the
explanatory variables, and the differences across groups are captured by the constant term alone (Greene 2011). Greene further illustrates that a major drawback of this model is that the effects time-invariant dependent variables are absorbed by the constant term as well and their coefficients cannot be estimated. In the random effects model, it is assumed that the unobserved heterogeneity is uncorrelated with the regressors. In this case, there is a high risk of that assumption being incorrect. The random parameters model is a highly versatile hierarchical model that allows for not only the variation of parameters across individuals, but also the mean value of the parameter distribution to be individual specific. In this research, LIMDEP (2007) software has been used for statistical modeling. The four available model types for binary choice (Probit, logit, Gompertz and complimentary log-log) have been implemented.
Chapter 3

ANALYSIS PROCESS AND EMPIRICAL SETTINGS

This chapter covers the data collection and VISSIM model calibration processes that were carried out in order to generate the simulation outputs containing data for the econometric modeling stages.

3.1 Analysis Process

Figure 3-1 on the ensuing page presents an outline of the main steps involved in this study. The first step was to calibrate the VISSIM software with the help of roadway inventory data, speed data and traffic volume data. The simulation was then iteratively calibrated in order to produce outputs that are consistent with the observed field data. Calibration is done by modifying the various driving behavior parameters in VISSIM. Once the desired level of accuracy of the simulation is achieved, speed and lane change count data are extracted from the model and arranged in a panel data format. A random parameters approach is then applied to model the effects of speed and roadway data on lane-changing probabilities.
Figure 3-1: Overview of this research
3.2 Study Area

The study area extends along a 32-mile-long freeway corridor of Interstate 5 (I-5) located in King County, Washington State. The corridor under study begins at ARM (Accumulated Route Mileage) 141.25 (Milepost 141.19) in Federal Way and ends at ARM 173.57 (Milepost 173.51) in Northgate, Seattle, spanning a total length of 32.32 miles. In addition to the northbound and southbound mainline segments, the study area also includes the reversible express lanes beginning at ARM 0.00 (Milepost 165.29) in Downtown Seattle and ending at ARM 7.14 (Milepost 172.43) in Northgate, Seattle. The mainline and reversible-expressway ramp terminals were also included in this study. The time period of analysis was one hour at average PM peak hour conditions. The study area is shown in Figure 3-2.

Travel time data and real time traffic maps from WSDOT indicate that I-5 experiences heavy congestion during peak hour conditions. Studies indicate that this stretch of I-5 (Federal Way to Seattle) performs the worst in comparison to other major routes of Washington State in terms of various congestion performance measures (WSDOT 2011).
Figure 3-2: Study Area – Interstate 5 in Washington State
3.3 VISSIM Network Construction

In this section, the various steps involved in constructing a VISSIM simulation network are discussed in detail. The steps involved are data collection, construction of links and connectors, desired speed distributions, traffic compositions and vehicle inputs, and traffic routing and driver behavior parameters.

3.3.1 Data Collection

Construction and calibration of the VISSIM simulation model requires various inputs including roadway inventory, traffic control data, traffic flow data, and speed and travel time information.

The following sources of data were utilized in this study:

- High-resolution aerial images from Google Maps (https://maps.google.com/): Screenshots of aerial photographs from Google Maps were downloaded and manually stitched to create larger images. These images were later scaled to serve as background images for the VISSIM network.

- Information about the number of lanes, lane and roadway width, SOV/HOV classification, ramp meter locations, and legal speeds were obtained in the WSDOT State Highway Log (WSDOT 2010)

- State route numbers, Related Route Type (RRT) codes, Milepost and Accumulated Route Mileage (ARM) were found in the WSDOT Interchange
Viewer Tool

- Street View for Google Maps (http://maps.google.com/help/maps/streetview/)
and the WSDOT SR Web Tool
(http://www.wsdot.wa.gov/mapsdata/tools/srweb.htm) were used to locate regulatory, warning and guide signs on the mainline as well as ramps.

- The Northwest Traffic Volume – Ramp and Roadway Report provided by the NW Region Traffic Management Center
(http://www.wsdot.wa.gov/Northwest/TrafficVolume/) contained information about the average PM Peak hourly volumes for HOV and SOV lanes of the mainline, reversible expressway and ramps. These volumes had to be balanced in order to be fed into the VISSIM software.

- Google Maps, WSDOT WebFLOW application
(http://www.wsdot.com/traffic/seattle/products/webflow.aspx) and WSDOT Seattle Area Travel Times (http://www.wsdot.com/traffic/seattle/traveltimes/)
were used to obtain information containing traffic maps, travel times and spot speeds.

In the next few sections, technical aspects of the VISSIM simulation model are described.
3.3.2 Links and Connectors

The base network comprising of “links” and “connectors” was constructed by tracing the network on the scaled background images. The attributes of these links and connectors include the number of lanes, lane widths, lane usage and lane-changing restrictions. These parameters allow for implementation of features such as the High-Occupancy Vehicle (HOV) lanes, exit only lanes etc. Additionally, each link is associated with some driver behavior parameters. These attributes are discussed in the following sections. Figure 3-3 shows links (in blue) and connectors (in pink) in a typical VISSIM network.

---

Figure 3-3: Links and Connectors in VISSIM
3.3.3 Desired Speed Distributions

Posted speeds on the mainline and ramps were used to define stochastic desired speed distributions. The “desired” speed is the speed a driver will travel at if not hindered by other vehicles or obstructions. These distributions were defined with the posted speed plus or minus 10 miles per hour as the limits. Figure 3-4 illustrates the desired speed distribution for a posted speed of 45 miles per hour. These distributions were spatially assigned to the links and connectors. The ideal location for a desired speed distribution would be the location of speed limit signs.

Figure 3-4: Desired Speed Distribution in VISSIM
3.3.4 Traffic Compositions and Vehicle Inputs

The next step involves defining vehicle compositions for the network to be populated with. For the freeway network under study, there are two sources of traffic loadings, the north and south ends of the mainline, and the on-ramps. With the help of WSDOT’s peak hour average volume data, these source nodes were assigned proportions of HOV and SOV vehicles class volumes. Additionally, the HOV and SOV vehicle classes were color coded for visualization purposes. After defining the vehicle volumes and HOV/SOV proportions, these vehicles were also assigned a speed distribution that governs the speed at with they enter the network.

In this study, passenger car was the only vehicle type used. For simplicity purposes and lack of data, trucks, buses, motorcycles etc. were not separately coded. This is a significant assumption even though the major proportion of vehicles in the field is of the passenger car type.

3.3.5 Traffic Routing

Now that the network is populated, the vehicles have to be “routed” to their respective destinations in order to simulate the field conditions. Routing decisions are rules that govern the path of vehicles on the network. Based on the off-ramp volumes from the WSDOT data, vehicles were accordingly routed to the off ramps. Theoretically, the start of a routing decision is a few hundred feet upstream of the guide sign location. When a vehicle crosses the start point of the routing decision (when the driver sees the
guide sign), it accordingly begins to change its path in order to either stay on the mainline or take the downstream off ramp.

Often vehicles need to change lanes in order to satisfy a particular routing decision. With the help of routing decision and link parameters, one can define the distance upstream of the off-ramp where a vehicle begins to change lanes in order to reach the off-ramp or stay on the mainline. These rules have a great impact on the lane-changing behavior of vehicles in the network.

### 3.3.6 Driver Behavior Parameters

This is the most important and also the most difficult and time consuming step in the calibration of the simulation model. Driver behavior parameters specify the car-following and lane-changing behavior in VISSIM. In the network, every link is associated with a driver behavior type. When a vehicle enters a link, it assumes the link specific driver behavior type. VISSIM software implements the widely accepted Wiedemann 99 car-following model for freeway networks. This model incorporates various parameters such as minimum headway, standstill distance, acceleration etc. The lane-changing parameters such as the minimum and maximum deceleration, rate of deceleration, minimum safety-distance etc. dictate the aggressiveness of the lane-changing process in the simulation.

These parameters greatly affect the safety, speed and flow rate at the corresponding location and also the overall performance of the network. Often, the default values for these parameters have to be modified in order for the simulation to
replicate field observations. The VISSIM Protocol (ODOT, 2011) specifies recommended and allowable ranges for the majorly impactful driver behavior parameters. With the help of the specified ranges and multiple simulation trials, these parameters were calibrated to represent a driving behavior both from the aerial perspective and the driver’s perspective that is closest to the field observations. Depending upon the location in the network, various driver behavior types to reflect basic freeway movement, light and heavy weaving, on-ramp merging etc. were defined and assigned to the corresponding links on the network. Shown below in Figures 3-5 and 3-6 are typical car following and lane-changing behavior parameter sets. Figure 3-7 and Figure 3-8 present a close-up view of a complex interchange segment and a zoomed out view of the 30 mile network respectively.

![Car-Following Parameters in VISSIM](image)

Figure 3-5: Car-Following Parameters in VISSIM
Figure 3-6: Lane-Changing Parameters in VISSIM

<table>
<thead>
<tr>
<th>Necessary lane change (route)</th>
<th>Own</th>
<th>Trailing vehicle</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum deceleration:</td>
<td>-14.99 ft/s²</td>
<td>-12.01 ft/s²</td>
</tr>
<tr>
<td>- 1 ft/s² per distance:</td>
<td>200.00 ft</td>
<td>150.00 ft</td>
</tr>
<tr>
<td>Accepted deceleration:</td>
<td>-3.51 ft/s²</td>
<td>-2.49 ft/s²</td>
</tr>
</tbody>
</table>

- Waiting time before diffusion: 60.00 s
- Min. headway (front/rear): 1.64 ft
- To slower lane if collision time above: 0.00 s
- Safety distance reduction factor: 0.50
- Maximum deceleration for cooperative braking: -14.99 ft/s²
- Overtake reduced speed areas: [ ]
- Advanced merging: [ ]

- Cooperative lane change: [ ]
- Maximum speed difference: 6.71 mph
- Maximum collision time: 10.00 s

Figure 3-7: Complex Interchange Segment
Figure 3-8: Finished 30 Mile network in VISSIM
3.4 VISSIM Calibration Results

The primary measures that were used to calibrate the VISSIM output are discussed in this section. The calibration goals were met by conducting trial runs of the simulation and adjusting the driving behavior parameters and link features until a desired level of confidence in the simulation output was achieved.

3.4.1 Volumes

The first step in calibration is to ensure that the link volumes match the field measurements. Both SOV and HOV volumes were measured on the mainline and HOV reversible lanes. These volumes were then compared to the field-observed volumes that were fed into VISSIM as simulation inputs. Additionally, FHWA (2004) recommends the use of GEH statistic (named after Geoffrey E. Havers) to compare the observed flows to the simulation model flows. The GEH statistic is given as

\[ GEH = \sqrt{\frac{2(M - C)^2}{M + C}} \]

where

- \( M \) is the output traffic volume from the simulation model and
- \( C \) is the observed field volume (simulation input)

A value less than 5 for GEH is considered an acceptable fit.
From Table 3-1 and Table 3-2 below, it can be seen that the percentage differences are below 7% and the GEH statistic is acceptable at all the given locations.

### Table 3-1: Northbound SOV and HOV Volumes (veh/hr)

<table>
<thead>
<tr>
<th>Northbound</th>
<th>SOV Volumes</th>
<th>HOV Volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simulation</td>
<td>Field</td>
</tr>
<tr>
<td>Federal Way</td>
<td>5718</td>
<td>5700</td>
</tr>
<tr>
<td>Before Southcenter</td>
<td>6271</td>
<td>6147</td>
</tr>
<tr>
<td>Southcenter Interchange</td>
<td>3757</td>
<td>3622</td>
</tr>
<tr>
<td>Southcenter Interchange 0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>After Southcenter</td>
<td>6517</td>
<td>6210</td>
</tr>
<tr>
<td>Before I-90</td>
<td>6645</td>
<td>6891</td>
</tr>
<tr>
<td>I-90 Interchange</td>
<td>4571</td>
<td>4660</td>
</tr>
<tr>
<td>After I-90 Interchange</td>
<td>4554</td>
<td>4660</td>
</tr>
<tr>
<td>I-5/I-90 Collector Distributor</td>
<td>3675</td>
<td>3589</td>
</tr>
<tr>
<td>Downtown Seattle</td>
<td>2246</td>
<td>2304</td>
</tr>
<tr>
<td>NB Reversible Lane Begin</td>
<td>1649</td>
<td>1660</td>
</tr>
<tr>
<td>NE 85th St</td>
<td>3678</td>
<td>3980</td>
</tr>
<tr>
<td>NE 85th St Reversible Lane</td>
<td>2546</td>
<td>2564</td>
</tr>
<tr>
<td>Northgate way</td>
<td>6591</td>
<td>6793</td>
</tr>
</tbody>
</table>

### Table 3-2: Southbound SOV and HOV Volumes (veh/hr)

<table>
<thead>
<tr>
<th>Southbound</th>
<th>SOV Volumes</th>
<th>HOV Volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Simulation</td>
<td>Field</td>
</tr>
<tr>
<td>Begin Northgate Way</td>
<td>5370</td>
<td>5730</td>
</tr>
<tr>
<td>NE 85th Street</td>
<td>4128</td>
<td>4446</td>
</tr>
<tr>
<td>I-5/I-90 Collector Distributor</td>
<td>1511</td>
<td>1625</td>
</tr>
<tr>
<td>Downtown Seattle</td>
<td>3000</td>
<td>3250</td>
</tr>
<tr>
<td>I-90 Interchange</td>
<td>2991</td>
<td>3250</td>
</tr>
<tr>
<td>After I-90 Interchange</td>
<td>6450</td>
<td>6847</td>
</tr>
<tr>
<td>Before Southcenter</td>
<td>9447</td>
<td>9436</td>
</tr>
<tr>
<td>Southcenter HOV Grade Separated</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Southcenter Interchange</td>
<td>6215</td>
<td>6613</td>
</tr>
<tr>
<td>After Southcenter</td>
<td>8146</td>
<td>8113</td>
</tr>
<tr>
<td>Federal Way</td>
<td>6405</td>
<td>6507</td>
</tr>
</tbody>
</table>
3.4.2 Travel Times

With the help of travel time data from WSDOT and Google Maps, travel times from the simulation output were compared against the field measurements. WSDOT’s congestion report (2011) provides average travel times for Federal Way to Seattle and Seattle to Federal way during peak hour of PM rush. Travel times from Google Maps during were averaged over a period of one week to serve as an additional benchmark. The time period of observation of Google Maps’ travel times was 5 PM to 6 PM on weekdays.

As shown in Table 3-3, the simulated travel times were within the acceptable 15% range (FHWA, 2011) of WSDOT and Google Maps’ times.

<table>
<thead>
<tr>
<th>Northbound</th>
<th>WSDOT</th>
<th>Google Maps</th>
<th>Simulation</th>
<th>% difference(WSDOT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Federal Way to Seattle (22 mi)</td>
<td>29</td>
<td>27</td>
<td>25.6</td>
<td>-11.72%</td>
</tr>
<tr>
<td>Seattle to Northgate Way (8 mi)</td>
<td>13.2</td>
<td>11.5</td>
<td>11.3</td>
<td>-14.39%</td>
</tr>
<tr>
<td>Southbound</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northgate Way to Seattle (8 mi)</td>
<td>12.3</td>
<td>10.5</td>
<td>11</td>
<td>-10.57%</td>
</tr>
<tr>
<td>Seattle to Federal Way (22 mi)</td>
<td>30</td>
<td>28</td>
<td>26</td>
<td>-13.33%</td>
</tr>
</tbody>
</table>

Note: WSDOT does not provide travel times for Seattle to Northgate (and vice versa). These times were approximated by using the average speed from Seattle to Everett (and vice versa).
In addition to the above calibration measures, a visual inspection of the model was performed. It was ensured that the on-ramp and off-ramp queuing, and lane utilization at various lane drop locations and the HOV lane were visually acceptable.

3.5 Simulation Output and Descriptive Statistics

3.5.1 Data Collection Locations

After the calibration and validation of the simulation runs, the acceptable simulation model was ready for the data extraction phase. The goal of this stage was to extract speed and lane-change count data from the simulation output. The panel data setup is achieved by extracting repeated measures involving time series of observations at fixed 5-minute intervals. The idea was to collect data at basically two points in space – the midpoint of interchange segments and the midpoint of non-interchange segments.

While the time intervals are much easier to define, identifying spatial locations of midpoints of interchange and non-interchange segments is more challenging. A brute force method was used here – the midpoint of the non-interchange segment was located first for the sequence of interchanges over the 30 mile corridor, travelling in both northbound and southbound directions. Subjectivity in defining the midpoints exists in the sense that midpoints are not identified on the basis of quantitatively defined stability of traffic flow, but in terms of the proximity of the upstream on-ramp and downstream off-ramp end points. Due to this midpoint identification protocol, conditioning on the identification of non-interchange midpoint, creates its own set of problems for midpoint
identification of interstate segments – for example, some interchange segments are complex and long, and a single midpoint for these segments located along the middle of the splined length between two successive non-interchange segments may not be representative of the segment as a whole in terms of traffic behavior.

The primary idea was to create a data collection structure that can be mimicked in both directions, while maintaining directional uniqueness in location and measurement points. After careful examination of the network, the following locations listed in Table 3-4 along the corridor were determined to be appropriately representative of the behavior on non-interchange and interchange segments:

<table>
<thead>
<tr>
<th>Location</th>
<th>MP</th>
<th>Feature</th>
<th>Interchange/Non Interchange</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>141.25</td>
<td>Enchanted Pkwy overpass</td>
<td>NI</td>
</tr>
<tr>
<td>2</td>
<td>142.05</td>
<td>Federal Way S 348th underpass</td>
<td>I</td>
</tr>
<tr>
<td>3</td>
<td>142.81</td>
<td>S 336th underpass</td>
<td>NI</td>
</tr>
<tr>
<td>4</td>
<td>143.85</td>
<td>S 320th overpass</td>
<td>I</td>
</tr>
<tr>
<td>5</td>
<td>145.25</td>
<td>Military Rd underpass</td>
<td>NI</td>
</tr>
<tr>
<td>6</td>
<td>146.84</td>
<td>S 272nd underpass</td>
<td>I</td>
</tr>
<tr>
<td>7</td>
<td>148.08</td>
<td>S 259th Underpass, S 248th</td>
<td>NI</td>
</tr>
<tr>
<td>8</td>
<td>149.2</td>
<td>SR 516, S Kent Des Moines Rd</td>
<td>I</td>
</tr>
<tr>
<td>9</td>
<td>150.35</td>
<td>S 216th overpass</td>
<td>NI</td>
</tr>
<tr>
<td>10</td>
<td>151.03</td>
<td>Military Rd Overpass</td>
<td>I</td>
</tr>
<tr>
<td>11</td>
<td>151.75</td>
<td>Angle Lake</td>
<td>NI</td>
</tr>
<tr>
<td>12</td>
<td>152.28</td>
<td>188th street underpass</td>
<td>I</td>
</tr>
<tr>
<td>13</td>
<td>153.16</td>
<td>178th St. overpass</td>
<td>NI</td>
</tr>
<tr>
<td>14</td>
<td>154.14</td>
<td>SR 518/ I 405 overpass</td>
<td>I</td>
</tr>
<tr>
<td>15</td>
<td>155.32</td>
<td>S 144th Overpass</td>
<td>NI</td>
</tr>
<tr>
<td>16</td>
<td>155.8</td>
<td>SR 599 overpass</td>
<td>I</td>
</tr>
<tr>
<td>17</td>
<td>156.48</td>
<td>Railroad Ave underpass, S 129th St Overpass</td>
<td>NI</td>
</tr>
<tr>
<td>18</td>
<td>157.77</td>
<td>Martin Luther King Jr. Way S</td>
<td>I</td>
</tr>
<tr>
<td>19</td>
<td>159.5</td>
<td>S Rose St/ Military Rd underpass</td>
<td>NI</td>
</tr>
</tbody>
</table>
3.5.2 Descriptive Statistics

The initial data set consisted of a balanced panel of 12 observations across 430 groups – a total of 5160 groups. However, it was observed that during certain time intervals in the PM peak hour of simulation, no lane-changes were reported. In such cases, inferences cannot be made about the ratio of incoming to outgoing lane-change ratio since the ratio is unknown. This renders the balanced panel unbalanced.

Balanced panel data models are far easier to estimate and are featured in most off the counter statistical packages such as LIMDEP. Estimating a model without ignoring
the nonresponses (zero-lane total lane-changes) requires a deeper knowledge of modeling and is not within the scope of this study. In order to work with a balanced panel, it was necessary to eliminate from the dataset the groups containing at least one observation with the total lane-change count being zero. As a result, the reduced data set consisted of 243 groups and 2916 observations, and hence the model becomes more selective. The problem of selection bias due to the reduced data set is addressed in the discussion section in Chapter 5.

From the VISSIM simulation output, more than 50 variables were measured at the midpoints of interchange and non-interchange segments. The following Table 3-5 lists the descriptive statistics of the variables that were considered in order to model the probability of high incoming lane-change ratio.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Mean</th>
<th>Std.Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCPIID</td>
<td>Data Collection Point ID</td>
<td>111</td>
<td>4024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCPL</td>
<td>Data Collection Point Location</td>
<td>1</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIR1</td>
<td>Northbound Direction Dummy</td>
<td>0.56</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>DIR2</td>
<td>Southbound Direction Dummy</td>
<td>0.44</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>INTCH</td>
<td>Interchange Dummy</td>
<td>0.58</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LTD</td>
<td>Lane Type ID</td>
<td>2.19</td>
<td>1.72</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>LANE1</td>
<td>Lane Number 1 Dummy</td>
<td>0.19</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LANE2</td>
<td>Lane Number 2 Dummy</td>
<td>0.32</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LANE3</td>
<td>Lane Number 3 Dummy</td>
<td>0.27</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LANE4</td>
<td>Lane Number 4 Dummy</td>
<td>0.14</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LANE5</td>
<td>Lane Number 5 Dummy</td>
<td>0.07</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LANE6</td>
<td>Lane Number 6 Dummy</td>
<td>0.01</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LANE7</td>
<td>Lane Number 7 Dummy</td>
<td>0.00</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LTD1</td>
<td>Northbound SOV Lane Type Dummy</td>
<td>0.39</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LTD2</td>
<td>Southbound SOV Lane Type Dummy</td>
<td>0.39</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LTD3</td>
<td>Northbound Reversible Lane Type Dummy</td>
<td>0.12</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>LTD4</td>
<td>Northbound Collector Distributor Lane Type Dummy</td>
<td>0.02</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
The dummy variables DIR1 and DIR2 represent the northbound and southbound directions respectively. The mean of DIR1 is 0.56 and the mean of DIR2 is 0.44. There are more groups in the northbound direction because of the express lanes that operate in
the northbound direction during the PM peak hour time. These lanes operate in addition
to the northbound mainline lanes.

The interchange dummy $INTCH$ has a mean of 0.58 and a standard deviation of
0.49. $INTCH = 1$ represents a non-interchange segment and $INTCH = 0$ represents an
interchange segment. In this study, an interchange segment is defined by the farthest
diverge and merge ramp limits in the direction of travel. A non-interchange segment is
defined as a continuous travel segment between two interchange segments.

$LANE1$ to $LANE7$ are lane number dummies, $LANE1$ being the rightmost lane in
the direction of travel. $LANE1$, $LANE2$, $LANE3$ and $LANE4$ have means of 0.18, 0.32,
0.27 and 0.14 respectively. It is odd that there are fewer instances of $LANE1$ than
$LANE2$ and $LANE3$. This is due to the selection effect caused by excluding groups
having zero total lane-change counts. On a busy interchange corridor such as I5, the
rightmost lane (Lane 1) experiences high amounts of queuing due to on-ramp and off-
ramp vehicles. It is understandable that at there are a good number of lane 1 groups that
experience no lane-changes in the proximity of the midpoints of non-interchange
segments.

$LTD1$ to $LTD9$ are dummy variables representing various lane types observed in
the study area. $LTD1$ and $LTD2$ represent northbound and southbound SOV lanes (or
mixed lanes) respectively and have means of 0.39 and 0.39 respectively. $LTD3$ has a
mean of 0.12 and represents the lanes on the reversible expressways. During the PM Peak
hour, the reversible lanes operate in the northbound direction. $LTD8$ and $LTD9$ are
dummy variables for northbound and southbound HOV lanes. These lanes run in both
directions on I-5 between Federal Way and Northgate. In the northbound direction, the mainline HOV lane ends just before I-90, about a mile upstream of where the reversible lanes begin, and reappears after at Northgate where the reversible lanes end. In the southbound direction, the HOV lane exists intermittently between Northgate and downtown Seattle, and exists continuously from I-90 until Federal Way. LTD8 and LTD9 have a mean of 0.025. In addition to the mainline and reversible lanes, collector distributor lanes have also been coded into the network. These lanes run parallel to the I-5 mainline in both northbound and southbound directions near downtown Seattle and handle traffic entering and exiting I-5 and I-90. These are represented by LTD4 and LTD5 in northbound and southbound directions respectively. LTD4 has a mean of 0.017 and LTD5 has a mean of 0.033. The 430 groups in the original dataset also consisted of LTD6 and LTD7 that represented grade separated HOV lanes. But these groups were excluded from the data set because they are grade separated single lane segments and hence do not experience any lane-changing activity.

TFLOW, SOVF, and HOVF are continuous variables expressing five-minute flows through the midpoints of interchange and non-interchange segments. These variables are expressed in number of vehicles per 5 minutes. TFLOW had a mean of 124.98 and a standard deviation of 35.64. SOVF had a mean of 105.21 and a standard deviation of 41.08, and HOVF had a mean of 19.77 and a standard deviation of 19.88. The maximum values of SOVF and HOVF were 192 and 136 respectively, while TFLOW had a maximum value of 199 vehicles.
Several speed variables were extracted from the simulation output. $ILMIN$, the in-lane minimum speed had a mean of 40.65 MPH and a standard deviation of 11.54 MPH and minimum and maximum values of 0.2 MPH and 56.3 MPH respectively. $ILMAX$, the in-lane minimum speed had a mean of 65.61 MPH and a standard deviation of 4.77 MPH and minimum and maximum values of 12.9 MPH and 70 MPH respectively. The in-lane mean speed was expressed by $ILMSP$ and had a mean of 53.7 MPH and a standard deviation of 7.24 MPH. The minimum and maximum values were 7.2 MPH and 60.4 MPH respectively. The speed deviation $SPDEV$ variable had a mean of 5.34 MPH and a standard deviation of 2.65 MPH and minimum and maximum values of 0.73 MPH and 17.64 MPH respectively. The speed range (difference between the maximum and minimum speeds) variable $SPRANGE$ had a mean of 24.97 MPH and a standard deviation of 10.02 MPH. The minimum speed range was 1.9 MPH and maximum speed range was 65.6 MPH. The ratio of maximum to minimum speed $MAXMIN$ had a mean of 2.42 and a standard deviation of 5.17, and minimum and maximum values of 1.05 and 106 respectively.

In addition to the aggregate speed descriptors listed above, data for speed class distributions was also extracted. $SP\nu_i\nu_k$ is used to express the frequency of vehicles in the speed class $\nu_i\text{MPH}−\nu_k\text{MPH}$. There were a total of 14 (5 MPH) speed classes from 0 MPH to 70 MPH. These intervals adequately capture the variation in speeds.

The lane-change counts were extracted from the simulation output. Lane-change counts were measured within +/- 0.025 miles of the midpoints of interchange and non-interchange segments. The idea was to capture the number of incoming and outgoing
lane-changes occurring in this 0.05 mile envelope. \textit{INLCH} was used to express the ratio of incoming lane-change count to the total lane-change count. It had a mean value of 0.50 and a standard deviation of 0.34. Finally, the dependent variable in the analysis \textit{HICHNG}, is a dummy variable that assumes a value of 1 if \textit{INLCH} is greater than 0.5 and a value of 0 otherwise. \textit{HICHNG} had a minimum value of 0.41. This variable was used to represent the probability of incoming lane-change count being higher than the corresponding outgoing lane-change count.
4.1 Random Parameters Model Specification

4.1.1 Notational Convention

In what follows, matrix and vector notations and manipulations have been used. Scalar variables have been denoted with lowercase italic letters ($\gamma_{lt}, \epsilon_{lt}$ etc.) Column vectors (including scalar vectors) are denoted with boldface lowercase letters ($\mathbf{x}_{lt}, \mathbf{\beta}_l$ etc.), and all matrices are denoted with boldface uppercase letters ($\mathbf{Z}_l, \mathbf{X}_l$ etc.)

4.1.2 Model Specification

The generalized formulation of the random parameters Probit model is given by:

\[ y_{lt}^* = \mathbf{\beta}'_{lt}\mathbf{x}_{lt} + \epsilon_{lt}, \quad i = 1, \ldots, N \text{ and } t = 1, \ldots, T, \]

\[ y_{lt} = 1(y_{lt}^* > 0) \]

\[ y_{lt} = 0(y_{lt}^* \leq 0) \]

\[ \mathbf{\beta}_l = \mu + \Delta \mathbf{z}_l + \Gamma \mathbf{v}_l. \]

where

the data consist of $N$ observations on $\mathbf{Z}_l = (\mathbf{y}_l, \mathbf{X}_l)$ where $\mathbf{y}_l = (y_{l1}, y_{l2}, \ldots, y_{lT})$, $y_{lt}$ is the response variable, $\mathbf{x}'_{lt}$ contains the main explanatory variables in the $T$ rows of the $T \times K$ matrix $\mathbf{X}_t$, $\mathbf{\beta}_l = \mu + \Delta \mathbf{z}_l + \Gamma \mathbf{v}_l$.
$z_t$ is a $L \times 1$ vector of time invariant, individual (group) specific variables that influence the means of the random parameters $\beta_t$.

$v_t$ is a $K \times 1$ vector of random latent individual effects,

with the following assumptions:

\[ \varepsilon_{it} \sim N[0,1], \]
\[ v_t \sim N[0, I], \]
\[ E[v_t|X_t, z_t] = 0, \text{ and} \]
\[ Var[v_t|X_t, z_t] = I. \]

The structural parameters of the model are:

\[ \mu, \text{ a } K \times 1 \text{ vector of constant terms in the means of the random parameters}, \]
\[ \Delta, \text{ a } K \times L \text{ matrix of unknown parameters that multiply the covariates in the distribution of random parameters, and} \]
\[ \Gamma, \text{ a } K \times K \text{ lower triangular matrix of unknown variance parameters}. \]

It follows that $\beta_t$ is normally distributed with the moments

\[ E[\beta_t|X_t, z_t] = \mu + \Delta z_t, \text{ and} \]
\[ Var[\beta_t|X_t, z_t] = \Gamma \Gamma'. \]

In the two-level model specification above, the first level estimates the response variable $y_{it}$ as a function of the main covariates $x_{it}$. The four primary models covered in this study differ in the distributions of the outcome variable $y_{it}$ - Probit, logit, Gompertz, and complementary log-log.
In the second level, the panel data setting allows for a group specific coefficient ($\mathbf{\beta}_t$) that may be drawn from a distribution, the parameters of which are influenced by the observed dataset. In the estimation of the four primary models covered in this study, it is assumed that the random parameters are normally distributed. The means of these normal distributions are modeled as a function of group specific effects $z_t$. This method provides a means to capture and analyze the observed heterogeneous effects in the means of the random parameters. In the correlated version of the model, the random parameters are allowed to be correlated, in which case $\mathbf{\Gamma}$ is a lower triangular matrix with non-zero off-diagonal elements and the full covariance matrix of the random coefficients is given by $\mathbf{\Gamma}\mathbf{\Gamma}'$. In the random parameter model framework, non-random parameters can be accommodated by forcing the corresponding elements in $\mathbf{\Delta}$ and $\mathbf{\Gamma}$ to contain zeros.

Given the above specifications, the parameter estimates for the random parameters model are estimated by maximizing the likelihood function for the observed data. According to Greene (2002), the direct optimization of the likelihood function is not feasible due to the multidimensional nature of the likelihood integral. As an alternative, the maximum simulated likelihood by a quasi-Monte Carlo method based on Halton sequences of draws is used. Halton draws are symmetric draws that population the parameter space uniformly. This method allows for the computation intensive maximum simulated likelihood estimation to be completed in a time-efficient fashion.
4.2 Random Parameters Approach for Modeling Incoming Lane-Change Ratio

This section covers the four main models that have been implemented to predict the incoming lane-change ratio. The dependent variable is the ratio of incoming lane-change count to the total lane-change count for a given lane at a given location for a given observation time period.

\[ INLCH_{it} = \frac{No.\ of\ incoming\ lane\ changes_{it}}{Number\ of\ incoming\ lane\ changes_{it} + Number\ of\ Outgoing\ lane\ changes_{it}} \]

where \( i = 1, \ldots, 430 \) locations, and \( t = 1, \ldots, 12 \) time periods of observation.

The value of \( INLCH \) is bounded by 0 and 1 and can be coded as a discrete response variable as:

\[ HICHNG_{it} = \begin{cases} 1, & \forall \ INLCH_{it} > 0.5, \\ 0, & \forall \ INLCH_{it} \leq 0.5 \end{cases} \]

The variable \( HICHNG \) assumes the value of 1 when the ratio of incoming lane-change count to outgoing lane-change count exceeds 0.5 and is categorized as “high incoming lane-change.” \( HICHNG \) will be modeled as a binary outcome in the model framework outlined in the previous section.

The main explanatory variables in the first level of the random parameters model are expressed in matrix notation as

\[ x_{it} = \begin{bmatrix} \frac{1}{SPRANGE_{it}} \\ SP5055_{it} \\ SP5560_{it} \\ SP6065_{it} \\ INMS_{it} \end{bmatrix} \]
In the first level of the model, the binary outcome $HICHNG_{it}$ was modeled a function of time-varying aggregate speed descriptors and speed class distributions. In the vector $x_{it}$ above, the random parameters comprise the constant term and the speed range parameter ($SPRANGE$). The non-random parameters comprise the variables containing frequencies of speed-classes from 50 MPH to 65 MPH and the mean speed variable ($SP5055, SP5560, SP6065$ and $INMSP$). These parameters were found to be statistically significant in the modeling process.

The time invariant vector containing group (location) specific cross section effects that influence the heterogeneity in the means of the random parameters is expressed in matrix notation as

$$
\begin{bmatrix}
LANE2_{it} \\
LTD3_{it} \\
LTD8_{it} \\
LTD38_{it} \\
INCHI_{it}
\end{bmatrix}
$$

After preliminary runs of the model, some elements of the parameter matrix $\Delta$ turned out to be statistically insignificant. In the final runs of the models, these elements were restricted to zeros. The matrix $\Delta$ is expressed as

$$
\Delta' =
\begin{bmatrix}
\delta_{ONE\_LANE2} & \delta_{SPRANGE\_LANE2} \\
\delta_{ONE\_LTD3} & 0 \\
0 & 0 \\
\delta_{ONE\_LTD8} & 0 \\
0 & \delta_{SPRANGE\_INCHI}
\end{bmatrix}
$$

In the second level of the model, it was expected that the time invariant vector containing cross-sectional effects such as lane position, lane type, interchange and direction dummy variables would capture group-specific heterogeneity. After preliminary
runs, interchange dummy, lane position 2, and the lane types 3, 8 and 38 were found to be significant.

Lane position 2 corresponds to the second lane from the right in the direction of travel. Lane type 38 is a union of the two dummy variables lane type 3 and lane type 8, i.e. northbound reversible lanes and northbound HOV lanes respectively. Preliminary runs of the model showed that lane type 3 and lane type 8 were both significant in the mean of the constant term. It was noted that the reversible lanes are active in the northbound direction during the analysis period of this study (PM hours.) Further, it was also observed that the northbound HOV lane ends where the reversible lanes viaduct begins and reappears where the viaduct ends. The lane type 38 variable was created to account for the aforementioned observations.

4.3 Model Estimation Results for Random Parameters Probit Model

The findings for the random parameters Probit model are presented in Table 4-1.
4.3.1 Non Random Parameters

The speed class frequencies $SP5055, SP5560, SP6065$, and the in-lane mean speed $INMSP$ had significant coefficients. $INMSP$ has a negative coefficient of $-0.024$. This effect was offset by the positive coefficients of $0.008$, $0.008$ and $0.017$ respectively for the higher speed classes from 50MPH to 65MPH. This is expected since faster are likely to experience more incoming lane-change events than outgoing lane-change events. Additionally, as opposed to a fixed slope across the 50-65 MPH range, the 60-65 range has maximum influence on the probabilities of high incoming lane-change ratios.

The above finding suggests that the combined positive effect of the four parameters $SP5055, SP5560, SP6065$ and $INMSP$ on the probability of high incoming
lane-change ratio is most substantial at segments experiencing high flows at high mean speeds. At low flows and high mean speeds, the negative effect of $\text{INMSP}$ is more prominent, producing an overall negative effect on the probability.

### 4.3.2 Means of Random Parameters

A mean value of -0.546 and a standard deviation of 1.336 for the constant term indicate that 65.87% of all the 243 groups will have a negative intercept in the equation of high incoming lane-change ratio ($\text{HICHNG}$).

The mean speed parameter ($\text{SPRANGE}$) has a mean of 0.014 and a standard deviation of 0.001 indicating that in almost 100% of the groups, an increase in the speed range is associated with an increase in the probability of high incoming lane-change ratio. A higher range of speeds on a lane can be associated with higher gaps that allow for more incoming lane-change events.

Of particular interest are the groups where $\text{SPRANGE}$ has maximum influence. An acceptable measure of this influence is the product of the variable $\text{SPRANGE}$ and its parameter mean. In groups 2312, 2313, and 1522, i.e. lanes 2 and 3 of segment 23 and lane 2 of segment 15, it was observed that the effect of $\text{SPRANGE}$ was most substantial. Interestingly, these are lanes 2, 3, and 2 respectively of non-interchange segments within 1000 feet of the downstream off-ramp. In both cases, lane 1 happens to be a high volume auxiliary lane connecting the upstream acceleration lane and downstream deceleration lane.
4.3.3 Heterogeneity in the Means of Random Parameters

The time invariant variables interchange dummy (\textit{INTCH}), lane position dummy (\textit{LANE2}), and lane type dummy (\textit{LTD3, LTD8}) were found to be statistically significant in the heterogeneity in the means of random parameters. \textit{LANE2} has a coefficient of 0.490 in the heterogeneity in the means of the constant term. This implies that in the equation of the probability of \textit{HICHNG}, there is an increase in the heterogeneity of the mean of the intercept associated with the lane 2 effect. In the simulation model, lane 2 corresponds to the second lane from the right in the direction of travel. In the proximity of midpoints of non-interchange segments on a busy corridor, lane 2 generally experiences the following necessary lane-change events:

- incoming lane-changes by vehicles entering from the upstream on-ramp
- both incoming and outgoing lane-changes by vehicles intending to take the downstream off ramp.

The effect of lane 2 can be attributed to higher incoming lane-changes than outgoing lane-changes due to the vehicles entering from the upstream on-ramp on non-interchange segments. In addition to that, descriptive statistics for lane 2 indicate that it has the highest mean among the lane position indicators. It has to be noted that there is a selection effect caused by a relatively higher number of lane 1 groups in the initial 430 groups experiencing zero total lane-changes. Due to this, the mean of lane 2 dummy is higher than the mean of lane 1 dummy.

\textit{LTD3} and \textit{LTD8} have coefficients of 0.546 and -0.729 respectively in the heterogeneity of the means of the constant term. The northbound HOV dummy (\textit{LTD3})
has a positive effect on high incoming lane-change probability. The overall volume increases in the northbound direction from north to south and it is expected that higher number of drivers will try to enter the HOV lanes.

\( LANE2 \) and \( INTH \) dummy variables were found to produce significant effects in the heterogeneity of the means of \( SPRANGE \) with coefficients of -0.014 and 0.004 respectively. This suggests that in lane 2 groups, the mean of the speed range parameter is lower than other that of other lane groups and in non-interchange segments, the mean of the speed range parameter is slightly higher than that of interchange segments. It has to be noted that \( LANE2 \) itself has a strong effect on the mean of the constant term.

### 4.4 Comparing Estimates of Probit, Logit, Gompertz and Complementary Log Log Models

In Table 4-2 on the following page, the results of the random parameter estimation are compared across the four density specifications.
Table 4-2: Comparison of Model Estimates

<table>
<thead>
<tr>
<th>Model</th>
<th>Probit</th>
<th>Logit</th>
<th>Gompertz</th>
<th>Comp Log Log</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coeff.</td>
<td>b/s.e</td>
<td>Coef.</td>
<td>b/s.e</td>
</tr>
<tr>
<td>Fixed Parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed Group 50-55</td>
<td>0.0086</td>
<td>4.1630</td>
<td>0.0103</td>
<td>4.1650</td>
</tr>
<tr>
<td>Speed Group 55-60</td>
<td>0.0081</td>
<td>4.7520</td>
<td>0.0098</td>
<td>4.7410</td>
</tr>
<tr>
<td>Speed Group 60-65</td>
<td>0.0173</td>
<td>4.5460</td>
<td>0.0212</td>
<td>4.6220</td>
</tr>
<tr>
<td>In-Lane Mean Speed</td>
<td>-0.0240</td>
<td>-4.0080</td>
<td>-0.0283</td>
<td>-3.8870</td>
</tr>
<tr>
<td>Random Parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.5465</td>
<td>-1.6930</td>
<td>-0.7172</td>
<td>-1.8250</td>
</tr>
<tr>
<td>Speed Range</td>
<td>0.0138</td>
<td>2.9240</td>
<td>0.0173</td>
<td>3.0230</td>
</tr>
<tr>
<td>Standard Deviations of Random Parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.3361</td>
<td>12.8160</td>
<td>2.3179</td>
<td>12.1750</td>
</tr>
<tr>
<td>Speed Range</td>
<td>0.0007</td>
<td>0.6700</td>
<td>0.0013</td>
<td>0.7550</td>
</tr>
<tr>
<td>Percentage of Parameter Distribution that is positive P(x&gt;0)=1-P(X&lt;=0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>34.1256</td>
<td>37.8501</td>
<td>54.8681</td>
<td>19.0173</td>
</tr>
<tr>
<td>Speed Range</td>
<td>100.000</td>
<td>100.000</td>
<td>98.0812</td>
<td>100.000</td>
</tr>
<tr>
<td>Heterogeneity in the means of random Parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant_Lane2</td>
<td>0.6012</td>
<td>3.0480</td>
<td>0.6012</td>
<td>3.0940</td>
</tr>
<tr>
<td>Constant_Lane Type 3</td>
<td>0.6638</td>
<td>4.4920</td>
<td>0.6638</td>
<td>4.4850</td>
</tr>
<tr>
<td>Constant_Lane Type 8</td>
<td>-0.9291</td>
<td>-2.6510</td>
<td>-0.9291</td>
<td>-2.7010</td>
</tr>
<tr>
<td>Speed Range_Lane2</td>
<td>-0.0177</td>
<td>-2.4630</td>
<td>-0.0177</td>
<td>-2.4940</td>
</tr>
<tr>
<td>Speed Range_Lane2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range_Interchange</td>
<td>0.0045</td>
<td>1.8020</td>
<td>0.0045</td>
<td>1.6990</td>
</tr>
<tr>
<td>Likelihood</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Restricted Log Likelihood for Baseline Random Effects Model</td>
<td>-1587.8</td>
<td>-1590.9500</td>
<td>-1600.2440</td>
<td>-1582.9390</td>
</tr>
<tr>
<td>Restricted Log Likelihood for Final Random Parameters Model</td>
<td>-1607.6</td>
<td>-1611.1890</td>
<td>-1618.6990</td>
<td>-1604.2850</td>
</tr>
<tr>
<td>Chi Squared Probability</td>
<td>0.0001</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
4.4.1 Random and Non-Random Parameters

The estimates of the logit model were very similar to those of the Probit model, both in terms of parameters and the $\beta/SE$ ratios. The non-random parameters and means of the random parameters of the probit model were in the 80-85% range of the corresponding logit model coefficients.

In both the probit and the logit models, the estimate for the mean of the constant term were of relatively weaker strength with $\beta/SE$ ratios of -1.693 and -1.875 respectively. The estimate for the constant term in the Gompertz model had the weakest strength across the four models with a $\beta/SE$ ratio of 0.503. The complimentary log-log model had the strongest estimate for the constant parameter with a $\beta/SE$ of -3.778. The Gompertz model also had the weakest estimate for the mean of the speed range parameter ($\beta/SE=2.455$) while the complimentary log-log model had the strongest estimate for the mean of the speed range parameter across the four groups ($\beta/SE=3.261$). In the heterogeneity of the means of the random parameters, the estimate for the coefficient of interchange dummy in the mean of speed range parameter had a weak strength across all the four models, suggesting that the relationship might not exist.

Greene (2004) points out that the conventional significance tests for individual structural parameters do no truly indicate the significance and strength of a relationship among the model variables. Wooldridge (2000) states that rejecting the evidence of a relationship based on the $\beta/SE$ ratio in a correlated model could be a “serious error.” Although the individual parameter might not be significant, they could be strongly correlated with other parameters in the model. In the four models discussed above,
although individually the constant term does not have a high statistical significance, it was significantly correlated with the speed range parameter. Greene (2004) adds that the latent variation in the normally distributed random vector $\mathbf{v}_i$ might have a significant effect on the random parameters.

### 4.4.2 Maximum Simulated Likelihood

The restricted likelihood computed by LIMDEP is the final likelihood for the fixed parameter equivalent of the random parameters model. The true baseline likelihood however, should be the random effects model, with constant term being the lone random parameter and the lone variable.

For all four models, based on the likelihood ratio test, the hypothesis of homogeneity of the model is rejected. The Complementary log-log model resulted in the best improvement in maximum likelihood estimates from the base random parameters model, the Gompertz model resulting in the least.

### 4.4.3 Symmetry of Parameter Distributions

While the Probit and the logit distributions are symmetric and similar in shape, the Gompertz and complementary log-log models are asymmetric. The cumulative distribution functions (CDF’s) of the Gompertz and complimentary log-log models are closely related. The CDF of the complimentary log-log distribution with a scale parameter $\beta$ is given by:
\[ F(x) = 1 - \exp(-\exp(\beta x)), \]

and the CDF of the Gompertz distribution with a scale parameter \( \beta \) is given by withdrawing form unity the above CDF as:

\[ F(x) = \exp(-\exp(\beta x)). \]

Due to their complementary nature, the Gompertz and the complementary log-log models are correspondingly asymmetric on either side of zero. The former is widely used in growth-curve modeling, in fields such as plant pathology, where the growth of bacteria is expected to increase over time. In such models, the probability density to the right of zero is greater than that to the left of zero. The complementary log-log model on the other hand is used in survival analysis (death of bacteria). It has a greater probability density to the left of zero than to the right of zero.

In the incoming lane change ratio probability models estimated in this study, the suitability of each of the above four models is driven by the threshold used to define \( HIC\)NG. As per the definition used, the binary outcome of \( HIC\)NG is ‘1’ when the ratio of the incoming lane-change count to the total lane-change count exceeds 0.5, and ‘0’ otherwise. It has to be noted that majority of the observed values for this ratio are less than the threshold value of 0.5, indicating the default suitability of complementary log-log, and hence its relatively better fit. Lowering the threshold from 0.5 may imply a more symmetric distribution (Probit or logit) and further lowering it may yield better likelihoods for the Gompertz model. Furthermore, this aspect of asymmetry can be useful in examining the lane-change balance across lanes and along the freeway. For instance, in a typical 5 lane cross-section, lanes 3 and 4 might have a different lane-change balance.
compared to the rightmost and the HOV lanes (in general, and in interchange and non-interchange segments).

4.5 Marginal Effects and Elasticities

The LIMDEP software allows for computation of marginal effects for four binary outcome random parameters models in this study. The software computes marginal effects at the means of the random parameters. However, marginal effects do not indicate whether a variable is worth considering as a policy variable. For instance, LIMDEP reported the marginal effect of (partial derivative with respect to) the mean speed parameter as -0.0086. A value of -0.0086 for \( \frac{\partial P(y=1)}{\partial x} \), where \( P \) is the probability of high incoming lane-change ratio and \( x \) is the mean speed, is difficult to interpret and meaningless without prior knowledge of baseline effects. For this reason, it is preferable to talk about elasticities than marginal effects.

In non-linear probability models, generally, elasticity is a function of the probability, the independent variables and the coefficients. Unlike linear models that allow for elasticity to be expressed by a constant term, non-linear relationships make it difficult to interpret the coefficients and express the effect of the change in a variable on the change in probability.

Furthermore, in models such as the random parameters model, varying parameters make measuring elasticity even trickier. In what follows, an attempt to approximate elasticities for the random parameters probit model has been made.
• Step 1: In the original model estimation process, the residuals are recorded as \( RES_{it} \).

• Step 2: Update one particular independent variable \( x_k = 1.01 \times x_k \)

• Step 3: Compute the new residuals \( NEWRES_{it} \) with the updated independent variable

• Step 4: Compute the predicted probabilities in Step 1 and Step 3.

Predicted Probability = Actual Probability – Residual

\[
P_{\text{old},i} = HICHNG_{it} - RES_{it}
\]

\[
P_{\text{new},i} = HICHNG_{it} - NEWRES_{it}
\]

• Step 5: Elasticity is given by the percentage change in the predicted probabilities for a 1% change in the explanatory variable. In this method, elasticity is approximated by averaging over the number of terms in the dataset the following expression:

\[
\frac{(P_{\text{new},i} - P_{\text{old},i})}{P_{\text{old},i}} 
\]

Elasticities were computed by using the above method and presented in Table 4-3.

<table>
<thead>
<tr>
<th></th>
<th>Probit</th>
<th>Logit</th>
<th>Gompertz</th>
<th>Comp Log Log</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Speed Range</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed Class 50-55</td>
<td>0.2933</td>
<td>0.2264</td>
<td>0.2093</td>
<td>0.1554</td>
</tr>
<tr>
<td>Speed Class 55-60</td>
<td>0.2289</td>
<td>0.1307</td>
<td>0.1637</td>
<td>0.0958</td>
</tr>
<tr>
<td>Speed Class 60-65</td>
<td>0.4076</td>
<td>0.1887</td>
<td>0.2906</td>
<td>0.1379</td>
</tr>
<tr>
<td>Mean Speed</td>
<td>-1.3210</td>
<td>0.4502</td>
<td>-0.8968</td>
<td>0.2052</td>
</tr>
</tbody>
</table>
The average elasticities are similar across the four models. Though the estimated coefficients in the logit model were higher than those of the probit model, the magnitude of the average elasticities in the probit model was higher for all the parameters. Elasticities of individual observations can also be interpreted. Of particular interest are the observations where the elasticity is greater than or in the neighborhood of 1 in absolute value. Table 4-4 lists the design vectors that have an elasticity value greater than 1 for the SPRANGE parameter.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>37</td>
<td>9</td>
<td>4</td>
<td>38.9</td>
<td>59.8</td>
</tr>
<tr>
<td>16</td>
<td>20</td>
<td>9</td>
<td>40.1</td>
<td>59.7</td>
</tr>
<tr>
<td>18</td>
<td>11</td>
<td>1</td>
<td>36.1</td>
<td>55.6</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>1</td>
<td>34.1</td>
<td>55.9</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
<td>1</td>
<td>35.2</td>
<td>54.7</td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>9</td>
<td>44.2</td>
<td>61.4</td>
</tr>
<tr>
<td>30</td>
<td>17</td>
<td>6</td>
<td>38.9</td>
<td>61.5</td>
</tr>
<tr>
<td>27</td>
<td>15</td>
<td>9</td>
<td>43.1</td>
<td>57.6</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>32</td>
<td>52.8</td>
</tr>
<tr>
<td>22</td>
<td>12</td>
<td>0</td>
<td>35.2</td>
<td>58.4</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>0</td>
<td>32.1</td>
<td>54.5</td>
</tr>
<tr>
<td>9</td>
<td>16</td>
<td>1</td>
<td>33</td>
<td>54.3</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>3</td>
<td>30.9</td>
<td>52.9</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>53.1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
<td>35.2</td>
<td>52.9</td>
</tr>
<tr>
<td>6</td>
<td>9</td>
<td>1</td>
<td>28.5</td>
<td>59.9</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0</td>
<td>25.5</td>
<td>55.9</td>
</tr>
</tbody>
</table>
It can be seen that the elasticity of speed range is highest in segments having a speed range of 50 to 60 MPH. Additionally, these segments have a mean speed of 30 to 45 MPH. Given the mean speeds are low and volume is high, greater variations in speed allow for more gaps in headway and hence more incoming lane changes.
5.1 Residuals Discussion

In the computation of elasticities in Chapter 4, residuals for the base random effects model were also computed in order to measure the improvement due to the introduction of random parameters. The residuals are listed in Table 5-1.

<table>
<thead>
<tr>
<th></th>
<th>Probit</th>
<th>Logit</th>
<th>Gompertz</th>
<th>Comp Log</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random Effects Model</td>
<td>3.10E-02</td>
<td>3.75E-03</td>
<td>-1.14E-02</td>
<td>7.16E-02</td>
</tr>
<tr>
<td>Final Random Parameters Model</td>
<td>2.64E-02</td>
<td>3.65E-03</td>
<td>-6.22E-03</td>
<td>6.08E-02</td>
</tr>
<tr>
<td>% Improvement</td>
<td>14.71%</td>
<td>2.85%</td>
<td>45.59%</td>
<td>15.13%</td>
</tr>
</tbody>
</table>

There is an improvement in average residuals from the base random effects model to the final random parameters model across all specifications. The Gompertz model has the best improvement in residuals and the logit model has the worst. It has to be noted that the base models themselves have very good residuals. This is because of the many random effects driving the 0/1 dichotomy in $HICNG$ (the binary outcome of high incoming lane-change ratio.) This dichotomy is arbitrary in a way. In the definition of $HICNG$, the threshold for high lane-change ratio was set at 0.5, i.e., $HICNG$ takes the value of 1 if the ratio of incoming lane-change count to the total lane-change count is greater than 0.5. It can be argued that the latent process governing $y^*$ is volatile enough
that a strict threshold of 0.5 might be restrictive. It is quite possible that there is a
threshold range where the random effects are most prominent - say for instance the 0.3 to
0.7 range. Then, two possibilities exist. It maybe that \( y = 1 \) if \( P > 0.3 \) to \( y = 1 \) if
\( P > 0.7 \) may represent behavioral effects in terms of mixing among lanes or it maybe
that very low mixing and very high mixing are insensitive to group specific effects,
meaning the lane position, lane type and interchange effects do not matter as much.

Additionally, the percent improvement in the residuals in the full random
parameters model due the inclusion of the vector of independent variables (\( X \)) could vary
a lot depending upon what was correctly predicted as 1 as well as 0 in both the base
model and the full model. This variation can be visualized in a 2x2 matrix with the
elements: number of correctly predicted 0 probabilities, number of correctly predicted 1
probabilities, number of incorrectly predicted ‘0’ probabilities and number of incorrectly
predicted ‘1’ probabilities. These aggregate matrices could yield useful inferences about
the effect of \( X \).

5.2 Discussion on Selection Effect

The disadvantage of having zero total lane-change counts in the dataset has been
stated in chapter 3. Excluding the zero observations results in a selection effect. This is a
fairly complex issue and can be visualized as a two-level problem. At the top-level is the
binary probability of a section (or time interval) being in the zero lane-change state
versus the non-zero lane-change state. Given that the ratio of incoming to outgoing lane-
changes is measured only for non-zero lane-change states, the lower level of this problem
is the model that is estimated in this study (the sterile model excluding the groups containing zero lane-change observations). The traditional way to estimate this two level model is the James Heckman (1979) approach. In this method, the probability of a section (or time interval) being in the non-zero lane-change state is computed for all the sections, including ones that are in the zero lane-change state. This probability is then factored into the lower level model of high incoming lane-change ratio for the non-zero state groups. The application of this approach in the random parameters framework is yet to be demonstrated.

Some problems arise with the Heckman approach in the traditional sense. First, the Heckman approach assumes that the selection probability (in the top level of the two-level model) is a smooth function. Such functions return probabilities in the entire 0-1 spectrum, causing a zero state section to change to a non-zero state section in the future. In the lane-change counts case, the probability is not as straightforward. Suppose a future simulation results in non-zero lane-changes in a zero-lane-change section. It is most likely that the counts are not greater than 1 or 2 lane-changes. This type of oscillation tends to produce discrete masses of probability, but not a continuous spectra probability (as assumed in the Heckman approach). For instance, in a future simulation, a zero lane-change count section can have either 1 incoming and 0 outgoing, or 1 outgoing and 1 incoming lane-change. This leads to absurd results in the sense that in the former case the $HICHTNG$ ratio is 1 and in the latter case it is 0. Similarly the case of 2 incoming and 0 outgoing, 1 incoming and 1 outgoing, and 0 incoming and 2 outgoing. The Heckman approach is not very useful in this scenario; it appears more like a combination problem.
One possible approach is to estimate say, a bivariate count data model of 0, 1 and 2 counts for incoming and outgoing lane-changes, the coefficients of which determine the total expected lane-change counts. And then, $HICNNG$ can be computed and plugged into the lower level random parameters model. Alternatively, combinatorial possibilities of 0, 1, and 2 lane-changes across the dataset could yield more insights on estimation. The above computational experiments if performed can pave the way for further research in the field.

5.3 Conclusions and Recommendations

Findings of this research validate the effectiveness of micro-simulation as a tool to generate large and detailed datasets for the purpose of microscopic traffic analyses. The complex nature of VISSIM allows the user to incorporate and modify various parameters that concern traffic-flow and traffic-control. The simulation model built for the purpose of this study can be further expanded to include more input parameters such as heavy vehicles, variable speed limits (now operational on I-5), both AM and PM hour analyses, dynamic origin-destination data etc. Further, VISSIM is capable of providing data at a much higher resolution both in terms of time intervals and spatial location of data collection points. The interchange/non-interchange five minute interval level of resolution can further be improved to yield much larger and richer datasets. From the modeling point of view, larger data sets measured at intervals less than five minutes provide for more time dependence across observations in terms of dynamic models with lagged time effects (such has previous period)
Knowledge of lane changing can be highly useful in developing real time adaptive traffic management strategies to manage congestion and emergency situations. In case of incidents and emergency situations where there is a chance of lane blockage, there may be need to force lane-changes at locations upstream of the lane-blocking incident. This can be achieved by factoring into the management of variable speed limit signs the relationships between speed variables and lane change probabilities.

It is hoped that this research can motivate further analysis of microscopic lane change behavior with the help of rich simulated data states and versatile modeling frameworks such as the random parameters framework.
References


Goswami, V. and G. H. Bham, An Empirical Study of Microscopic Lane-Changing Behavior on a Multilane Freeway, International Conference on Applications of Advanced Technologies in Transportation, Chicago, IL, Aug. 2006


Jin, W.L., Macroscopic characteristics of lane-changing traffic, Transportation Research Record: Journal of the Transportation Research Board, 2010


LIMDEP, Econometric Software, 2007


Oregon Department of Transportation, Protocol for VISSIM Simulation, June 2011.


