

1. Report No. <b>SWUTC/16/600451-00036-1</b>		2. Government Accession No.		3. Recipient's Catalog No.	
4. Title and Subtitle <b>REAL OPTION ANALYSIS TO VALUE MANAGED LANES USING BIG DATA</b>				5. Report Date <b>May 2016</b>	
				6. Performing Organization Code	
7. Author(s) <b>Sunghoon Lee</b>				8. Performing Organization Report No. <b>Report 600451-00036-1</b>	
9. Performing Organization Name and Address <b>Texas A&amp;M Transportation Institute Texas A&amp;M University System College Station, Texas 77843-3135</b>				10. Work Unit No. (TRAIS)	
				11. Contract or Grant No. <b>DTRT12-G-UTC06</b>	
12. Sponsoring Agency Name and Address <b>Southwest Region University Transportation Center Texas A&amp;M Transportation Institute Texas A&amp;M University System College Station, Texas 77843-3135</b>				13. Type of Report and Period Covered	
				14. Sponsoring Agency Code	
15. Supplementary Notes <b>Supported by a grant from the U.S. Department of Transportation, University Transportation Centers Program and general revenues from the State of Texas</b>					
16. Abstract <p>Managed lanes (MLs) provide a backup option for travelers even when they do not use MLs. For example, travelers have the option to use MLs when they encounter unexpected congestion. The option value of MLs refers to travelers' willingness to pay for having the ML option for possible use in the future. Despite the potential benefit of MLs, earlier studies have only considered the actual use benefits of MLs, such as travel time savings.</p> <p>This research used detailed travel data from both MLs and general purpose lanes (GPLs) of the Katy Freeway (I-10) in Houston. From these data, revealed preferences between MLs and GPLs of all travelers with a transponder in 2012 were identified. This research examined two potential definitions of travelers who valued MLs as a travel option. These definitions included 1) travelers who used MLs at least once in 2012 and 2) all travelers with transponders (even those who never used MLs). This research found that the travelers who never used the MLs in 2012 were extremely unlikely to use the MLs in all of 2013. Thus, this research recommends ML option users to be defined as only those travelers who used MLs at least once in 2012.</p> <p>This research used the Small-Rosen log sum method (1981) and the Black-Scholes option pricing method (1973) to estimate the option value of MLs. The log sum method estimates the option value by measuring change in consumer surplus between the situation where both MLs and GPLs are available and the situation where only GPLs are available. The Black-Scholes method was originally developed to price options in stock markets and was modified to estimate the option value of MLs in this research. This research found that the log sum method frequently provided a poor estimate of the option value. Thus, this research recommends the Black-Scholes method to estimate the option value of MLs. The option value of the MLs was found to be similar to the value of travel time savings from the MLs for the ML option users. Thus, the option value of MLs is an important component of the total value of MLs.</p>					
17. Key Words <b>Managed Lanes, General Purpose Lanes, Option Value, Value of Travel Time</b>			18. Distribution Statement <b>No restrictions. This document is available to the public through NTIS: National Technical Information Service 5285 Port Royal Road Springfield, Virginia 22161</b>		
19. Security Classif.(of this report) <b>Unclassified</b>		20. Security Classif.(of this page) <b>Unclassified</b>		21. No. of Pages <b>205</b>	
				22. Price	

## DISCLAIMER

The contents of this report reflect the views of the author, who is responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the Department of Transportation, University Transportation Centers Program in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

REAL OPTION ANALYSIS TO VALUE MANAGED LANES USING BIG DATA

A Dissertation

by

SUNGHOON LEE

Submitted to the Office of Graduate and Professional Studies of  
Texas A&M University  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Mark Burris
Committee Members,	Luca Quadrifoglio
	Bruce Wang
	Wei Li
Head of Department,	Robin Autenrieth

May 2016

Major Subject: Civil Engineering

Copyright 2016 Sunghoon Lee

## ABSTRACT

Managed lanes (MLs) provide a backup option for travelers even when they do not use MLs. For example, travelers have the option to use MLs when they encounter unexpected congestion. The option value of MLs refers to travelers' willingness to pay for having the ML option for possible use in the future. Despite the potential benefit of MLs, earlier studies have only considered the actual use benefits of MLs, such as travel time savings.

This research used detailed travel data from both MLs and general purpose lanes (GPLs) of the Katy Freeway (I-10) in Houston. From these data, revealed preferences between MLs and GPLs of all travelers with a transponder in 2012 were identified. This research examined two potential definitions of travelers who valued MLs as a travel option. These definitions included 1) travelers who used MLs at least once in 2012 and 2) all travelers with transponders (even those who never used MLs). This research found that the travelers who never used the MLs in 2012 were extremely unlikely to use the MLs in all of 2013. Thus, this research recommends ML option users to be defined as only those travelers who used MLs at least once in 2012.

This research used the Small-Rosen log sum method (1981) and the Black-Scholes option pricing method (1973) to estimate the option value of MLs. The log sum method estimates the option value by measuring change in consumer surplus between the situation where both MLs and GPLs are available and the situation where only GPLs

are available. The Black-Scholes method was originally developed to price options in stock markets and was modified to estimate the option value of MLs in this research. This research found that the log sum method frequently provided a poor estimate of the option value. Thus, this research recommends the Black-Scholes method to estimate the option value of MLs. The option value of the MLs was found to be similar to the value of travel time savings from the MLs for the ML option users. Thus, the option value of MLs is an important component of the total value of MLs.

## ACKNOWLEDGEMENTS

I would like to especially thank my advisor, Dr. Mark Burris, for his encouragement and guidance to me during the exploration of this research.

I would like to express my gratitude to Dr. Luca Quadrifoglio, Dr. Bruce Wang and Dr. Wei Li for serving on my committee and for their suggestions and support throughout the course of this dissertation.

I am grateful for all funding agencies that have supported me during my studies at Texas A&M University. Specifically, I would like to thank the Southwest Region University Transportation Center, which is funded by a grant from the U.S. Department of Transportation University Transportation Centers Program. I also gratefully acknowledge the kind support from the Transportation Policy Research Center at the Texas A&M Transportation Institute. I also appreciate the sharing of data by TxDOT and Harris County Toll Road Authority (HCTRA) without which none of this effort would have been possible.

I would also like to thank my friends and colleagues in the transportation engineering division, for the happiness and encouragement they have brought to me. I have greatly enjoyed being with them at Texas A&M University.

Most importantly, I am deeply indebted to my wife and family for their love, support, and encouragement.

## NOMENCLATURE

AVI	Automatic Vehicle Identification
BCA	Benefit-Cost Analysis
BT	Benefit Transfer
CE	Choice Experiment
CMM	Choice Modeling Method
CVM	Contingent Valuation Method
CS	Consumer Surplus
G	Generalized Trip Cost
GPLs	General Purpose Lanes
HCTRA	Harris County Toll Road Authority
HOV	High-Occupancy Vehicle
MLs	Managed Lanes
NPTS	Nationwide Personal Transportation Survey
NUV	Non-Use Value
OV	Option Value
RP	Revealed Preference
SAS	Statistical Analysis System
SH 6	State Highway 6
SOV	Single-Occupancy Vehicle
SP	Stated Preference

TEV	Total Economic Value
TT	Travel Time
TxDOT	Texas Department of Transportation
VOT	Value of Travel Time
VTTS	Value of Travel Time Savings
WTA	Willingness to Accept
WTP	Willingness to Pay



## TABLE OF CONTENTS

	Page
ABSTRACT .....	II
NOMENCLATURE.....	IV
TABLE OF CONTENTS .....	VII
LIST OF FIGURES .....	IX
LIST OF TABLES .....	XI
1. INTRODUCTION.....	1
2. OBJECTIVES .....	8
3. LITERATURE REVIEW .....	11
3.1 Classification of Total Economic Value .....	11
3.2 Valuation of Total Economic Value in Transportation .....	16
3.2.1 Basic Principle and Methods of Economic Valuation.....	16
3.2.2 Valuation of Actual Use Value .....	24
3.2.3 Valuation of Option Value .....	27
3.3 Importance of Option Value Estimation .....	34
4. DATA.....	39
4.1 Katy Freeway (I-10).....	40
4.2 Data Sources.....	43
4.2.1 Automatic Vehicle Identification (AVI) Data.....	43
4.2.2 Harris County Toll Road Authority (HCTRA) Toll Data .....	44
4.3 Revealed Preference (RP) Data on Katy Freeway .....	45
4.3.1 Processing Raw Data (AVI and HCTRA Data) .....	46
4.3.2 Identifying Actual Choices (Trips) on Katy Freeway .....	47
4.3.3 Identifying Alternative Choices (Trips) on Katy Freeway .....	50
4.3.4 Cleaning and Correcting Actual and Alternative Choice (Trip) Data.....	56
4.3.5 Final Data set.....	66
4.4 Statistics of Revealed Preference (RP) Data on Katy Freeway .....	67

5. METHODOLOGY AND SAMPLE CALCULATIONS.....	71
5.1 Defining Managed Lane Option Users on the Katy Freeway .....	72
5.2 Utility Functions of ML and GPL Choices on the Katy Freeway.....	76
5.3 Methodology to Estimate the Option Value based on the Log Sum Method .....	80
5.4 Sample Estimations Using the Log Sum Method .....	85
5.5 Methodology to Estimate the Option Value based on the Black-Scholes Method .....	89
5.6 Sample Estimations Using the Black-Scholes Method.....	99
5.7 Comparisons of Sample Estimations Using the Log Sum and the Black-Scholes Methods.....	103
6. OPTION VALUE OF MANAGED LANES ON KATY FREEWAY .....	107
6.1 Option Value of the MLs using the Log Sum Method.....	108
6.2 Option Value of the MLs Using the Black-Scholes Method .....	112
6.3 Comparison of the Option Values of the MLs using the Log Sum and the Black-Scholes Methods.....	117
6.4 Comparison of the Option Values for the Option Users in Case 2 .....	121
6.5 Sensitivity Analysis.....	123
6.5.1 Effect of the Value of Travel Time on the Option Value of the MLs .....	124
6.5.2 Effect of Standard Deviation of Generalized GPL Trip Costs on the Option Value of the MLs.....	127
6.5.3 Comparison of Applying Short-Term and Long-Term Standard Deviations of the Generalized Trip Costs on the GPLs to the Black-Scholes Method .....	129
6.5.4 Effect of the Risk-Free Return on the Option Value of the MLs .....	132
6.6 Comparison of the Option Values with the Value of Travel Time Savings from the MLs .....	133
6.7 Summary of the Option Value of the MLs on the Katy Freeway .....	135
7. CONCLUSIONS .....	143
REFERENCES.....	150
APPENDIX A .....	155
A.1 Descriptions of Transportation Cost Categories (Litman 2009) .....	155
APPENDIX B .....	156
B.1 SAS Program Code to Generate RP Data on Katy Freeway .....	156

## LIST OF FIGURES

	Page
Figure 1 Classification of Total Economic Value (Source: Bateman et al. 2002, selected and modified) .....	12
Figure 2 Demand Curve and Change in Consumer Surplus by Transportation Service Improvement (Source: Small 2007, selected and modified) .....	18
Figure 3 Demand Curve and Change in Consumer Surplus by Availability of MLs .....	20
Figure 4 Katy Freeway (Source: Harris County Toll Road Authority).....	41
Figure 5 Katy Freeway AVI Sensor Locations .....	43
Figure 6 Example of Actual Trip and its Alternative Trip .....	51
Figure 7 Five Sections (where unreasonable alternative trips could be generated) .....	52
Figure 8 Examples of Actual Trips (that include repeated detections) .....	58
Figure 9 Examples of Reverse Direction Trips .....	59
Figure 10 Example of the Actual Trip that Omitted GPL Sensors .....	61
Figure 11 Trip Routes of both Actual and Alternative Trips detailed in Table 12 .....	62
Figure 12 Trip Sections where Travelers would not Consider a Choice between ML and GPL Trips.....	66
Figure 13 Relationship between the option value using the log sum method and the difference between the generalized trip costs on the MLs and GPLs.....	89
Figure 14 Distribution of the Generalized Trip Costs on the GPLs and a Normal Distribution .....	98
Figure 15 Relationship between the option value using the Black-Scholes Method and the standard deviation of the generalized trip costs on the GPLs .....	103
Figure 16 Distribution of the Option Value using the Log Sum Method .....	112
Figure 17 Distribution of the Option Value Using the Black-Scholes Method .....	116
Figure 18 Option value of the MLs depending on the value of travel time .....	126

Figure 19 Total option value of the MLs depending on change in the standard deviation of the generalized trip costs on the GPLs using the Black-Scholes Method .....	129
Figure 20 Total option values of the MLs depending on the risk-free return in the Black-Scholes Method .....	133

## LIST OF TABLES

	Page
Table 1 Comparison of SP and RP Methods (Source: Fujiwara et al. 2011).....	23
Table 2 Transportation Cost Categories (Source: Litman 2009, selected and modified) .....	25
Table 3 Benefit Categories in Managed Lane Benefit-Cost Analysis.....	26
Table 4 Studies on Estimation of Option Value (Source: Chang et al. 2012, selected and modified with additional literature) .....	29
Table 5 Option Values Estimated in Chang (2010) .....	33
Table 6 Option Values Estimated in Chang et al. (2012).....	34
Table 7 Estimates of the Sum of Option and Non-use Values (Source: Laird, Geurs, and Nash 2009, modified to express as US dollar).....	36
Table 8 Size of the Sum of Option and Non-use Values relative to User Benefits and the Present Value of Benefits (Source: Laird, Geurs, and Nash 2009) .....	38
Table 9 Toll Schedules and Rates on Katy Freeway Managed Lanes (Source: Harris County Toll Road Authority).....	42
Table 10 Examples of Alternative Trip Generation in the Data Set .....	54
Table 11 Average Speed on Each Lane of the Freeway by Time of Day (Source: Burris and Danda 2014).....	55
Table 12 Detailed Trip Information of the Example in Figure 10 .....	62
Table 13 Criteria that can Identify Omitted Toll Sensors in Each Alternative Trip .....	64
Table 14 The Number of ML and GPL Trips in 2012 .....	68
Table 15 The Number of ML and GPL Trips by Time of Day in 2012.....	68
Table 16 The Number of ML and GPL Trips by Day of Week in 2012.....	69
Table 17 The Number of ML and GPL Trips by Direction in 2012 .....	70
Table 18 Classification of Katy Freeway Travelers by percentage of ML trips .....	75
Table 19 Summary of Cases 1 and 2.....	76

Table 20 Coefficient Estimates for the Utility Functions of ML and GPL Choices .....	80
Table 21 Correlation Coefficient between Travel Time and the Amount of Toll Paid....	80
Table 22 Three Option Users' Choice Records in 2012 .....	86
Table 23 Option Value for Each GPL Choice (Trip) Using the Log Sum Method .....	88
Table 24 Descriptions of Variables in the Black-Scholes Method .....	91
Table 25 The Number Option Users by their Total Number of ML Trips in 2012 .....	95
Table 26 Values of the Variables in the Black-Scholes Method for Each Choice (Trip)	100
Table 27 Option Value for Each GPL Choice (Trip) using the Black-Scholes Method	102
Table 28 Comparison of the Option Value using the Log Sum Method and the Option Value using the Black-Scholes Method.....	105
Table 29 Option Value of the MLs using the Log Sum Method.....	109
Table 30 Quantiles of the Option Value of the MLs Using the Log Sum Method .....	111
Table 31 Option Value of the MLs using the Black-Scholes Method .....	113
Table 32 Quantiles of the Option Value of the MLs Using the Black-Scholes Method	115
Table 33 Comparison of the Option Values of the MLs in 2012 in Both Methods .....	118
Table 34 The Number of the Poor Estimates of the Option Value in Both Methods ....	121
Table 35 Comparison of the Option Values of the Option Users in Categories a and b	122
Table 36 Option Value of the MLs depending on the Value of Travel Time .....	125
Table 37 Option Value of the MLs Depending on the Standard Deviation of the Generalized Trip Costs on the GPLs Using the Black-Scholes Method .....	128
Table 38 Comparison of Applying the Short-Term and the Long-Term Standard Deviations to the Option Value Estimation in the Black-Scholes Method ....	131
Table 39 Option Value of the MLs Depending on the Risk-free Return in the Black- Scholes Method .....	132
Table 40 Comparison of the Option Values with the Value of Travel Time Savings from the MLs for the Option Users in Case 1 .....	135

Table 41 Summary of the Results .....	137
Table 42 Number of the Option Users who never used the MLs in 2012 by their ML Trip Frequency until 2013 and 2014 .....	139

## 1. INTRODUCTION

The total economic value of an object comprises multiple parts, including actual use and an option value. Actual use value refers to benefits accruing from the actual use of services. Option value refers to the willingness to pay (WTP) for the option of having the service available for possible use at some time in the future, even if the option may never be used (Wallis, Wignall, and NZ Transport Agency 2012). The concept of ‘option value’ is widely applied and is applicable to all public transportation services (rail, bus, etc.) and can be extended to road infrastructure (Geurs, Haaijer, and Van Wee 2006; UK Department for Transport 2009).

The concept of option value for transportation services is fairly straightforward. If travelers have a mode choice, they benefit from having each option available, even if they choose only one mode. For example, even travelers who have never used transit may value the availability of transit service. Specifically, they may value transit as a backup mode if their automobile breaks down because transit can be used to travel instead (Chu and Polzin 1998; Roson 2000; ECO Northwest, Ltd. 2002; Geurs, Haaijer, and Van Wee 2006; Laird, Geurs, and Nash 2009; Chang 2010; Chang et al. 2012).

It is hypothesized that this value can be also applied to managed lanes (MLs). For example, a traveler whose average travel time for work is 30 minutes might sometimes take 60 minutes when using general purpose lanes (GPLs) — as GPLs have a high variability of travel time. As a result, when GPLs are the only option available on the freeway, the traveler has to leave home 60 minutes early to ensure they are not late for



work even if the average travel time is 30 minutes. Generally, MLs provide faster and less variable travel time than GPLs. Assume that MLs can provide an average travel time of 25 minutes with 5 minute variability for the trip on the same route. Therefore, the traveler can now leave home 30 minutes before work when both MLs and GPLs are available. In this case, due to the existence of the MLs, the traveler benefits from spending 30 minutes more for other activities, such as sleeping and morning exercise even when the traveler does not use the MLs. Thus, travelers might favor having a managed lane as an option available and may willingly pay for this benefit despite rarely, or never, using it. Despite this potential benefit of MLs, earlier studies of MLs have only considered the actual use benefits, such as travel time savings and vehicle operating cost savings.

The Katy Freeway (I-10) MLs in Houston provide an excellent opportunity to better understand managed lane travelers, including the option value travelers place on having a managed lane travel choice. After opening the Katy Freeway MLs in 2008 (with tolled access starting in April 2009), the performance of the Freeway, including both the MLs and the GPLs, has been continuously monitored by the Texas Department of Transportation (TxDOT) and the Harris County Toll Road Authority (HCTRA) using automatic vehicle identification (AVI) sensors. This monitoring has resulted in a massive quantity of detailed travel data on the freeway. The available data includes each vehicle's transponder ID, time of travel, section of the freeway where a vehicle passed, and the amount of a toll paid by each vehicle that installed a transponder. Since each traveler's information is identified by the transponder ID, the data gives a unique

opportunity to recognize a traveler's lane-choice decision between MLs and GPLs. Using these data, this research estimates the option value of the Katy Freeway (I-10) MLs. The results of this research add a new dimension to our understanding of the value of these MLs, the option value. In addition, this research develops a new methodology for estimation of the option value specifically as it relates to MLs.

The funding challenge facing transportation investment in the United States is well documented. The investment shortfall results in limited provision of new freeways. As an alternative, MLs are an increasingly popular option for expanding infrastructure (particularly in Texas where 14 MLs are planned), because some portion of the cost needed for the expansion can be covered by tolls. Decision making for providing MLs is partially based on benefit-cost analysis (BCA). The BCA procedure considers only benefits that accrue from the actual use of MLs such as travel time savings and operating cost savings; however, the option value is not considered. It may present a potential problem as the value of MLs is underestimated and makes it harder to justify the construction of needed capacity. Long term, this would result in inaccurate decision making when choosing whether or not to invest in MLs.

Many studies of existing MLs conclude that a vast majority of travelers use MLs only occasionally (Sullivan 2000; Collier and Goodin 2002; Burris and Stockton 2004; Patil et al. 2011). In that case, the benefits of MLs estimated in a conventional BCA procedure may be underestimated because it estimates only the actual use benefits. However, MLs could provide a choice opportunity as a backup (which can be interpreted as an option value) for travelers even when they do not use MLs. MLs offer added

flexibility in travel decisions, such as time of departure. Additionally, travelers have the option to use MLs depending on their situation such as when they encounter unexpected congestion and are in a hurry. Thus, it is obvious that the value of a freeway with MLs is different from the value of a freeway without MLs to a given traveler even if that traveler does not intend to use the MLs. However, this ML option use benefit has not been considered when estimating the benefits of MLs. This may cause an underestimation of ML benefits and hinder their implementation despite data showing how MLs add capacity and in managing traffic congestion.

Most previous research regarding option values in transportation has focused on option value of transit facilities, not the option value of MLs on highways. One reason may be because highways with both GPLs and MLs are relatively new. Also, highways are generally considered as a primary option for travelers. The fact that a highway can now have two classes of service, a primary and optional ‘preferred’ service, changes that perception. Thus, this research first uses the most suitable methodology to evaluate the option value of MLs based on the previous methodology used for transit. The research then develops a new methodology specifically designed for MLs, borrowing from the other economics field, financial economics.

Most previous research evaluates the option value using stated preference (SP) data. This can be collected using contingent valuation methods (CVM) or choice experiments (CE). CVM directly asks people their WTP for a specific transportation service, while CE estimates a value by surveying preferences among several choice alternatives to travelers. SP data is more suitable than revealed preference (RP) data to

examine large changes, such as providing an alternative transit service. Also, transit users' RP data is not able to capture WTP of the option users (potential users) as the option users of transit service are not precisely defined. Implementation of MLs on highways, however, represents a relatively small change in travelers' route choice. In fact it is just a different lane on the existing route. Thus, the option users of MLs can be considered as those travelers who occasionally or never use the MLs but use GPLs on the freeway. Clearly, those could be others, but we expect those to be very few. The data used in this research tracked Katy Freeway travel records of all travelers with a transponder for approximately one year, specifically from January to November in 2012. Thus, such ML users' RP data might be suitable to estimate the option value of MLs and provide more precise results as it provides actual choice results depending on traffic situations, not latent preferences or stated preferences in a survey. This research bases its methodology on such RP data to evaluate the option value of MLs.

In order to estimate the option value of the Katy Freeway (I-10) MLs, this research first identifies revealed preferences of travelers on the Katy Freeway using the AVI sensor data. Since these data only provide actual choices of the travelers and their attributes, such as travel time and the amount of toll paid, this research generates alternative choices that the travelers could have chosen at the time of travel but did not; hence the attributes of those trips not chosen were estimated using data from other travelers who chose the other lanes. By combining the actual and alternative choices and their attributes, each traveler's revealed preference is identified.

Using the revealed preference data, this research then examines the option value of the MLs by measuring a change in consumer surplus between the situation where both MLs and GPLs are available and the situation where only GPLs are available. For a measurement of the change in consumer surplus, Small and Rosen's log sum method is used (Small and Rosen 1981): Most experts believe it to be the most comprehensive method to estimate consumer surplus in discrete choice cases (Cambridge Systematics 1998).

This research also examines the option value of the MLs using the Black-Scholes call-option pricing method. The Black-Scholes method was developed by Black and Scholes (Black and Scholes 1973) and has been widely used to evaluate the value of a call-option in financial economics. A call-option in stock markets gives its holder the opportunity to buy a stock at a lower price, if the price of an underlying stock increases. This relationship between a stock and its call-option is similar to a relationship between GPLs and MLs because MLs provide an alternative lane to travelers alongside GPLs with a lower trip cost if a trip cost on the GPLs increases. Thus, this research applies the Black-Scholes method to estimate the option value of MLs; furthermore, variables in the model are modified according to the needs of this research, which includes an estimation of the option value of MLs.

The methods and results could contribute to a better understanding of the values of MLs. Ultimately, this can greatly impact investment in MLs, and may help allocate scarce transportation funding to this most needed project.

The rest of this research is organized as follows: Section 2 presents the key objectives of this research. Section 3 describes the classification of the total economic value used in this research and reviews the existing efforts on option value measurement in the transportation field. Section 4 identifies travelers' revealed preferences between MLs and GPLs in 2012 using the AVI sensor data on the Katy Freeway. The revealed preferences are used to estimate the option value of the MLs on the Katy Freeway. Section 5 provides the two methodologies to estimate the option value of the MLs and the sample estimations to help better understand applications of the methodologies. Section 6 presents the results and the sensitivity analysis for the option value of the MLs in 2012. Finally, Section 7 presents the conclusions and the discussion of the results in this research and describes possibilities for further research.

## 2. OBJECTIVES

The primary objective of this research is to estimate the option value of MLs on the Katy Freeway (I-10) using a large volume of revealed preference data. The specific objectives of this research are as follows:

- 1) To identify travelers' revealed preferences between MLs and GPLs on the Katy Freeway in 2012 which are needed to estimate the option value of the MLs. The AVI sensor data only provided actual choices and their attributes for all travelers who installed a transponder. Thus, this research also generates alternative choices and their attributes based on other travelers' actual choices' attributes.
- 2) To define option users of the MLs on the Katy Freeway using the revealed preference data. No previous studies were found that defined the option users of MLs. The revealed preference data contains chronological choices (trips) for all of 2012 of each traveler who installed a transponder. Thus, by analyzing each traveler's choices on the freeway, option users were identified from that data to estimate the option value of MLs.
- 3) To identify the utilities of ML and GPL choices on the freeway for the option users defined. The main purpose of economic valuation for a non-market service is estimation of change in consumer surplus. The change in consumer surplus is generally measured by a change in utility caused from using the service. Thus, estimation of the option value of the MLs begins with

identifying the utilities of ML and GPL choices for the option users. Using option users' revealed preferences in 2012, this research identified the utilities of ML and GPL choices. .

- 4) To estimate WTP for the option of having the MLs available (the option value of the MLs). This WTP is quantified by the change in consumer surplus depending on whether the ML option exists or not using Small and Rosen's log sum method (Small and Rosen 1981). For the quantification of consumer surplus depending on whether the ML option exists or not, the utilities of ML and GPL choices for the option users were used.
- 5) To develop an ML specific Black-Scholes call-option pricing method (Black and Scholes 1973). The Black-Scholes method was originally developed to estimate a call-option value in stock markets in financial economics field. Thus, this research appropriately modifies the method to estimate the option value of the MLs. The option value of the MLs is then estimated using the modified method.
- 6) To identify parameter inputs that affect the option value estimates and a range of possible option value estimates in the two methods (the log sum method and the Black-Scholes method). Parameter inputs, such as the value of travel time, can change depending on traffic and economic circumstances and these changes could affect the option value estimates in the two methodologies. Thus, in order to identify the range of possible option value estimates, the sensitivity analysis was conducted.



- 7) To identify the importance of the option value estimation for MLs. The option value estimate in this research cannot be determined whether it is meaningful or not by itself. If the estimate is too small compared to other benefits of the MLs, the estimation of the option value of MLs could only add complexity to the estimation of ML benefits. Thus, this research compares the option value of the MLs with the other benefit of the MLs, travel time savings. By comparing it with the travel time savings, importance of the option value estimation for the MLs is finally identified.

The option value of transportation services has been recently focused in transportation economics and the fact that people are willing to pay for the option of having services available was identified in some previous studies (Chu and Polzin 1998; Roson 2000; ECO Northwest, Ltd et al. 2002; Geurs, Haaijer, and Van Wee 2006; Laird, Geurs, and Nash 2009; Chang 2010; Chang et al. 2012). However, no previous studies were found that estimated the option value of MLs. One reason may be because highways with both GPLs and MLs are relatively new. Also, highways are generally considered the primary option for travelers.

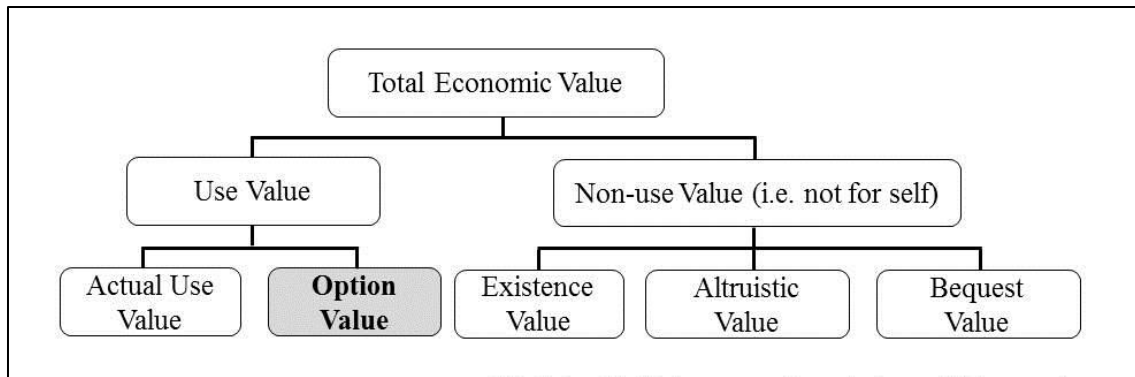
Finally, the results of this research will provide crucial evidence of the option value of MLs. This will add a new dimension to our understanding of their value and will lead to improved transportation planning and BCA for implementation of MLs.

### 3. LITERATURE REVIEW

For the objectives outlined in Section 2, this section first classifies total economic value including the option value, and provides definitions for components of the total economic value. This section also reviews the valuation of total economic value in earlier transportation studies. How benefits of a project or policy are evaluated in transportation field is reviewed in terms of actual use and option values. However, since the main purpose of this research is to estimate the option value of MLs, this review mostly focuses on the option value. In conclusion, based on the existing studies, this section identifies the importance of option value estimation.

#### **3.1 Classification of Total Economic Value**

A transportation project or policy can change each individual's welfare. Transportation appraisal is, therefore, mainly interested in measuring this welfare change for valuation of the project or policy. Each individual's welfare change is mostly measured by WTP, which is based on evaluating each change in monetary terms. WTP refers to the amount of money that an individual would pay for a proposed change. The net sum of all the relevant WTPs defines the total economic value (TEV) of any change in welfare due to a project or policy. The diagram in Figure 1 is the classification of TEV adopted in this research.



Note: The shaded and bolded area is the main focus of this research.

**Figure 1** Classification of Total Economic Value (Source: Bateman et al. 2002, selected and modified)

TEV is first classified into use and non-use values. Use values refer to benefits caused from actual use or possible (optional) use of a transportation service (Bateman et al. 2002). Option value from the possible (optional) use, in the transportation context, is the willingness to pay to preserve the option of using a transportation service for trips not yet expected or currently undertaken by other modes (Laird, Geurs, and Nash 2009). It is fairly obvious that use values include actual use value. However, the issue of whether an option value is included in use values has been discussed and the option value is classified differently in the literature. Early studies usually considered the option value as a component of the non-use value (Bristow et al. 1991; Crockett 1992; Painter et al. 2002; Humphreys and Fowkes 2006). Recent research considers the option value as a component of the use value (Bateman et al. 2002; Geurs, Haaijer, and Van Wee 2006) or independently recognizes the use, option and non-use values (UK Department for Transport 2009; Laird, Geurs, and Nash 2009; Chang 2010; Chang et al. 2012).

Despite different classifications in earlier literature, this research considers option value as use value because individuals are eventually willing to use the option if they encounter specific travel events, even if those events rarely or never occur. Furthermore, this research distinguishes between the option value and actual use value. This is because individuals may perceive these two values very differently and, therefore, separately allocate a WTP for each value (Chang 2010). This option value can be considered from the two perspectives, contingency value and value of choice. Contingency value represents an individual's value of a secondary alternative for possible use in a contingency situation. For example, in unexpected circumstances, such as ice and snow storms, passenger car travelers are able to recognize the value of transit as an option and may use it, even if they do not intend to use it under normal circumstances. Value of choice refers to the value of having the opportunity to choose between alternatives rather than having only a single alternative. Even if this concept is rarely applied in the transportation field, the economic literature has numerous works on the concept in the context of the value of product diversity (Sattinger 1984; Perloff and Salop 1985; Suen 1991; Anderson and Palma 1992; Weitzman 1992). In economic theory, this value is defined for an additional product as change in consumer surplus without changes in prices of related products or services (Suen 1991). If the price of an additional product remains the same, consumer surplus does not change in consumer theory. However, a change in consumer surplus may arise due to the value of choice. This is because the additional product provides an additional opportunity that a new product/service gives to find a product/service that better suits the potential user's

preference (Sattinger 1984; Suen 1991). Therefore, a large set of alternatives is valuable even if a consumer chooses one specific alternative. Furthermore, some consumers have preferences towards diversity itself (Train 1994). Even though the contingency value and the value of choice are conceptually different, in practice, these two values could be the same as the option value. For example, in contingency situations, the motive of the value of choice could be the same as the motive of the contingency value. Thus, this research does not consider these two values as different types of option value. Note that, even if this research considers them as the same option value, the option value estimated from the first methodology (the Small-Rosen log sum method) in this research is based on the concept of the value of choice, while option value estimated from the second methodology (the Black-Scholes call-option pricing method) is based on the concept of the contingency value.

Non-use value represents willingness to pay to maintain some services even if there is no actual or possible (option) use (Bateman et al. 2002). This includes various types of non-use values. A convenient classification is in terms of (a) existence value, (b) altruistic value, and (c) bequest value (although in practice, it is usually not possible to disaggregate individual types of non-use value). Existence value is defined as WTP to maintain a good in existence in a context where the individual has no actual use for himself/herself or for anyone else. Motivation of this value includes having a feeling of concern for the asset itself (e.g. endangered species) or a 'stewardship' motive, which assumes some responsibility for the asset (Bateman et al. 2002). Pure existence value is unlikely to occur from transportation infrastructure (Geurs, Haaijer, and Van Wee 2006).

Altruistic value arises when individuals are willing to pay to preserve a service that should be available to others in the current generations, such as relatives and friends.

Bequest value is similar to altruistic value but WTP to preserve a service is for the next and future generations (Bateman et al. 2002). Transportation infrastructure could provide these two non-use values because individuals may be willing to pay to construct transportation infrastructure for others. For example, individuals may value transit for the sake of its future availability to their children if the children are unable to drive and are willing to pay to preserve transit, even if they do not use it themselves.

Based on the definitions above, actual use value of MLs can easily be recognized as benefits to be gained from actual use of MLs, such as travel time savings and operating cost savings. Option value of MLs can be considered from two perspectives. One is related to contingency value. If travel time variability on GPLs is large, travelers who primarily use the GPLs have to consider this variability in time and leave earlier than their preferred travel time. However, if MLs exist on the same route with the GPLs, irrespective of use of the MLs, those travelers can leave at their preferred time because MLs generally provide a more reliable travel and allow the travelers a backup option if they are faced with delay on the GPLs. The other is related to value of choice. That is, travelers' welfare change that can result from having MLs as choice rather than only GPLs is also part of the option value of MLs. This choice opportunity ultimately provides a backup option for the travelers if they encounter situations where they need a contingent option.

Pure existence value is unlikely to exist for MLs because MLs are transportation infrastructure. However, MLs can provide altruistic and bequest values. GPL users may favor having MLs for others or the next generations, even though they will never use MLs. Whereas, from the perspective of a BCA, some double-counting may occur with the inclusion of the altruistic and bequest values of MLs, especially when their motives are related to personal gain/loss. For example, when vehicles on the GPLs change their route choice decisions to MLs, travel time savings on the GPLs are possible due to the reduced traffic volume. If those GPL users favor having MLs to save their travel time on the GPLs at present or in the future as in that case, there will be double-counting of altruistic and bequest values with actual use value. This research, therefore, estimated only the option value of MLs, and excluded non-use value.

### **3.2 Valuation of Total Economic Value in Transportation**

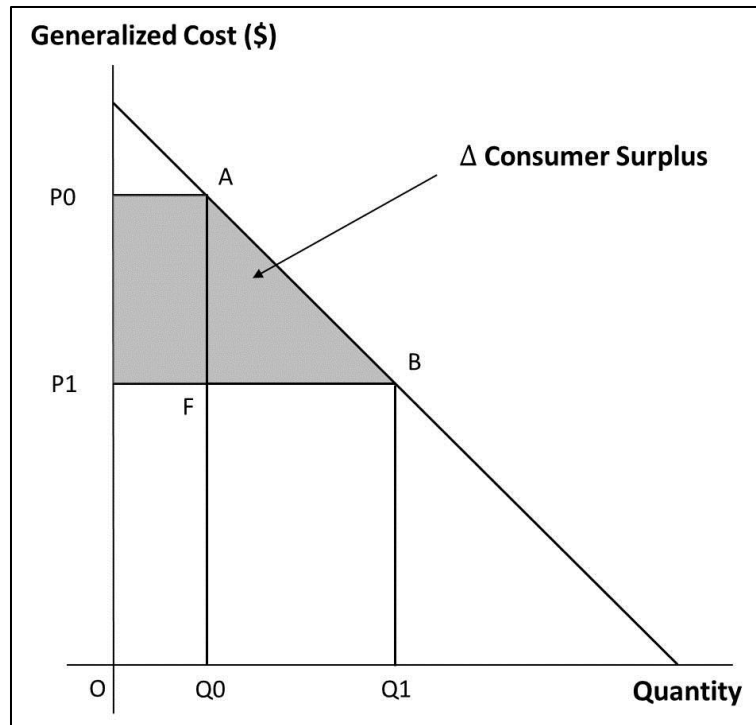
This section reviews the valuation of total economic value in transportation research. This section first outlines the basic principle of economic valuation. Based on this principle, this research then identifies the option value of MLs. How benefits of a project or policy are evaluated in transportation is also reviewed in terms of actual use and option values. However, since the main purpose of this research is to estimate the option value of MLs, this review mostly focuses on the option value.

#### *3.2.1 Basic Principle and Methods of Economic Valuation*

Benefits are generally explained in terms of individual preferences. To measure the value of such a benefit, we usually measure how much a person is willing to give up to get the benefit. This represents an important characteristic of economic valuation:

There is no absolute measure of value because there exists only relative values between one thing and another. Thus, the measure simply uses a relative valuation which is revealed in individual preferences. This approach enables a meaningful measurement of all the benefits in a single dimension were we to adopt one specific type of benefits as a standard. We would then express all other benefits in terms of a specific standard using an individual's own preferences. Conventionally, for the standard of measurement in economics, money is used (Bateman et al. 2002). If money is adopted as the standard to measure welfare, the measure of benefit is willingness to pay (WTP) that is the amount of money a person would pay to gain the benefit. Furthermore, a US government report guidance (US Department of the Treasury 2003) says that benefits of non-market goods, such as public goods, are generally best evaluated in terms of people's WTP for marginal changes in supply. WTP has a relationship with a demand curve. Figure 2 presents the usual depiction of a demand curve for bus trips by users with generalized cost of travel. Suppose that the generalized cost of bus trips decreases from  $P_0$  to  $P_1$  due to a service improvement. For existing users at  $Q_0$ , change in consumer surplus due to the improvement is the area  $P_0AFP_1$ . New users,  $Q_1-Q_0$ , are also attracted by the improvement and their consumer surplus is the area  $ABF$ . Thus, total change in consumer surplus due to the improvement is the area  $P_0ABP_1$ . Change in consumer surplus (CS) represents WTP to obtain non-market services. Thus, the total WTP for the improvement from both existing and new users is also the area  $P_0ABP_1$  (Small 2007).





**Figure 2** Demand Curve and Change in Consumer Surplus by Transportation Service Improvement (Source: Small 2007, selected and modified)

The option value of MLs can also be understood by a depiction of a demand curve for GPL trips by users with generalized costs of travel. Suppose that demand elasticity of the GPL trips with respect to the generalized costs of travel is  $e_T$ . In general, demand elasticity is affected by availability of substitutes. That is, the more substitutes, the more elastic the demand will be (“Economics Basics: Elasticity” 2013). For example, if the price of orange juice went up by \$0.25 (25 cents), consumers could replace their juice with apple juice (or other substitutes). However, if the apple juice is not available for the consumers, they could be less elastic in the increase of the orange juice price because there is no substitute for the apple juice. As in the example, availability of an

ML trip could affect  $e_T$  because the ML trip can be a substitute for a GPL trip.  $e_T$  is calculated as:

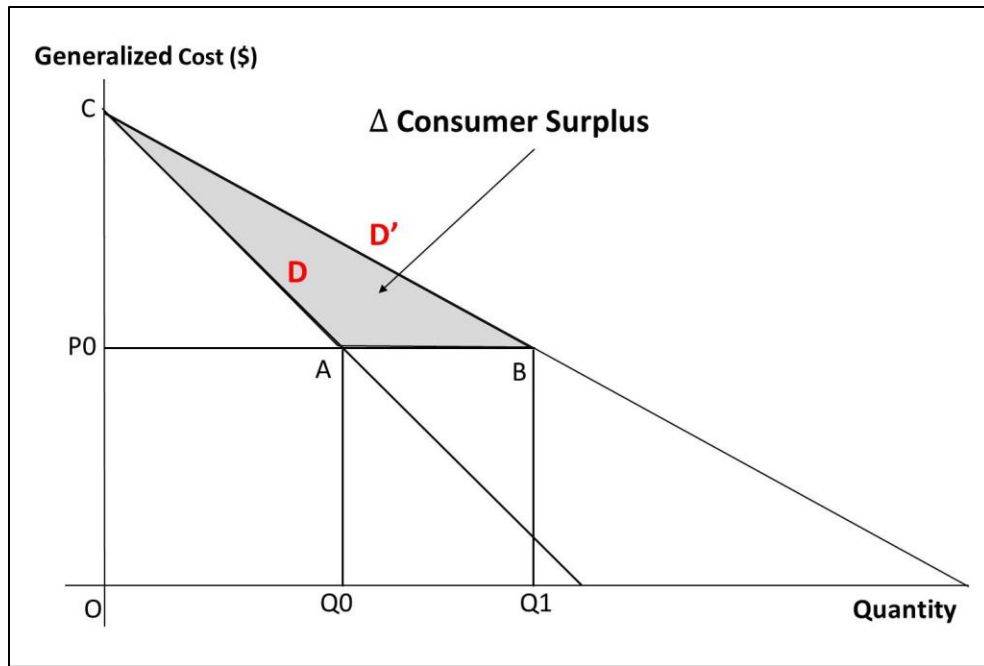
$$e_T = \frac{\Delta Q}{\Delta P} \times \frac{P}{Q} \quad (1)$$

where: P = the generalized cost of travel on GPLs, Q = quantity.

The first term,  $\frac{\Delta Q}{\Delta P}$ , is a reciprocal of slope of the demand curve. Thus, we are able to infer the following relationship:

$$e_T \propto \frac{1}{\text{slope of demand curve}} \quad (2)$$

Figure 3 depicts the demand curve for GPL trips by users with generalized cost of travel. Assume that the demand curve is D when GPLs are only available. The availability of MLs could increase  $e_T$ . Since  $e_T$  is the reciprocal of the slope of a demand curve, the slope of the demand curve (D) decreases to D'. When the generalized GPL travel cost is P0, and the GPLs are only available, consumer surplus is the area P0AC. If the MLs become available without change in the generalized GPL travel cost, P0, consumer surplus becomes the area P0BC due to the decrease in the slope of the demand curve from D to D'. Thus, change in consumer surplus is caused from the availability of MLs is the area ABC. From the change in consumer surplus, this research is able to identify the option value of MLs.



**Figure 3** Demand Curve and Change in Consumer Surplus by Availability of MLs

The main purpose of economic valuation for non-market goods and services is estimation of change in consumer surplus in terms of net change in wellbeing (or welfare, or utility). This change in consumer surplus represents WTP to obtain non-market goods and services. To estimate WTP for non-market goods, three broad categories of economic valuation are widely applied: stated preference (SP), revealed preference (RP), and benefit transfer (BT) which depends on estimates from RP and/or SP studies.

SP methods uncover WTP by directly asking for an individual's WTP, such as 'What are you willing to pay for a service?' or by asking for respondents' preferences across some set of alternatives. These methods are used when the relevant information cannot be inferred from markets or from individual decisions. Since many transportation

services do not have markets, these methods are frequently used to estimate the benefits. SP methods are also used to simulate individuals' behavior in the potential market place.

SP methods are classified into contingent valuation methods (CVM) and choice modeling methods (CMM). CVM focuses on the valuation of a non-market good (product) or service as a whole while CMM focuses on valuing specific attributes of a non-market good or service. CVM presents a hypothetical situation to respondents and directly infers respondents' WTP for that situation. The question for this method can be presented in various ways, including the open-ended question, bidding game, payment card, and dichotomous choice elicitation formats (Fujiwara et al. 2011). To estimate values of specific attributes, CMM questionnaires present respondents with a series of alternative descriptions of a good or service. Choice modeling methods (CMM) include contingent ranking, choice experiment (CE), contingent rating, and paired comparison methods (Hanley, Mourato, and Wright 2001).

RP methods also elicit WTP by using evidence on how people behave in the face of real choices. These methods are used when relevant WTP information can be inferred from an individual's actual decisions (Bateman et al. 2002). For example, if a traveler pays  $x$  dollars for a specific transportation service, WTP for the service can be inferred as at least  $x$  dollars. There are several methods, including the hedonic pricing method, travel cost method, discrete choice models (including random utility model), damage costs, control or prevention costs, and compensation rates. Among these methods, the common RP methods are:

- Hedonic pricing method: This method is based on the premise that non-market goods or services affect the price of market goods in well-functioning markets. Thus, this method elicits values for non-market goods or services from their effect on market prices, such as housing values and wages. For example, if prices of houses on heavy traffic streets are lower than prices of other comparable houses on low traffic streets, the cost of traffic can be estimated according to its effect on real estate value. This is an example of the hedonic pricing method (Litman 2009).
- Travel cost method: This method uses visitors' travel cost (monetary expenditures and time) to measure consumer surplus provided by a recreational site such as a park (Litman 2009).
- Discrete choice models: These methods are based on the premise that choices between alternative options reflect the wellbeing (utility) that accrues from those options (Bateman et al. 2002).
- Random utility model: In this method, individual's choices are expressed as the probability of choosing one option rather than another. The choices define the random utility model, and then utilities of the choices are indirectly measured. This model underlies discrete choice models (Bateman et al. 2002).

Table 1 compares SP and RP methods. Both can be used to estimate use value, and either method can be selected depending on whether information (from proxy markets) is available in practice. As defined in the previous section, this research defines the option value as the use value. Thus, both methods can be applied to estimate the

option value of MLs. However, note that non-use value can only be estimated by SP methods (Bateman et al. 2002).

**Table 1** Comparison of SP and RP Methods (Source: Fujiwara et al. 2011)

	Stated Preference (SP)	Revealed Preference (RP)
Advantages	<ul style="list-style-type: none"> <li>• Wide application and specific valuations</li> <li>• Allow one to explore the reasons behind preferences</li> <li>• Ex-ante application</li> <li>• Widely used and researched</li> <li>• Relatively easy to describe and explain</li> </ul>	<ul style="list-style-type: none"> <li>• Estimates based on real economic choices</li> <li>• Cost-effective</li> </ul>
Disadvantages	<ul style="list-style-type: none"> <li>• Hypothetical bias: non-commitment bias and strategic bias</li> <li>• Protest valuations (e.g., a zero WTP for an ethical objection)</li> <li>• WTP – WTA<sup>1</sup> disparity</li> <li>• Costly</li> <li>• Survey-related biases: information bias, interviewer bias, non-response bias, fatigue, and frustration</li> </ul>	<ul style="list-style-type: none"> <li>• Market imperfections</li> <li>• Measuring WTP for non-marginal changes (e.g., hedonic pricing method)</li> <li>• Marshallian and Hicksian demand</li> </ul>
Methods	<ul style="list-style-type: none"> <li>• Contingent valuation methods</li> <li>• Choice experiments</li> </ul>	<ul style="list-style-type: none"> <li>• Hedonic pricing method</li> <li>• Travel cost method</li> <li>• Discrete choice models</li> <li>• Random utility model</li> <li>• Damage costs</li> <li>• Control or prevention costs</li> <li>• Compensation rates</li> </ul>

---

<sup>1</sup> Willingness to accept (WTA) refers to the minimum amount of money that would need to be compensated for foregoing a good (Fujiwara et al. 2011).

The last category of economic valuation is benefit transfer (BT). BT refers to the process of taking information about economic values from one context (the ‘study site’) and applying it to another context (the ‘policy site’) (Fujiwara et al. 2011). Obviously, both SP and RP methods are resource-intensive because they require the collection of considerable samples. Once a reasonable number of studies has been conducted, it may be possible to understand uniformity or to find some rules in estimates of WTP. This feature in the theorem of BT enables evaluation of WTP without carrying out another survey. In BT, an average WTP estimate from one primary study, WTP estimates from a meta-analysis of several studies, or parameters in a WTP function from one site can be transferred to a new site (Bateman et al. 2002). There is still no previous research related to estimation of option value of MLs. Thus, BT cannot be applied in this research.

### *3.2.2 Valuation of Actual Use Value*

Actual use value of transportation services is accrued from benefits that are obtained from actual use of the services. A benefit can be defined as a reduction in costs, and a cost can be defined in terms of reduced benefits.

**Table 2** Transportation Cost Categories (Source: Litman 2009, selected and modified)

Internal/External	Cost	Market/Non-Market
Internal	<b>Vehicle ownership</b>	Market
	<b>Vehicle operating</b>	Market
	<b>Internal: parking</b>	Market
	<b>Travel time</b>	Non-market
	<b>Internal: crash</b>	Non-market
	<b>Internal: activity</b>	Mixed
External	Operating subsidies	Market
	External: parking	Market
	Road facilities	Market
	Roadway land value	Market
	Traffic services	Market
	Transportation diversity	Non-market
	Air pollution	Non-market
	Noise	Non-market
	Barrier effect	Non-market
	Water pollution	Non-market
	Waste disposal	Non-market
	Congestion	Mixed
	Greenhouse gas	Mixed
	Resource consumption	Mixed
	Land use impacts	Mixed
	External: crash	Mixed
	Healthful activity	Mixed

Note: Boldfaced items under Cost (Internal) are considered as the benefits that can be counted as having actual use value. Descriptions for each category are provided in Appendix A.

Litman (2009) summarized transportation cost categories which were studied in previous literature (see Table 2). These categories can be classified as internal and external costs. Users endure internal costs directly and external costs are imposed on non-users. Therefore, reduced internal transportation costs (i.e., vehicle ownership,



vehicle operating costs, internal parking costs, travel time, internal crash costs, and internal activity) can be considered as benefits obtained from the actual use of transportation services. Furthermore, each cost is divided into market and non-market costs. Market cost is associated with goods traded in the competitive markets, such as vehicles and fuel. Non-market cost is associated with goods that are not regularly traded in markets, such as clean air, crash risk, and noise.

Most highway cost allocation and investment evaluations mainly focus on internal costs such as travel time, vehicle operating costs, and crash damages. External and non-market costs tend to be undervalued because they are difficult to measure (Litman 2009). Table 3 provides benefit categories that were considered in BCA studies for MLs. Table 3 also shows that the BCA studies have been mainly focused on estimation of the actual use values (reduced internal transportation costs).

**Table 3** Benefit Categories in Managed Lane Benefit-Cost Analysis

	Burris and Sullivan (2006)	Sullivan and Burris (2006)	Burris and Patil (2009)	Sisiopiku et al. (2009)	Parsons and Brinckerhoff (2010)	SRF Consulting Group (2013)
Travel Time Savings	✓	✓	✓	✓	✓	✓
Vehicle Operating Cost Savings	✓	✓		✓	✓	✓
Safety Benefits				✓	✓	✓
Emission Reduction	✓	✓		✓		

### *3.2.3 Valuation of Option Value*

Only a few studies on option value of transportation services can be found in the literature. In addition, all tend to estimate the option value of public transit, ignoring other travel modes. This may be because, as mentioned in Section 1, private travel modes are generally recognized as a primary option for travelers, while the option value concept has been applied to backup options.

This section separates previous studies on option value in the transportation field based on the two economic valuation methods, RP and SP methods; moreover, no previous studies were found that adopted the benefit transfer (BT) method. Table 4 provides a brief overview of the literature. Most research adopted SP methods because transit service option users' choices are not fully revealed in RP data. Option users of transit are the users who potentially intend to use transit in situations where their primary mode, such as an automobile, is not available. Thus, it is impossible to identify who they are, and their intention cannot be captured. However, most option users of MLs could be those who occasionally or never use the MLs but use GPLs on the freeway. Thus, if such users' trip decisions were continuously documented, RP methods could be applied to estimate the option value of MLs. As can be seen from Table 4, most research is restricted to the estimation of option value for local service, and not for nationwide long distance services. To estimate the option value of transit, the catchment areas and affected travelers first need to be defined. However, for long distance services, it is difficult to precisely define these travelers and areas because of the uncertainty in impacts of such services. This research was, therefore, restricted to estimating the option

value of a specific segment of MLs, specifically the Katy Freeway (I-10) MLs. Finally, it must be noted that compared to SP data, RP data provides numerous samples (see Table 4) especially when sample size increases, and bias between a true estimate and a sample estimate approaches zero (Montgomery 2007a). Thus, it can be expected that RP data provides a more reliable result than SP data.

### **Valuation of Option Value Based on Revealed Preference (RP) Method**

Chu and Polzin (1998) quantified option value of public transit in the context of the logit mode choice model. Their methodology was based on Small and Rosen (1981), who measured consumer surplus when choice situations that consumers faced were a discrete choice. The Small–Rosen log sum method makes it possible to measure how consumer surplus changes due to variations in price, quality, and the number of options available. This methodology is regarded by most experts as the most comprehensive method to measure benefits. It is also preferred as the method that avoids double-counting (Cambridge Systematics, Inc. 1998). The Small-Rosen methodology was then applied to daily personal travel in the United States using the 1995 Nationwide Personal Transportation Survey (NPTS). Average (aggregate) values of each mode travel were used to estimate the generalized travel costs instead of disaggregate personal data. The estimated US total value of choice was then compared with the total cost of public transit including operation and maintenance costs as well as amortized transit capital investments. Option value varies from 0 cents per person-trip (for trips with no public transit available) to 70 cents depending on modes available to travelers—the average value is 5 cents for every daily person-trip. These unit option values totaled 18.6 billion-

**Table 4** Studies on Estimation of Option Value (Source: Chang et al. 2012, selected and modified with additional literature)

	Chu and Polzin (1998)	ECO Northwest, Ltd et al. (2002)*	Bristow et al. (1991)	Crockett (1992)	Painter et al. (2002)	Humphreys and Fowkes (2006)	Geurs, Haaijer, and VanWee (2006)	Chang (2010)	Chang et al. (2012)
Estimated value	Option value	Option value	Option and non-use values**	Option and non-use values**	Option and non-use values**	Option value	Option value	Option value	Option value
Methodology	RP (change in CS)	RP (Black–Scholes call-option pricing)	Closed-ended CVM (bidding)	Closed-ended CVM (bidding)	Open-ended CVM (direct question)	Open-ended CVM (payment card) and CE	CE	Closed-ended CVM (dichotomous choice)	Closed-ended CVM (dichotomous choice)
Mode	Bus, trolley and rail	All transits	Local buses	Railways that connect rural to urban	Rural buses	Commuter railways	Railways that connect rural to urban	Intercity railways	Intercity, metropolitan, and urban bus services
Study area type	National (U.S.)	NA	Local	Local	Local	Local	Local	National (Korea)	Local and major corridors in Korea
Unit of analysis	Average of individuals	Individual	Household	Household	Individual and household	Household	Individual	Individual	Individual
Sample size	95,360	NA	60	34	170	178	2,665	242	3,000
Base year	1995	NA	1990	1992	1999	2002	2004	2008	2009
Currency	US dollar	NA	UK pound	UK pound	US dollar	UK pound	Euro	Korean won***	Korean won***
Option value	0.09–0.70	NA	NA	NA	NA	154	94	9.3–22.8	3.0–16.20
Option + non-use values	NA	NA	58	36	56	190	242	17.0–38.9	5.3–31.6

\*: Since ECO Northwest, Ltd et al. (2002) only provides a methodology and its example, specific estimates are not provided.

\*\*.: Separate estimates for option values are not available in Bristow et al. (1991), Crockett (1992), and Painter et al. (2002) because they included option value in non-use value.

\*\*\*: 1 Korean won is approximately the same as 0.001 US dollar.

Note: The values are expressed as Korean won per kilometer and per hourly number of services in Chang (2010) and Chang et al. (2012), the values in Chu and Polzin (1998) are expressed as US dollar per person-trip, and the others are represented by their own currencies per the unit of analysis per year.

dollars in 1995 in the United States. The comparison results indicate that the aggregated option value is two-thirds the total cost of transit service, which equaled about 26 billion dollars in 1995.

The book, “Estimating the benefits and costs of public transit projects: a guidebook for practitioners” published by the National Academy Press (ECO Northwest, Ltd. 2002) used the Black-Scholes formula to estimate the option value of transit by applying it to RP data. The authors who put together this book helped add to a series of reports. The report referred to here was the Transit Cooperative Research Program, 78 (TRCP 78), which indicated that the option value of transit to an automobile user is very much akin to the call-option. In financial markets, explicit agreements that offer traders the option to buy a security at a set future date and price is referred to as a “call-option.” This helps to cover uncertainty in stock prices. The risk of loss lies in volatility; hence, the cost of automobile travel can suddenly become very expensive due to bad weather, breakdowns, accidents, etc. Transit provides a way to buy alternative services. From this similarity, in order to estimate transit’s option value, the TRCP 78 report (2002) recommended the Black–Scholes call-option pricing formula to evaluate financial option prices. A detailed explanation about this formula is provided in Section 5.4. The TCRP 78 report (2002) does not examine the option value of a specific transit service, but provides a simple example of transit’s option value. The provided example gives two meaningful results. First, the higher the cost of transit travel, the lower is the option value of the transit. This is because the high cost of the transit travel acts as a barrier against the use of transit. Second, the higher the volatility of automobile travel costs, the

higher the option value of transit. This is because this high volatility increases the chances that transit may be needed. Based on these results, a higher option value of MLs is expected:

- when the cost of managed lane travel is low; and
- when variability (volatility) in travel time on GPLs is high.

### **Valuation of Option Value based on SP Method**

Based on SP data, the option value of transit was identified in Humphreys and Fowkes (2006), Geurs, Haaijer, and Van Wee (2006), Chang (2010), and Chang et al. (2012).

Humphreys and Fowkes (2006) provide an empirical study, which separates the different components of the total economic value of transit. The option value of the Edinburgh to North Berwick rail service in Scotland was estimated. They used the contingent valuation method (CVM) to measure the train users' consumer surplus. Using a payment card method, respondents were directly asked the maximum amount of money that they would be willing to pay in order to ensure the continued availability of the rail service. The option value was established per household at between 150 British pounds and 172 British pounds per year, with a weighted average value of 154 British pounds per year.

Geurs, Haaijer, and Van Wee (2006) estimated the option value of regional rail links in the Netherlands. Through an Internet-based survey, they conducted a choice experiment (CE) to quantify WTP for the option use of specified rail links. Two study areas were selected: the Arnhem–Winterswijk light rail link and Leiden–Gouda railway

link. The results showed that rail service users were willing to pay a large amount for the continued availability of the railway links. An average option value of the rail users, in both cases, was 9 Euros per person per month. Car drivers also expressed WTP for the use of train in case of unforeseen situations, such as breakdown of their car. The option value was estimated at between 11 Euros and 14 Euros per person per month. These two studies, Humphreys and Fowkes (2006) and Geurs, Haaijer, and Van Wee (2006), conclude that travelers are willing to pay to maintain rail services irrespective of their actual use.

Chang (2010) estimated the option value of the two types of intercity rail services in Korea; high-speed rail and conventional rail services. Chang (2010) also defined option value as a traveler's WTP for reserving rail services, similar to the two previous studies (Humphreys and Fowkes 2006; Geurs, Haaijer, and Van Wee 2006). A CVM based on double-bounded dichotomous choice was conducted. Depending on the primary travel mode, respondents were classified into four user groups: auto, bus, high-speed rail, and conventional rail users. Table 5 tabulates the estimation results. Auto users seem to favor the rail services for their unexpected situations more so than transit users. This may be due to the higher uncertainty in auto travel than in public transit travel. In this research, auto travel on GPLs also poses uncertainty, such as fluctuation of travel time due to unexpected congestion. Thus, this uncertainty could also be the reason for such option values of MLs.

**Table 5** Option Values Estimated in Chang (2010)

User group by primary travel mode	Option value of high-speed rail	Option value of conventional rail
Auto	0.0214	0.0155
Bus	0.0120	0.0087
High-speed rail	-	0.0146
Conventional rail	0.0143	-

Note: Values are expressed by US dollar per kilometer per hourly number of services (\$/km/the number of services per hour). This research converted Korean won currency to USD currency using the ratio of 1065 to 1 which is based on the foreign exchange rate in November, 2013.

Chang et al. (2012) evaluated values of three types of bus services in the Seoul metropolitan region and in three major corridors. These were urban, metropolitan and intercity bus services. CVM based on dichotomous choice survey was conducted to quantify the option values. As in Chang (2010), depending on users' primary travel mode, user groups were categorized into:

- auto, high-speed rail and conventional rail users for the intercity bus service;
- auto and metropolitan rail users for the metropolitan bus service; and
- auto, taxi, and urban rail users for the urban bus service.

Table 6 contains the option value estimates of the three services in Chang et al. (2012). As in Chang (2010), auto users are willing to pay more for having backup bus services than users of other modes.



**Table 6** Option Values Estimated in Chang et al. (2012)

Target for the option value	User group by primary travel mode	Option value
Intercity buses	Auto	0.0152
	High-speed rail	0.0129
	Conventional rail	0.0107
Metropolitan buses	Auto	0.0080
	Metropolitan rail	0.0059
Urban buses	Auto	0.0033
	Urban rail	0.0028
	Taxi	0.0030

Note: Values are expressed by US dollar per kilometer per hourly number of services (\$/km/the number of services per hour). This research converted Korean won currency to USD currency using the ratio of 1065 to 1 which is based on the foreign exchange rate in November, 2013.

### 3.3 Importance of Option Value Estimation

In the previous sections, the concept of option value (OV) is outlined. This section, therefore, identifies the importance of option value estimation in terms of its amount.

Based on the literature review, option value of transportation services has not been an active research area in the transportation field. As a result, only a few estimates of the option value were found in the literature. Furthermore, even these estimates are for option values of public transportation, not for highway facilities. Laird, Geurs, and Nash (2009) compared five estimates of combined option and non-use values (NUV) (see Table 7). The reason for comparison of the combined values is that, depending on the studies, different use/non-use classifications were used and three of the studies (Bristow et al. 1991; Crockett 1992; Painter et al. 2002) have not separately identified the option value. To compare those estimates, they also converted the estimates to a

common currency in year 2002 dollars. In Table 7, the combined values vary from 66 dollars to 304 dollars per population unit per year. There is no obvious pattern for the values. However, the authors view that population unit of each combined value (individual or household) and the availability of alternative transportation services would cause the differences in the values. Differences between bus and rail would also contribute to the differences in the values, rather than uncertainty in the values (Laird, Geurs, and Nash 2009). Humphreys and Fowkes (2006) found that the non-use value comprised 25 percent of the sum of the option and non-use values while Geurs, Haaijer, and Van Wee (2006) found that it comprised 40–60 percent. If these percentages are applied to estimate the range of the option values for the public transportation, the option values may range from 26 dollars to 228 dollars per household per year.

**Table 7** Estimates of the Sum of Option and Non-use Values (Source: Laird, Geurs, and Nash 2009, modified to express as US dollar)

OV+NUV	Painter et al. (2002)	Bristow et al. (1991)	Crockett (1992)	Geurs, Haaijer, and Van Wee (2006)	Humphreys and Fowkes (2006)
Survey year	1999	1990	1992	2004	2002
Population unit	Mixture of household and individual	Household	Household	Individual	Household
Mode	Bus	Bus	Rail	Rail	Rail
Alternative public transportation	No	No	Existing bus service and alternative rail line/train station	No	Half hourly bus service
Converted values as GBP currency (2002 pounds)	£41	£104	£59	£125	£190
Converted values as USD currency* (2002 dollars)	\$66	\$166	\$94	\$200	\$304

\*: This research converted GBP currency to USD currency using the ratio of 1.0 to 1.6 which is based on the foreign exchange rate in October, 2013. The converted values imply the value per the population unit per year.

Table 8 provides the size of total option and non-use values relative to user benefits for five rail schemes in Laird, Geurs, and Nash (2009). The percentages of the values relative to user benefits range from 9 percent to 561 percent. To clarify why these percentages so greatly varied, the area type and transit service features in each area need to be considered. In Table 8, three schemes (Highland Rail Developments 2000; Halcrow 2006; Highland Rail Partnership 2003) serve remote and sparsely populated areas while the other two schemes (Jacobs 2006; UK Department for Transport 2006) are for areas near large metropolitan areas. In the sparsely populated areas, generally, the rail network is lightly used, which means user benefits are low. Consequently, the high

percentages were provided. One extreme case, Cononbridge station re-opening where the option and non-use values are almost six times the user benefits, also resulted from the similar reason that user benefits are low, rather than option and non-use values being high (Laird, Geurs, and Nash 2009). From these results, inclusion of the option in a BCA value would substantially affect judgement about whether construction of MLs is economically beneficial, especially for the MLs where user benefits are low.

Currently, it is uncertain how large option values of MLs will be. However, based on the previous research reviewed in this section, it is clear that alternative transportation services have option values and the option value of MLs should be investigated.

**Table 8** Size of the Sum of Option and Non-use Values relative to User Benefits and the Present Value of Benefits (Source: Laird, Geurs, and Nash 2009)

Study	Scheme	Area type	# of households affected	Annual passengers on line (single trips in opening year)		Option and non-use values as percentage of	
				Do nothing	Do something	Transport user benefits (%)	Present value of benefits (%)
Highland Rail Developments (2000)	Beaully station re-opening	Remote community in North Scotland	550	125,000	148,000	87	84
Halcrow (2006)	Cononbridge station re-opening	Remote community in North Scotland	1,000	250,000	270,000	561	117
Highland Rail Partnership (2003)	Internet-provision of services	Remote communities in North Scotland	2,600	485,000	557,000	64	57
	(a) Increase in service frequency by 50%		700	145,000	210,000	23	20
	(b) Provision of commuter service		1,900	340,000	347,000	197	178
Jacobs (2006)	Airdrie–Bathgate line re-opening	Small to medium-sized communities within commuting distance of large conurbations	7,400	0	4,000,000	Not known	4
UK Department for Transport (2006)	Rail closure	Rural communities in a part of southern England near to some medium conurbations	3,700	3,798,000	0	9	9

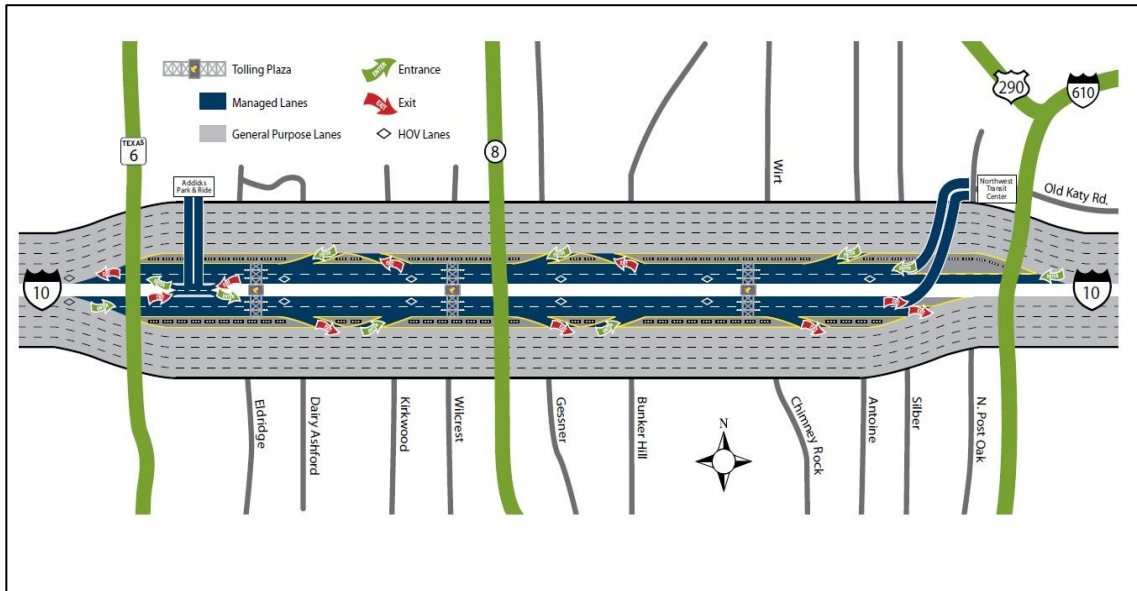
#### 4. DATA

A massive quantity of detection records from 2012 automatic vehicle identification (AVI) sensors on Katy Freeway (I-10) were used in this research. These records include transponder (toll tag) IDs, times of detection, and AVI sensor information from each vehicle that installed a transponder. However, these records are not meaningful themselves because they are not trip information but only detection information from one specific AVI sensor. Thus, this research first chains consecutive detections for the same vehicle to identify trips on the freeway. Then, detailed attributes of the trips, such as travel time and the amount of toll paid, are estimated. In order to estimate the option value of MLs, this research requires revealed preference (RP) data between MLs and GPLs of travelers on the freeway. However, the identified trips from chaining consecutive detections only provided actual choices of the travelers and their attributes. Thus, this research generated alternative choices that the travelers could have chosen at the time of travel but did not choose and their attributes estimated using other actual trips' attributes. By combining the actual and alternative choices and their attributes, each traveler's revealed preference was finally identified in our data set. This section explains these procedures. However, since there were missing records in December, 2012, our data set only includes detection records from January to November, 2012. Another variance is that in this section, an ML trip implies a paid trip and a GPL trip implies a toll-free trip.

#### **4.1 Katy Freeway (I-10)**

Katy Freeway (I-10) is one of the major highways in Houston, Texas and connects the city of Katy in the west to downtown Houston in the east. The freeway has a total length of 40 miles and was built in the 1960s. The initial design had three lanes and two frontage lanes per direction to accommodate about 80,000 vehicles per day. In the late 1990s and 2000s, traffic volume often reached three times the volume the freeway was built for, which caused chronic congestion lasting up to 11 hours a day. This led to the TxDOT undertaking a major reconstruction project for a 12-mile section of the freeway between west of State Highway 6 (SH 6) and the I-10/I-610 interchange. The project was started in 2003 and ended in October, 2008. The approximately 2.79 billion dollar project was partially funded by a combination of federal and state funds and toll revenues. The project widened the 12-mile section to accommodate up to six GPLs and two MLs per direction where toll rates varied depending on the time of day and occupants of a vehicle in each direction in the middle of the freeway. Figure 4 represents a detailed map of Katy Freeway. The two lanes in the middle of the freeway are MLs with four entries and four exits. High-occupancy vehicles (HOVs) with two or more occupants do not have to pay a toll during HOV-free hours but have to pay the same toll rate as single-occupancy vehicles (SOVs) during all other hours. HOV-free hours are Monday to Friday between 5 AM and 11 AM, and between 2 PM and 8 PM. To avoid the toll, HOVs must also drive in the HOV lane, the inside lane of the MLs in each direction to avoid being charged the toll during the HOV-free hours. GPLs are free

at all times. In general, the MLs provide faster and more reliable travel compared to the GPLs.



**Figure 4** Katy Freeway (Source: Harris County Toll Road Authority)

The freeway has three toll plazas in each direction near the cross streets of Eldridge, Wilcrest, and Wirt. All tolls are electronically collected at these plazas, and vehicles need to have a transponder (toll tag) in order to pay the toll. All vehicles passing through any toll plaza are identified with the transponder. Except for the HOVs during the HOV-free hours, all vehicles identified are charged a toll that varies depending on time of day and the toll plaza passed. The Harris County Toll Road Authority (HCTRA) is responsible for operation and toll collection for the freeway. The toll rates in 2012 are



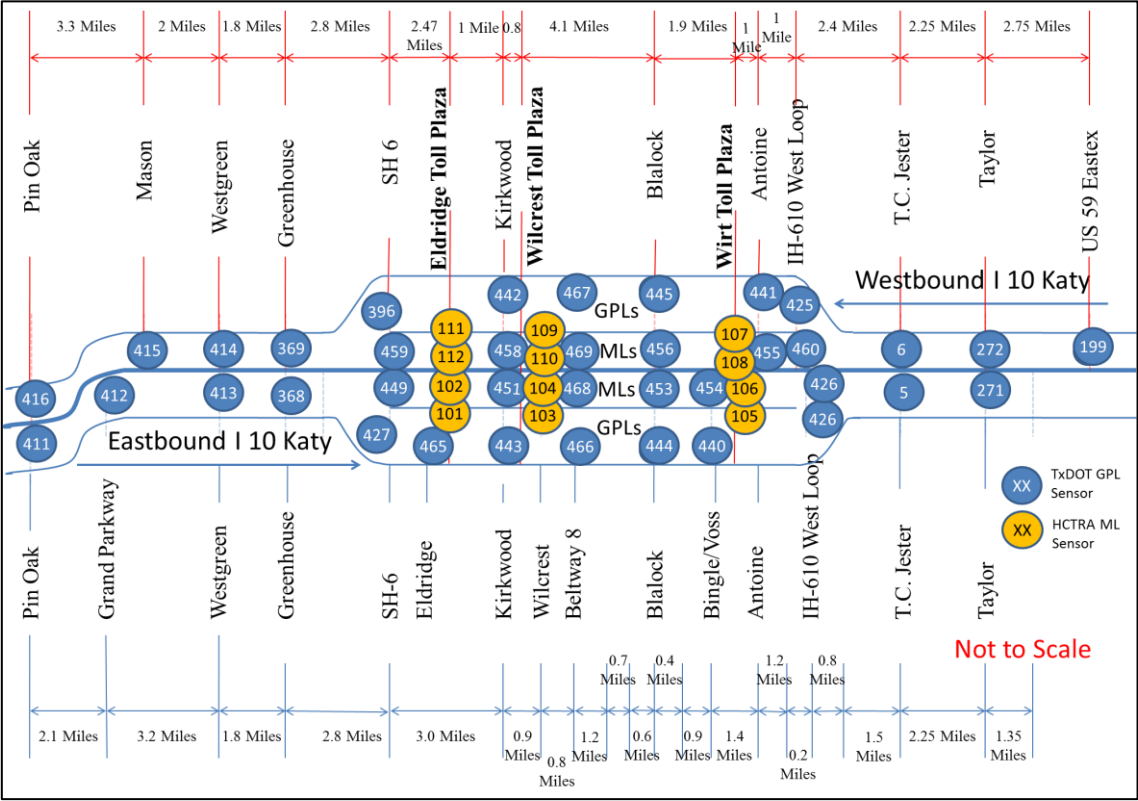
provided in Table 9. Note that the toll schedules and rates changed on September 8<sup>th</sup>, 2012.

**Table 9** Toll Schedules and Rates on Katy Freeway Managed Lanes (Source: Harris County Toll Road Authority)

Date	Time Period	Toll Plaza		
		At Eldridge	At Wilcrest	At Wirt
Before September 8, 2012	Peak Hours: Weekdays 7–9 AM Eastbound 5–7 PM Westbound	\$1.60	\$1.20	\$1.20
	Shoulder Hours: Weekdays 6–7 AM and 9–10 AM Eastbound 4–5 PM and 7–8 PM Westbound	\$0.80	\$0.60	\$0.60
	Off-peak Hours: All other times	\$0.40	\$0.30	\$0.30
As of September 8, 2012	Peak Hours: Weekdays 7–9 AM Eastbound 4–6 PM Westbound	\$2.20	\$1.40	\$1.40
	Shoulder Hours: Weekdays 6–7 AM and 9–10 AM Eastbound 3–4 PM and 6–7 PM Westbound	\$1.10	\$0.70	\$0.70
	Off-peak Hours: All other times	\$0.40	\$0.30	\$0.30

AVI sensors are installed along the MLs and GPLs in both directions on the freeway. The section of the freeway examined in this research has 38 AVI sensors operated by TxDOT (see Figure 5). Each sensor has a unique number, which is used to identify location and direction of travel for vehicles passing the sensor. Only vehicles that have a valid transponder ID are detected at these sensors. When each sensor detects a vehicle, the sensor records the time of detection, and the unique transponder ID of the

vehicle. For this research, these records and the locations of the AVI sensors were used to identify vehicles' trips on the freeway. Each transponder ID was assigned a unique random ID, and the original transponder ID was removed from our data set. Thus, it is impossible to trace the records back to the traveler who made a specific trip. Figure 5 provides the locations of the AVI sensors used in this research.



**Figure 5** Katy Freeway AVI Sensor Locations

## 4.2 Data Sources

### 4.2.1 Automatic Vehicle Identification (AVI) Data

The AVI data contains detection records of vehicles that installed a transponder at all AVI sensor locations on the freeway in 2012. The data was obtained from TxDOT. Each record has a time stamp (time of detection), each vehicle's unique transponder ID, and the AVI sensor number that detected the vehicle. The data were processed and used to identify trips on the MLs and GPLs. The data were combined with the HCTRA toll data, and then each transponder ID was assigned a unique random ID. In this way, it is impossible to identify the original transponder ID and the traveler who made a particular trip. For 2012, 225,118,768 records were obtained from the 38 AVI sensors. These contain 1,993,347 unique transponder IDs for the entire year.

#### *4.2.2 Harris County Toll Road Authority (HCTRA) Toll Data*

The toll data on the freeway in 2012 was obtained from the HCTRA, which used the data to charge an appropriate toll to vehicles that used the MLs based on the toll rates in Table 9. The data contains all vehicles with a valid transponder ID that were detected at the toll plaza(s) along the MLs on the freeway. Each record in the data set contains a time stamp, the vehicle's unique transponder ID, location, toll plaza ID and lane ID. This data set was used to supplement the AVI detection records in order to better identify trips along the MLs. Based on the location, the toll plaza ID, and the lane ID, a unique sensor number was assigned to each toll sensor. All 12 toll sensors were assigned sensor numbers from 101 to 112. Figure 6 provides locations of these toll sensors. This data set was used to assign a correct toll to each trip identified as an ML trip or properly identify non-tolled vehicles in the case of toll-free HOVs on the MLs. For 2012, 14,769,730 toll detection records were obtained from the 12 toll sensors.

### **4.3 Revealed Preference (RP) Data on Katy Freeway**

To estimate the option value of MLs, this research required revealed preference (RP) data that contains key attributes of the alternatives considered by travelers to choose between the MLs versus the GPLs on Katy Freeway. Using the data sets (AVI and HCTRA toll data sets), a new data set which includes all identified actual choices (trips) in 2012 and their possible attributes was generated. Some of the trip attributes include a random ID for each vehicle, travel time, standard deviation of travel time at 10-minute intervals during the time of travel, the amount of toll paid, time of day when the trip was made (peak, shoulder, and off-peak hours), and trip length. After identifying all actual choices (trips) in 2012, the alternative choices (trips) that travelers could choose but did not choose at the time of actual choices (i.e., GPL trip for ML trips and ML trips for GPL trips) and their attributes were also generated in the same data set in order to construct the RP data set. The attributes for each alternative choice (trip) were estimated using attributes of other actual trips that were made on the same section as the actual trip at the same time (at the same 10-minute interval). However, there were times when too few alternative trips were made to determine attributes of alternative trips based on real trips. Thus, to estimate some unknown attributes, reasonable assumptions were made. After identifying the actual and alternative choices and their attributes, this research removed and corrected some erroneous observations. Some of these erroneous observations include reverse direction trips on the freeway, alternative trips that do not correctly count the amount of toll paid, and unreasonably fast trips, such as 906 mph. In addition, some observations that are not fit for the objectives of this research were

excluded. For example, trips that were only on sections where MLs were unavailable were excluded because travelers who traveled those sections did not value the MLs as a backup option. Last, to obtain the RP data, this research used SAS (Statistical Analysis System) software and made an SAS program code included in Appendix B.

#### *4.3.1 Processing Raw Data (AVI and HCTRA Data)*

The first step of identifying RP data on Katy Freeway was to clean the raw data sets, AVI and HCTRA toll data sets. No detection record with incomplete information was found in the AVI and HCTRA toll data, but a few duplicated detections were found and removed. In the HCTRA toll data, toll sensors were given sensor numbers to enable them to merge with the AVI data. In the initial HCTRA data set, each toll plaza was identified by a toll plaza ID. In order to merge the AVI and the HCTRA data sets, a unique sensor number was assigned to each toll plaza instead of using the plaza ID. All attributes except the time stamp, sensor number, and transponder ID were excluded in the HCTRA data set and then two data sets were combined. After the merge, a unique random ID was assigned to each transponder ID, and the original transponder IDs were deleted. This procedure made it impossible to trace the original ID from the random IDs. This merged and randomized data set was used for all subsequent analysis. Detection records of random IDs that were only detected at one sensor (a single location) were removed because no trip could be identified by a single detection. Thus, a data set with detection records of random IDs that were detected more than once was generated. Last, since there were missing records in December, 2012, our data set excluded all detection records in December, 2012. After these initial procedures, the total number of the

detection records (individual transponder reads) in 2012 (from January to November) was 176,026,903 in the data set.

#### *4.3.2 Identifying Actual Choices (Trips) on Katy Freeway*

The detection records were sorted by chronological order and random ID. Thus, consecutive detections by the same random ID were chained to trace a trip along the freeway. For example, a specific random ID was identified at sensor 427, then 465, then 443, and finally, 466 within the same time period and was converted to a single trip which started at sensor 427 and ended at sensor 466 (see Figure 5 for the locations of the sensors). If the time difference between two consecutive detections for the same random ID was more than 15 minutes, the two detections were separated as a part of two different trips. Using all detections during the peak and shoulder hours, this research tested how many actual trips could increase from the separation using the 15-minute time limit. Only a few trips (about 0.6% of total trips during the peak and shoulder hours) increased due to the separation. Thus, the separation had a minimal impact on the identification of all actual trips on the freeway in 2012. In this way, all actual choices (trips) on the freeway in 2012 were identified.

After identifying the actual trips on the freeway, four attributes (travel time, trip length, the amount of toll paid, and standard deviation of travel times at the 10-minute interval during the time of travel) were estimated for each actual trip. First, travel time was calculated by recording the difference in time between the first detection and the last detection. Similarly, trip length was calculated by measuring the distance between the first and last sensor of each actual trip. Distances between all sensors were measured

using Google Earth and are provided in Figure 5. The amount of toll paid at the time of detection was estimated for trips that were detected at least once at one of the toll sensors (plazas) based on the toll schedules and rates in Table 9. Trips that were detected on the HOV lane of the MLs during the HOV-free hours (Monday to Friday between 5 AM and 11 AM, and between 2 PM and 8 PM) were considered as toll-free trips. Total amount of toll paid for one complete trip was equal to the sum of tolls paid along the complete trip at up to three different toll plazas. Standard deviation of travel time was calculated as a measure of travel time reliability that might be useful to explain lane-choice behaviors between the MLs and the GPLs. The standard deviation was calculated per 10-minute interval for each sensor pair (between trip start and end sensors) using travel times of all trips that passed that sensor pair at the same 10-minute interval (see Equation 3).

$$\sigma_{X,Y} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2} \quad (3)$$

where:

$\sigma_{X,Y}$  = standard deviation of travel time between sensors X and Y on the freeway,

$x_i$  = travel time of trip  $i$  that was made on the section between sensors X and Y at the 10-minute interval,

$\bar{x}$  = average travel time between sensors X and Y at the 10-minute interval,

$N$  = total number of observations at the 10-minute interval.

Despite the large number of trips in our data set, some 10-minute intervals had too few trips to estimate the standard deviation. If there were less than three trips for one sensor pair but adjacent segments were located between the sensor pair with enough trips, the standard deviation for the sensor pair was estimated using the regression equation shown in Equation 4:

$$\sigma_{X,Y} = 0.48 \times \sum \sigma_i + 2.20 \times S + 6.37 \quad (4)$$

where:

$\sigma_{X,Y}$  = standard deviation of travel time between sensors X and Y on the freeway,

$\sigma_i$  = standard deviation of travel time for all adjacent segments located between sensors X and Y,

$S$  = the number of adjacent segments located between sensors X and Y.

For example, if there were less than three trips that start at sensor 413 and ended at sensor 444 while passing through sensors 368 and 443 during the same 10-minute interval, the standard deviation for the trips is given by Equation 5 (see Figure 5 for the locations of the sensors):

$$\sigma_{413,444} = 0.48 \times (\sigma_{413,368} + \sigma_{368,443} + \sigma_{443,444}) + 2.20 \times 3 + 6.37 \quad (5)$$

Note that if there were less than two trips on one of the adjacent segments, the standard deviation for the segment was allowed to be zero. However, these were very



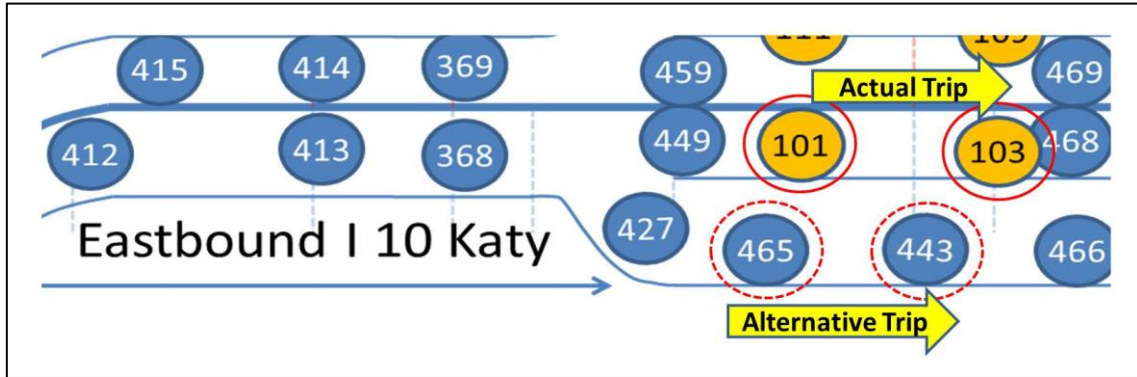
few in the data set. Equation 4 was adopted from Burris and Danda (2014), which also used the AVI and HCTRA data sets. In their study, this regression equation was developed using data from all sensor pairs where at least five trips were identified between a given sensor pair during the same 10-minute interval in April, 2012. The regression model has an R-squared value of 0.49. They tested various models to improve the R-squared value, but no significant improvement was observed.

#### *4.3.3 Identifying Alternative Choices (Trips) on Katy Freeway*

In the previous section, this research identified actual choices (trips) of travelers who installed a transponder on the freeway in 2012 using the AVI and HCTRA data sets. However, in order to construct RP data on the freeway, this research requires an alternative choice (trip) for each actual choice (trip). The alternative trip is a trip that was available during a traveler's trip time but was not chosen. This means, for every trip made on the MLs, an alternative trip on the GPLs was created and vice versa.

Generating an alternative trip was generally straightforward. Since, in our data set, each actual trip included all sensor information that was passed by a traveler (transponder) who made the trip, we first generated alternative sensors against all sensors of each actual trip. Then, by chaining all alternative sensors of the actual trip, this research generated an alternative trip and estimated its attributes, such as travel time, and standard deviation of travel time obtained from all trips that were made on the alternative lanes during the same 10-minute interval as the actual trip. This was done for all actual trips in the data set. Finally, by combining each actual trip and its alternative

trip, this research constructed RP data. Figure 6 is provided to better explain this procedure.

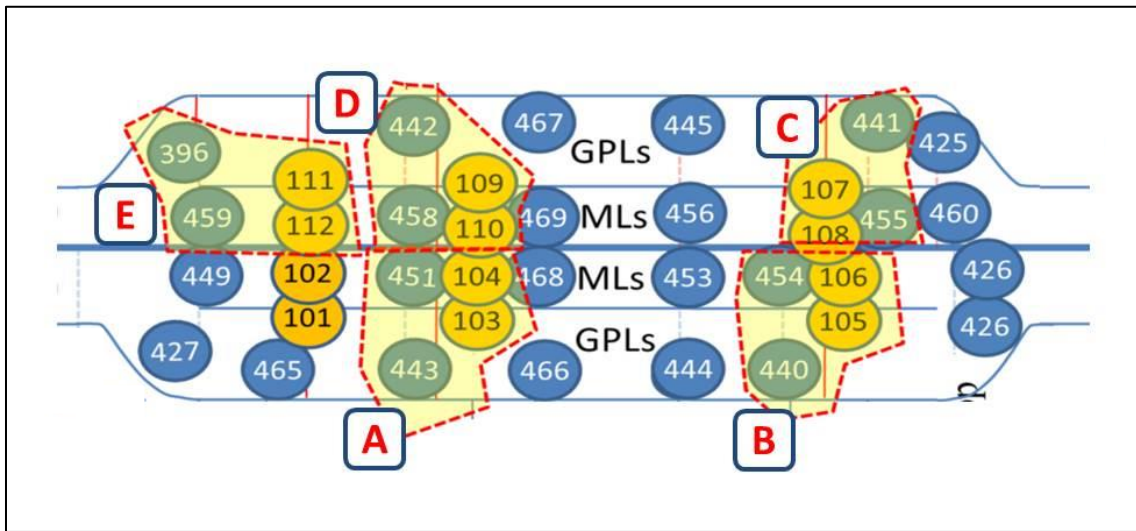


**Figure 6** Example of Actual Trip and its Alternative Trip

Suppose that one transponder was sequentially detected at sensors 101 and 103. From these detections, its actual trip is identified as an ML trip between sensors 101 and 103. Since its alternative trip is a GPL trip, this research first generates two alternative sensors 465 and 443, and then, by combining them, an alternative trip between sensors 465 and 443 is generated. Thus, an alternative was generated such that it passed through the same section of the freeway but on the other set of lanes (i.e., MLs for GPLs and vice versa). Last, alternative trip's attributes were estimated from all trips that were made from sensor 465 to sensor 443 during the same 10-minute interval as the actual trip.

In the alternative trip generation procedure, to generate reasonable alternative trips, each sensor has to have one unique alternative sensor on sections where both MLs and GPLs are available. However, on each of the five sections on the freeway (Sections

A, B, C, D and E in Figure 7), there is only one GPL sensor while there are three sensors on the MLs, one ML sensor operated by TxDOT, and HCTRA toll and HOV lane sensors. This could generate unreasonable alternative trips and consequently our data set could contain unreasonable revealed preferences on the freeway.



**Figure 7** Five Sections (where unreasonable alternative trips could be generated)

In Figure 7, in Section E, there is only one GPL sensor (Sensor 396) whereas there is an ML sensor (sensor 459), and toll and HOV lane sensors (sensors 111 and 112). Thus, for all three sensors on the MLs, an alternative GPL sensor is the same as sensor 396. If one transponder was detected at sensors 111 and 459 on the MLs, its alternative GPL trip was generated as a trip between sensors 396 and 396 in the procedure. Obviously, this alternative GPL trip is not a reasonable trip because it includes only the one sensor. This unreasonable alternative trip could also be generated on the other

sections A, B, C, and D and was frequently observed in our data set after the initial alternative trip generation.

In order to correct unreasonable alternative trips on the five sections, there is no choice but to remove all detection records at sensors 451, 454, 455, 458 and 459 in our data set. This is because there is only one GPL sensor on each of the five sections and the HCTRA toll and HOV lane sensors cannot be removed. The HCTRA toll and HOV lane sensors were used to estimate the amount of toll that had to be paid by a traveler when s/he passed through the MLs. In addition, no transponder can be detected at both toll and HOV lane sensors in the same trip because those are located at exactly the same location on the MLs although, in order to consider toll-free HOV lane trips during the HOV free hours, this research separated two sensors. Thus, this research does not need to remove one of the HCTRA toll sensors and one of the HOV lane sensors. Note that removal of all detection records at the ML sensors 451, 454, 455, 458 and 459 does not correct all trips that contain any of five ML sensors but removes only trips that contain an unreasonable alternative trip. This is explained by the examples in Table 10.

In Table 10, traveler A traveled from the sensor 101 to the sensor 103. S/he spent four minutes (from 10:01 AM to 10:05 AM) for the trip and paid a toll twice at the sensors 101 and 103. Thus, the detection at sensor 451 does not give any meaningful information for the actual trip. In addition, for its alternative trip, it only generated the same alternative sensor 443 with the sensor 103 (see the fourth column in Table 10). Thus, without the detection at the sensor 451, essential actual and appropriate alternative trip information can be identified (see the fifth column in Table 5). However, for the

traveler B, sensor 451 generated an unreasonable alternative trip between the same sensor 443 and consequently our data set included the unreasonable revealed preference. Thus, this research had to remove all detection records at the ML sensors 451, 454, 455, 458 and 459. After the removal of the detection records, the total number of the detection records (individual transponder reads) in 2012 was 169,383,445 in the data set.

**Table 10** Examples of Alternative Trip Generation in the Data Set

Random ID	Time of Detection	Sensor (Actual Trip)	Alternative Sensor 1 (Initial Generation)	Alternative Sensor 2 (After the Removal)
A	03JAN2012 /10:01:00	101	465	465
A	03JAN2012 /10:03:00	451	443	-
A	03JAN2012 /10:05:00	103	443	443
B	03JAN2012 /10:01:00	451	443	-
B	03JAN2012 /10:03:00	103	443	443

After the removal, this research regenerated alternative trips for all actual trips. Then, four attributes of the alternative trips (travel time, trip length, the amount of toll paid, and standard deviation of travel times at the 10-minute interval during the time of travel) were estimated. The trip length, the amount of toll paid, and the standard deviation were estimated in the same way as the actual trip identification. However, the travel time of each alternative trip was estimated by taking the average of travel times on the alternative lane on the same section of the freeway during the same 10-minute

interval when its actual trip was made on the same day. Note that the start time of each alternative trip was assumed to be exactly the same start time of its actual trip. Despite a large number of trips in the data set, some intervals do not have any trip on the alternative lane during the 10-minute intervals. In such cases, average speed of each lane depending on time of day (peak, shoulder, and off-peak hours) was used to estimate travel time of alternative trips. The average speed was adopted from Burris and Danda (2014), which also used the AVI and HCTRA data sets provided in Table 11. The average speed was calculated using all actual trips on each lane during the same time of day (peak, shoulder, and off-peak hours) with their travel time information.

**Table 11** Average Speed on Each Lane of the Freeway by Time of Day (Source: Burris and Danda 2014)

Time of Day	Average Speed on the MLs	Average Speed on the GPLs
Peak Hours	53.2 mph	42.8 mph
Shoulder Hours	61.3 mph	55.6 mph
Off-peak Hours	68.1 mph	65.3 mph

Note: Peak, shoulder, and off-peak hours refer to the same time periods in Table 9.

After estimating the travel time of each alternative trip, the estimated travel time had to be adjusted because each alternative trip length could be different from its actual trip length (up to  $\pm 3.75$  miles) due to different locations of ML and GPL sensors on the same section of freeway. For example, suppose that one traveler traveled from toll sensor 103 to toll sensor 105. In this case, the actual trip length between toll sensors 103

and 105 is 6.0 miles while its alternative trip length between the sensors 443 and 440 on the GPLs is 5.5 miles (see Figure 5 for each sensor location and distance). In our data set, about 75 percent of the generated alternative trips had a different trip length from their actual trip.

This research adjusted the estimated travel time of each alternate trip corresponding to its actual trip length. This adjustment is provided in Equation 6. Note that the adjustment in Equation 6 assumed that average speed on the unmatched portion of each alternative trip was the same as the average speed of the alternative trip.

$$TT_{New} = TT_{Old} \times \frac{Actual\ Trip\ Length}{Alternative\ Trip\ Length} \quad (6)$$

where:

$TT_{New}$  = travel time of an alternative trip after the adjustment,

$TT_{Old}$  = travel time of an alternative trip before the adjustment.

#### *4.3.4 Cleaning and Correcting Actual and Alternative Choice (Trip) Data*

In the previous sections, this research identified actual and alternative choices and their attributes to construct RP data on the freeway. However, after reviewing the data in detail, this research observed some trips that contain critical errors. In addition, this research identified trips that do not fit for the purposes of this research. This section explains procedures to correct or to clean those trips.

### **Cleaning Trips that Include Repeated Detections at One Sensor**

A few actual trips that include repeated detections at one sensor were identified. Figure 8 provides real examples of these trips in our data set. In the first example, traveler 892628909 (see 'RandID' column in Figure 8) was detected seven times at the sensor 425 for 339 seconds ( $=37+34+30+67+69+42+60$ ) (see 'Allsensor' and 'timediff' columns in Figure 8 for the number of detections and the dwell time at sensor 425). In the second example, traveler 891111687 was detected twice at sensor 5 for 864 seconds. These trips may be considered as trips caused from extreme congestion. However, it is more reasonable that the trips are considered as trips caused from system errors rather than congestion because range of transponder recognition by the AVI and the HCTRA sensors is 31.5 feet ("EZ TAG" 2014). It seems unlikely that vehicles cannot move 31.5 feet forward for more than five minutes on the freeway even in extreme congestion. In addition, in the first example, the average speed on the section between the sensors 425 and 441 (the nearest sensor from the sensor 425) was 60.4 mph and, in the second example, the average speed on the section between the sensors 5 and 271 (the nearest sensor from the sensor 5) was 64.7 mph. These average speeds were calculated using travel speeds of other vehicles that traveled those sections at the same 10-minute interval as the travelers 892628909 and 891111687, respectively. Thus, this research eliminated all trips that include repeated detections at one sensor. After this elimination, the total number of the detection records (individual transponder reads) in 2012 was 169,337,527.



VIEWTABLE: Katy2.ZeroLength

	RandID	datetime	timediff	Allsensor	seglen
217	892797048	20APR2012:15:15:46	24	469,469	0
218	892797048	20APR2012:15:17:20	94	469,469,469	0
219	892797048	20APR2012:15:17:45	25	469,469,469,469	0
220	892797048	20APR2012:15:18:21	36	469,469,469,469,469	0
221	892695678	29JUN2012:17:41:13	44	469,469	0
222	892695678	29JUN2012:17:42:32	79	469,469,469	0
223	892628909	30JUN2012:13:02:12	37	425,425	0
224	892628909	30JUN2012:13:02:46	34	425,425,425	0
225	892628909	30JUN2012:13:03:16	30	425,425,425,425	0
226	892628909	30JUN2012:13:04:23	67	425,425,425,425,425	0
227	892628909	30JUN2012:13:05:32	69	425,425,425,425,425,425	0
228	892628909	30JUN2012:13:06:14	42	425,425,425,425,425,425,425	0
229	892628909	30JUN2012:13:07:14	60	425,425,425,425,425,425,425,425	0
230	892292940	16NOV2012:15:04:00	797	425,425	0
231	891486796	01OCT2012:16:16:03	797	427,427	0
232	891249720	30NOV2012:11:10:50	596	5,5	0
233	891111687	06SEP2012:12:27:03	864	5,5	0

**Figure 8** Examples of Actual Trips (that include repeated detections)

### Cleaning Reverse Direction Trips

A few actual trips that were made in a reverse direction on the freeway were identified (i.e., westbound trips on the eastbound freeway and vice versa). Figure 9 provides real examples of these trips in the data set. In the first example, traveler 999999052 traveled from sensor 453 to the sensor 411 (see ‘Allsensor’ column in Figure 9) in a westward direction on the eastbound freeway (see Figure 5 for each sensor location). The negative trip length in the ‘seglen’ column also indicates a reverse direction trip. In the second example, traveler 999697882 also had a negative trip length. Obviously, these trips could be caused from system errors. Thus, this research excluded

all reverse direction trips from the data set. After this elimination, the total number of the detection records (individual transponder reads) in 2012 was 169,306,545 in the data set.

	RandID	datetime	timediff	Allsensor	seglen
1	999999052	11JUL2012:04:28:48	1179648.453		0
2	999999052	11JUL2012:04:29:38	50 453,411		-17.93
3	999903179	06FEB2012:18:07:47	2130 444		0
4	999903179	06FEB2012:18:11:55	248 444,412		-15.83
5	999697882	27AUG2012:07:48:36	1094606 5		0
6	999697882	27AUG2012:07:58:45	608 5,271		2.2
7	999697882	27AUG2012:08:11:28	763 5,271,411		-26.61
8	999697882	27AUG2012:08:13:19	111 5,271,411,412		2.1
9	999697882	27AUG2012:08:25:19	720 5,271,411,412,465		10.3
10	999697882	27AUG2012:08:27:07	108 5,271,411,412,465,443		1.21

**Figure 9** Examples of Reverse Direction Trips

### **Cleaning Trips That Were Only on Sections Where MLs Were Unavailable**

On the freeway, sections where MLs are unavailable exist (outside the ML sections). For example, if one traveler traveled from the sensor 411 to sensor 368, s/he was unable to use the MLs (see Figure 5 for each sensor location). On those sections, travelers could not choose between the MLs and the GPLs and, therefore, revealed preferences for those sections did not exist. This research excluded those trips that were made only outside the ML sections. After this elimination, the total number of detection records (individual transponder reads) in 2012 was 155,543,880 in the data set.

### **Cleaning Trips That Changed Lanes**

Actual trips that switched from the GPLs to the MLs, or vice versa, on the middle of sections where both MLs and GPLs were available were identified. Knowing the

exact location of the switch was impossible to determine because vehicles were only detected at the AVI sensors. As a result, travel times for their alternative trips could not be estimated. Consequently, those trips that switched from the GLPs to the MLs, or vice versa, were removed from the data set. After this elimination, the total number of detection records (individual transponder reads) in 2012 was 131,799,004 in the data set. This represented 39,465,989 actual trips. This trip data still required cleaning and correcting procedures. The next sections explain how these procedures were handled for the RP data.

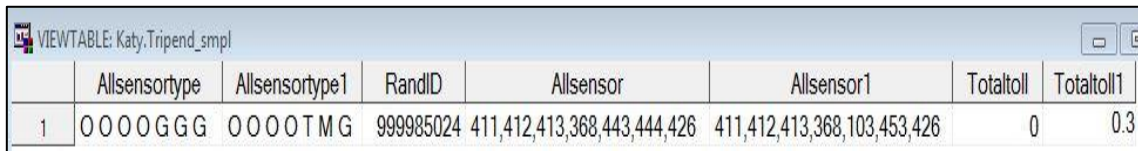
### **Cleaning Trips on the HOV Lanes During the HOV-free Hours**

All trips which were made on the HOV lanes during the HOV-free hours did not pay any toll, but received the same benefits as paid trips on the toll lanes. As mentioned in the previous section, this research required revealed preferences between a toll-free (GPL) trip versus a paid (ML) trip on the freeway. Thus, those trips were beyond the scope of this research and, therefore, were excluded from the data set. After this elimination, the total number of trips (choices) in 2012 was reduced to 35,131,739.

### **Correcting Omitted Toll Sensors and the Amount of Toll Paid for Alternative Trips**

The AVI sensors operated by TxDOT on the GPLs are not perfect and do miss some vehicles (transponders) as they were originally designed to gather enough data to provide accurate travel time information on the GPLs. As a result, not all GPL sensors passed by a vehicle that installed a transponder were recorded in the AVI data set for that vehicle. The vast majority of toll and HOV lane sensors passed by the vehicle are recorded in the HCTRA data set as they were designed for toll collection. Thus, it was

possible that an actual trip could pass by a GPL sensor and not be read, but instead lead to an alternative trip that should have passed an ML toll sensor but did not due the omitted GPL sensor. Figure 10 provides a real example of an actual GPL trip that omitted GPL sensors even though the GPL sensors were actually passed.



	Allsensortype	Allsensortype1	RandID	Allsensor	Allsensor1	Totaltoll	Totaltoll1
1	OOOOGGG	OOOOTMG	999985024	411,412,413,368,443,444,426	411,412,413,368,103,453,426	0	0.3

**Figure 10** Example of the Actual Trip that Omitted GPL Sensors

In Figure 10, ‘RandID’ column indicates the vehicle’s unique random ID. The ‘Allsensor’ column indicates all sensors where the vehicle was detected, and ‘Allsensor1’ column indicates all alternative trip sensors that were generated from the sensors in the ‘Allsensor’ column. The ‘Allsensortype’ column indicates sensor types of the sensors in the ‘Allsensor’ column and the ‘Allsensortype1’ column indicates sensor types of the sensors in the ‘Allsensor1’ column. In these columns, ‘O’ indicates a sensor on the outside of ML sections, ‘M’ indicates an ML sensor, ‘G’ indicates a GPL sensor, and ‘T’ indicates a toll lane sensor. The ‘Totaltoll’ column indicates dollar amounts of toll paid for the actual trip, and the ‘Totaltoll1’ column indicates the dollar amounts of tolls that are estimated from the sensors in the ‘Allsensor1’ column for the alternative trip.

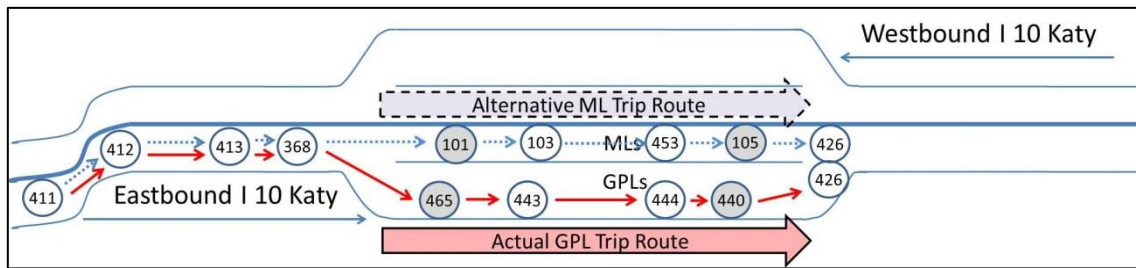
Since it was difficult to identify which GPL sensors were omitted in the actual trip in Figure 10, Table 12 and Figure 11 are provided together. Table 12 provides all

sensors and their types that were passed in the actual and alternative trips in Figure 10. In the actual GPL trip, the GPL sensors 465 and 440 were actually passed but were not recorded (see Table 12). This can be proven in Figure 11. Even though the vehicle was not detected at sensors 465 and 440, it likely passed them based on the actual GPL trip route in Figure 11. Since the GPL sensors 465 and 440 were not in the actual trip information, this research could not generate toll sensors 101 and 105 for the alternative trip in Table 12 from the previous alternative trip identification.

**Table 12** Detailed Trip Information of the Example in Figure 10

Trip Type	Detailed Trip Information									
Actual GPL Trip	All sensors	411	412	413	368	465	443	444	440	426
	All sensor types	O	O	O	O	G	G	G	G	G
Alternative ML Trip	All sensors	411	412	413	368	101	103	453	105	426
	All sensor types	O	O	O	O	T	T	M	T	G

Note: In the actual GPL trip, shaded areas indicate sensors that were actually passed but were not recorded. In the alternative ML trip, shaded areas indicate sensors that could not be generated because sensors 465 and 440 were not recorded in the actual trip.



**Figure 11** Trip Routes of both Actual and Alternative Trips detailed in Table 12

Calculating the amount of toll paid for alternative trips is based on toll sensors included in alternative trip information. Thus, since the alternative trip identification in the previous section did not generate toll sensors 101 and 105 but only toll sensor 103 for the alternative trip, the amount of toll paid was calculated as 0.3 dollars from sensor 103 (see 'Allsensor1' and 'Totaltoll1' columns in Figure 10). However, based on the alternative ML trip route in Figure 11, if the vehicle chose the alternative ML trip, it had to pay 0.7 dollars ( $=0.4+0.3$ ) more at toll sensors 101 and 105. Based on this example, it was possible that other alternative trips on the MLs may have also missed toll sensors and consequently the amount of toll paid was not correctly estimated in the previous section. Thus, this research needed to identify toll sensors which were not generated but were needed for each alternative trip and then to correct the amount of toll paid for it.

This research first established criteria that can identify toll sensors that were omitted in each alternative trip (see Table 13).

**Table 13** Criteria that can Identify Omitted Toll Sensors in Each Alternative Trip

Omitted Toll Sensor	Criteria
Sensor 101	a) An actual trip includes one of the sensors 411, 412, 413, 368, 427 and 449, and, b) An actual trip includes one of the sensors 443, 466, 444, 440, 426, 5, 271, 468, 453, 103, 104, 105 and 106, and, c) An actual trip does not include the sensors 101, 102 and 465.
Sensor 103	a) An actual trip includes one of the sensors 411, 412, 413, 368, 427, 465, 449, 101 and 102, and, b) An actual trip includes one of the sensors 466, 444, 440, 426, 5, 271, 468, 453, 105 and 106, and, c) An actual trip does not include the sensors 103, 104 and 443.
Sensor 105	a) An actual trip includes one of the sensors 411, 412, 413, 368, 427, 465, 443, 466, 444, 449, 468, 453, 101, 102, 103 and 104, and, b) An actual trip includes one of the sensors 426, 5 and 271, and, c) An actual trip does not include the sensors 105, 106 and 440.
Sensor 107	a) An actual trip includes one of the sensors 199, 272, 6, 425 and 460, and, b) An actual trip includes one of the sensors 445, 467, 442, 396, 369, 414, 415, 416, 456, 469, 109, 110, 111 and 112, and, c) An actual trip does not include the sensors 107, 108 and 441.
Sensor 109	a) An actual trip includes one of the sensors 199, 272, 6, 425, 441, 445, 467, 460, 456, 469, 107 and 108, and, b) An actual trip includes one of the sensors 396, 369, 414, 415, 416, 111 and 112, and, c) An actual trip does not include the sensors 109, 110 and 442.
Sensor 111	a) An actual trip includes one of the sensors 199, 272, 6, 425, 441, 445, 467, 442, 460, 456, 469, 107, 108, 109 and 110, and, b) An actual trip includes one of the sensors 369, 414, 415 and 416, and, c) An actual trip does not include the sensors 111, 112 and 396.

Thus, using the criteria in Table 13, this research identified the toll sensors that were omitted in each alternative trip. About 15 percent of alternative trips identified in the previous section missed toll sensor(s).

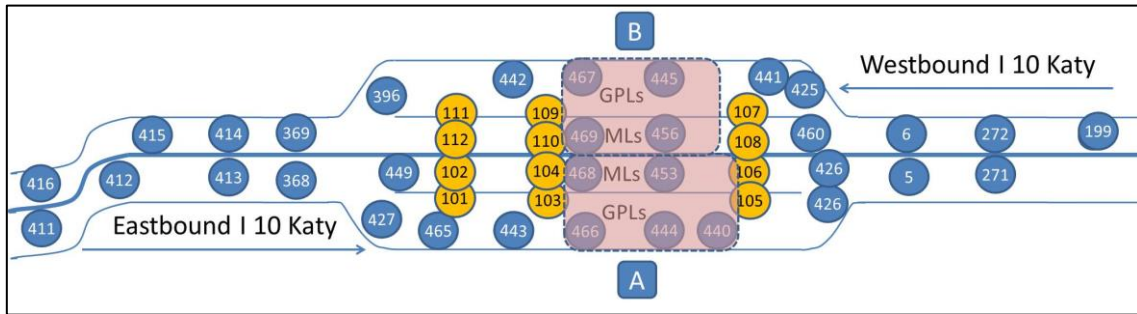
After identifying the toll sensors omitted in each alternative trip, this research also corrected the amount of toll paid for it based on the toll sensors omitted and time of day when the trip occurred. After this correction, the total number of trips (choices) in 2012 was reduced to 35,059,345.

## **Cleaning Trips That Are Unlikely to Occur in the Real World**

This research identified two types of trips that are unlikely to occur in the real world. One type represents trips where the average speed is unreasonably fast, such as 906 mph. These trips could be caused from system misreads. Thus, trips of which average speed is greater than 110 mph were removed. 2,828 trips were removed due to high speed. The other type represents trips where the alternative trip was only on ML sections but did not pass any toll plaza (sensors) even after omitting toll sensor identification in the previous section. That alternative trip cannot occur in real travel because once travelers enter into ML sections, they cannot leave the ML sections without passing a toll plaza. However, those alternative trips could be generated in the previous alternative trip identification. For example, for an actual trip from sensor 466 to sensor 444 on the GPLs, an alternative trip was generated as a trip from sensor 468 to sensor 453 on the MLs in this research (see Figure 12 for each sensor location). However, in real travel, travelers cannot leave this ML section between sensors 468 and 453 without passing the toll plaza (sensor) 105. Therefore they are unlikely to consider this ML travel option for the trips between sensors 466 and 444 because, after using the ML travel option, they have to make a detour to sensor 444. On the freeway, two GPL sections (one per each direction) can generate unreasonable alternative trips as in the example if their actual trips were made only on the GPL sections. These sections are provided in Figure 12 (refer to the Sections A and B). Since reasonable ML alternative trips cannot be generated on Sections A and B, actual trips that were made only on the



sections should not be included in our data set. After excluding these two types of trips, the total number of trips (choices) in 2012 was reduced to 33,245,549.



**Figure 12** Trip Sections where Travelers would not Consider a Choice between ML and GPL Trips

#### 4.3.5 Final Data set

Additional attributes that might be useful to explain revealed preferences between paid ML and toll-free GPL trips were determined for each trip in the data set. These attributes were time of day (peak, shoulder, and off-peak hours), day of week, the total number of trips in 2012 on the freeway by the same vehicle (transponder), the total number of ML trips in 2012 on the freeway by the same vehicle, and direction of each trip.

The final data set which had two records for each trip (choice) contained two records for each trip which represent the two potential options for the trip. One is the option that was actually chosen, and the other is the option that was not chosen. Trip attributes included in the final data set were the unique random ID, lane choice, trip type

(either paid ML trip or toll-free GPL trip), travel time, standard deviation of travel time, the amount of toll paid, trip length, time of day, day of week, the total number of trips in 2012, the total number of ML trips in 2012, and direction.

#### **4.4 Statistics of Revealed Preference (RP) Data on Katy Freeway**

This section provides statistics of actual trips (choices) from the final data set. This includes the number of ML and GPL trips in each month, the number of ML and GPL trips by time of day, the number of ML and GPL trips by day of week, and the number of ML and GPL trips in each direction. Remember that since there were missing records in December, 2012, the data set did not include trips in December, 2012. Thus, the statistics indicate the numbers of ML and GPL trips from January to November, 2012

A total of 33,245,549 trips were identified in the data set. ML trips account for 6.1 percent (2,011,283 trips) and GPL trips account for 93.9 percent (31,234,266 trips) in the data set. In May and August, 2012, a higher percentage of ML trips were identified than the other months (see Table 14).

**Table 14** The Number of ML and GPL Trips in 2012

Month	ML Trips	GPL Trips	Total Trips	Percentage of ML Trips
January	152,253	3,058,315	3,210,568	4.7%
February	167,367	2,968,234	3,135,601	5.3%
March	179,237	3,084,220	3,263,457	5.5%
April	178,824	2,897,333	3,076,157	5.8%
May	206,003	2,646,650	2,852,653	7.2%
June	181,862	2,727,491	2,909,353	6.3%
July	171,491	2,650,671	2,822,162	6.1%
August	216,255	2,526,105	2,742,360	7.9%
September	183,336	2,664,531	2,847,867	6.4%
October	201,212	3,075,595	3,276,807	6.1%
November	173,443	2,935,121	3,108,564	5.6%
Total in 2012	2,011,283	31,234,266	33,245,549	6.1%

A total of 4,043,126 trips during the peak hours, 3,825,338 trips during the shoulder hours, and 25,377,085 trips during the off-peak hours were identified in the data set. During the peak hours, the percentage of ML trips was much higher than the other times of day (see Table 15). Note that peak, shoulder, and off-peak hours refer to the same time periods in Table 9.

**Table 15** The Number of ML and GPL Trips by Time of Day in 2012

Time of Day	ML Trips	Percentage of ML Trips	GPL Trips	Percentage of GPL Trips	Total Trips
Peak Hours	764,745	18.9%	3,278,381	81.1%	4,043,126
Shoulder Hours	439,642	11.5%	3,385,696	88.5%	3,825,338
Off-peak Hours	806,896	3.2%	24,570,189	96.8%	25,377,085
Total	2,011,283	6.1%	31,234,266	93.9%	33,245,549

A total of 25,839,063 trips during weekdays in 2012, and 7,406,486 trips during weekends were identified. There was a higher percentage of ML trips during weekdays than during weekends (see Table 16).

**Table 16** The Number of ML and GPL Trips by Day of Week in 2012

Day of Week	ML Trips	Percentage of ML Trips	GPL Trips	Percentage of GPL Trips	Total Trips
Monday	329,986	6.9%	4,482,049	93.1%	4,812,035
Tuesday	385,819	7.4%	4,820,035	92.6%	5,205,854
Wednesday	385,628	7.4%	4,795,011	92.6%	5,180,639
Thursday	400,503	7.6%	4,882,720	92.4%	5,283,223
Friday	370,596	6.9%	4,986,716	93.1%	5,357,312
Weekdays	1,872,532	7.2%	23,966,531	92.8%	25,839,063
Saturday	84,597	2.1%	3,898,710	97.9%	3,983,307
Sunday	54,154	1.6%	3,369,025	98.4%	3,423,179
Weekends	138,751	1.9%	7,267,735	98.1%	7,406,486

A total of 14,986,355 trips in the eastbound direction, and 18,259,194 trips in the westbound direction were identified. However, the eastbound direction has more ML trips than the westbound direction. Recall, Katy Freeway connects the city of Katy in the west to downtown Houston in the east.

**Table 17** The Number of ML and GPL Trips by Direction in 2012

Direction	ML Trips	Percentage of ML Trips	GPL Trips	Percentage of GPL Trips	Total Trips
Eastbound	1,030,946	6.9%	13,955,409	93.1%	14,986,355
Westbound	980,337	5.4%	17,278,857	94.6%	18,259,194
Total	2,011,283	6.1%	31,234,266	93.9%	33,245,549

From the AVI and the HCTRA data sets, this section developed 33,245,549 actual trips (choices) and their alternative trips in 2012. The next sections explain the methodologies used to estimate the option value of the MLs on the freeway in 2012 using these trips.

## 5. METHODOLOGY AND SAMPLE CALCULATIONS

Through a review of the existing studies, this research identified the option value of MLs as a benefit of MLs that has not been considered. Option values in transportation have focused on transit with most existing studies adopting SP data to estimate the option value of public transit service. SP data has been used primarily because option users of public transit services are unknown; therefore their preferences cannot be obtained through RP data. However, option users of MLs could be those who occasionally or never use the MLs but use the parallel GPLs on the freeway. Therefore, the application of RP data to quantify the option value of MLs is possible if travel choices of each traveler on the freeway were obtained. This is exactly the data that were obtained for travelers on the Katy Freeway in 2012 (from January to November) using AVI data from TxDOT and HCTRA.

This section develops two methodologies to estimate the option value of MLs. The first methodology is based on the Small–Rosen log sum method (Small and Rosen 1981), and the second methodology is based on the Black-Scholes call-option pricing method (Black and Scholes 1973). For convenience, this research calls the Small–Rosen log sum method the log sum method and calls the Black-Scholes call-option pricing method the Black-Scholes method. These methodologies are then applied to the data set identified in the previous section. This data set reveals travelers' preferences on the freeway in 2012.

To develop the methodologies, the option users of the MLs are identified by examining each traveler's choices between the MLs and GPLs on the freeway. Then, a lane-choice behavior model between the MLs and the GPLs is developed to estimate utility functions of ML and GPL choices (trips) for the travelers. Recall, in this research, the ML trip implies a paid ML trip and the GPL trip implies a toll-free GPL trip. These utility functions are required for both methodologies. Finally, this section presents two methodologies to estimate the option value of the MLs based on the log sum method and the Black-Scholes method, respectively. This section also provides sample calculations using the methodologies. These calculations can help transportation planners better understand applications of the methodologies presented herein.

### **5.1 Defining Managed Lane Option Users on the Katy Freeway**

Option value refers to the willingness to pay for the option of having a service available for possible use at some time in the future, even if the option may never be used (Wallis, Wignall, and NZ Transport Agency 2012). From this definition, this research assumes that option users of the MLs are those travelers who occasionally or never use the MLs but use the GPLs on the freeway and have a transponder. These travelers could value the MLs for possible use at some time in the future because they can switch their lane to the MLs depending on traffic situations and travel needs.

When planning a trip, most travelers first choose their travel mode. They then choose a route to their destination. The choice between MLs and GPLs is a sub-decision of their entire trip route choices. Therefore, this research assumes all managed lane option users are those who use the Katy Freeway. In addition, in order to use the MLs on

the freeway as an SOV, vehicles have to install a toll tag (transponder). This indicates that users who did not install the toll tag had no intention of using the MLs as an SOV, even occasionally. Therefore, since our data set includes all Katy Freeway travelers who installed a transponder, this research assumes all ML option users and their preferences are included in our data set. Clearly, other ML option users may exist based on other preferences, but we expect very few of these users.

From these assumptions, this research was able to define the option users of the MLs on the Katy Freeway in 2012 as the users who occasionally or never used the MLs in 2012 as included in our dataset and travelers who always used the MLs in 2012 were not option users. The users who occasionally used the MLs in 2012 could easily be considered as the option users of the MLs because they occasionally changed their lane choice from the GPLs to the MLs depending on their needs and were able to value the availability of the MLs as a backup. However, whether the users who never used MLs in 2012 are option users of the MLs requires additional consideration.

As mentioned, automobile travelers may value public transit service as a backup option for cases where they are unable to use their automobile due to breakdown or diverted usage (for example, loaning a car to family member or friend). Even travelers who never used public transit service are able to value the availability of public transit service because they cannot use their automobile and need other alternatives in such cases. Conversely, travelers on the freeway are always able to use the GPLs. Thus, the fact that a traveler never used MLs in 2012 would likely indicate that the traveler will continue to exclusively use GPLs in the future regardless of traffic situations and, hence,



probably does not value MLs as an available option. Therefore, travelers who never used MLs in 2012 would probably not be option users of MLs. However, since it is still uncertain who the managed lane option users are and no previous studies have defined option users of MLs, this research considers two cases to estimate the option value of MLs in 2012: 1) Case 1 only includes the users who occasionally used the MLs in 2012 as option users and 2) Case 2 includes the users who occasionally or never used the MLs in 2012 as option users.

The data set identified in the previous section had 33,245,549 choices (trips) on the Katy Freeway in 2012. These trips include a total of 2,011,283 ML trips, and 31,234,266 GPL trips. From the dataset, this research identified a total of 1,604,804 travelers (vehicles) who traveled at least once with a toll tag (transponder) on the freeway in 2012. Table 18 classifies the 1,604,804 travelers by the percentage of trips that were on the MLs and the total number of trips in 2012 of each traveler. A total of 1,347,020 (83.9 percent) travelers never used the MLs in 2012 (see column a in Table 18), and a total of 214,859 (13.4 percent) travelers occasionally used the MLs (see column b in Table 18). A total of 42,925 (2.7 percent) travelers always used the MLs only in 2012 (see column c in Table 18), but they do not meet our definition of the option users. Thus, Case 1 considers the 214,859 travelers and Case 2 considers the 1,561,879 ( $=1,347,020+214,859$ ) travelers for the estimation of the option value of the MLs. The 214,859 travelers in Case 1 made a total of 13,503,494 trips on the freeway and 86.8 percent (11,721,575 trips) of the trips were GPL trips in 2012 (see Table 19). The 1,561,879 travelers in Case 2 made a total of 33,016,185 trips on the freeway and

94.6 percent (31,234,266 trips) of the trips were GPL trips in 2012 (see Table 19). In Case 1, the 13,503,494 trips are used to estimate utility functions of the ML and the GPL choices for the travelers, and, in Case 2, the 33,016,185 trips are used to estimate utility functions of the ML and the GPL choices for the travelers. Then, in each case, only GPL choices (trips) are used to estimate the option value of the MLs because, from ML choices, travelers value the MLs as the actual use value, not the option use value. These are 11,721,575 GPL trips and 31,234,266 GPL trips in Cases 1 and 2, respectively. A brief summary of the number of option users and their number of trips in each case is given in Table 19.

**Table 18** Classification of Katy Freeway Travelers by percentage of ML trips

Number of Travelers (Vehicles) on the Katy Freeway in 2012					
		Percentage of Trips that were on the MLs in 2012			
		a) 0%	b) 1-99%	c) 100%	Total
Total Number of Trips on the Katy Freeway in 2012	1-50	1,259,848	134,188	42,044	1,436,080 (89.5%)
	51-100	50,635	36,420	503	87,558 (5.5%)
	101-150	17,158	17,826	184	35,168 (2.2%)
	151-200	8,080	10,589	91	18,760 (1.2%)
	201-250	4,687	6,774	59	11,520 (0.7%)
	251-300	2771	4269	23	7,063 (0.4%)
	301-350	1812	2570	13	4,395 (0.3%)
	351-400	1054	1291	4	2,349 (0.1%)
	401+	975	932	4	1,911 (0.1%)
	Total	1,347,020 (83.9%)	214,859 (13.4%)	42,925 (2.7%)	1,604,804 (100.0%)

Note: a) indicates the users who never used the MLs in 2012, b) indicates the users who occasionally used the MLs in 2012, and c) indicates the users who always used the MLs in 2012.

**Table 19** Summary of Cases 1 and 2

	Option Users	Total Number of Option Users	Total Number of Option Users' Trips (MLs + GPLs Trips)	Total Number of Option Users' GPL Trips
Case 1	b)	214,859	13,503,494	11,721,575
Case 2	a) + b)	1,561,879	33,016,185	31,234,266

Note: a) indicates the users in column a and b) indicates the users in column b from Table 18

## 5.2 Utility Functions of ML and GPL Choices on the Katy Freeway

The main purpose of economic valuation for a non-market service is estimation of the change in consumer surplus. This change in consumer surplus represents WTP to obtain the service. The change in consumer surplus is generally measured by a change in utility caused from using the service. Thus, the economic valuation starts from estimating a utility function of the service that can estimate the utilities of all individuals who use the service.

Our methodologies to estimate the option value of the MLs begin with estimation of the option users' utility function of each choice (ML and GPL choices) as defined in the previous section. To estimate utility functions, it is necessary to model the choice each traveler makes between the MLs and the GPLs. Thus, this research first determined attributes of each alternative or individual who made the choice that needed to be included in the lane-choice behavior model. The most fundamental attribute that affects the choice of lanes could be the monetary cost of using each alternative, especially the amount of toll paid. Travelers generally place value on their time. Thus, the other crucial attribute could be the travel time of each alternative. Both attributes (cost and time) are generally included in most lane-choice behavior models. This research identified some

potential attributes, such as trip length and the standard deviation of travel time (a measure of travel time reliability). But, in general, trip length is significantly correlated with travel time. In addition, inclusion of travel time reliability in the model caused counter-intuitive results. Since the travel time reliability was measured by the standard deviation of travel time in this research, an increase in the standard deviation of travel time (decrease in the reliability) should lead to a decrease in utility. However, in the model, an increase in the standard deviation of travel time caused an increase in utility because the estimate of the standard deviation coefficient was positive. This research also examined additional attributes, including: direction, the time of day (peak, shoulder, and off-peak hours), and day of week. These attributes were also tested in the lane choice model, but no meaningful result was obtained. The direction attribute was significant but produced counter-intuitive results. Since the direction attribute has only two values, west and east directions, two values should have an opposite effect on utility; hence, if a westbound trip leads to a decrease in utility, an eastbound trip should lead to an increase in utility. However, both directions lead to a decrease in utility. The time of day improved explanatory power of the model but the time of day is already partially reflected in the travel time attribute and the effect of the time of day attribute on utility was much larger than the effect of the travel time attribute. For example, when this research included both travel time and time of day attributes in the model, their coefficient estimates were  $-0.1185$  and  $-1.2019$ , respectively. This seemed far too great an impact for time of day since toll and travel time were already in the model and accounted for those shifts. Finally, day of week was correlated with time of day because

weekends were included in off-peak hours. Except for the random ID, no information about each traveler, such as income, gender, and purpose of trip, was available. For these reasons, the lane choice model only included the amount of toll paid and travel time in the utility functions of ML and GPL choices. The utility functions were estimated using a standard logit model. The logit model inherently assumes that travelers have information about the attributes of each alternative, which is the amount of toll paid and travel time. The toll rate was provided in advance and could also be found online. Travelers could obtain travel time information through several sources including their own experience, media reports, roadside electronic message signs, and the Houston Transtar website. Thus, it could be reasonable to expect that travelers had a reasonable estimate of their expected travel time for their alternatives. Utility functions of ML and GPL choices are shown in Equation 7:

$$V_{MLS} = \beta_{TT} \cdot Travel\ Time_{MLS} + \beta_{Toll} \cdot Toll_{MLS} \quad (7)$$

$$V_{GPLS} = \beta_{TT} \cdot Travel\ Time_{GPLS} + \beta_{Toll} \cdot Toll_{GPLS}$$

where:

$V_i$  = utility derived by choosing lanes i,

i = MLs or GPLs,

$\beta_j$  = coefficients to be estimated,

j = TT (travel time) or toll,

$Travel\ Time_i$  = travel time on lanes i,

$Toll_i$  = the amount of toll paid on lanes i.

The Statistical Analysis System (SAS) was used to estimate the coefficients in the utility functions. For Case 1, revealed preferences of the option users were used to estimate the utility functions and these were 13,503,494 choices (trips). For Case 2, revealed preferences of the option users were used to estimate the utility functions with 33,016,185 choices. Table 20 provides the coefficient estimates for the utility functions in each case. Intuitively, decreases in travel time and the amount of toll paid should lead to increases in utility. In both cases, the coefficients of both travel time and the amount of toll paid attributes were negative as expected and were significant based on t-statistics. The results suggest the value of travel time of \$14/hour for the option users in Case 1 and the value of travel time of \$5.97/hour for the option users in Case 2. This appears reasonable as Case 1 includes only travelers who have used the MLs while Case 2 includes many travelers who have not used the MLs in 2012. The correlation coefficients between travel time and the amount of toll paid are provided in Table 21. Based on the correlation coefficients, the attributes are not correlated and independence between the attributes is well maintained in the utility functions in both cases. The rest of this chapter details the methodologies that were used to estimate the option value of MLs partially based on these utility functions.

**Table 20** Coefficient Estimates for the Utility Functions of ML and GPL Choices

Variables	a) Coefficient Estimates in Case 1	b) Coefficient Estimates in Case 2
Travel Time	***-0.2489	***-0.2394
t-statistics	-729.07	-778.12
Toll	***-1.0667	***-2.4061
t-statistics	-1401.8	-2865.7
Log-likelihood at Zero	-9359909	-22885076
Log-likelihood at Convergence	-7857854	-14207810
$\rho^2$ w.r.t. Zero*	0.1605	0.3792
PCP for All Trips** (%)	84.84	93.94
PCP for ML Trips (%)	7.08	0.97
PCP for GPL Trips (%)	96.66	99.25
Value of Travel Time (\$/hour)	14.00	5.97

\*: McFadden's Likelihood Ratio Index.

\*\* : PCP represents a percentage of correct predictions.

\*\*\*: Statistically significant at 0.01 significance level ( $p < 0.01$ ).

**Table 21** Correlation Coefficient between Travel Time and the Amount of Toll Paid

	Case 1	Case 2
Correlation Coefficient between Travel Time and Toll	0.1695	0.1202

### 5.3 Methodology to Estimate the Option Value based on the Log Sum Method

The first methodology in this section estimates the option value of the MLs by measuring change in consumer surplus between the situation where both MLs and GPLs are available and the situation where only the GPLs are available based on the log sum method (Small and Rosen 1981). This methodology assumes that travelers would have a higher consumer surplus when both MLs and GPLs are available rather than when only GPLs are available, even though utilities of the GPL choice remain the same in both

situations. This methodology was used by Chu and Polzin (1998) to estimate the option value of public transportation, but this research modified it according to our choice alternatives, the ML and the GPL choices. The following is a description of the steps for this methodology:

### **Step 1: Estimating Generalized Trip Costs of GPL and ML Trips**

Equation 7, in the previous section, provides utility functions for ML and GPL trips (choices). This equation can be re-written as:

$$V_{MLS} = \beta_{Toll} \left( \frac{\beta_{TT}}{\beta_{Toll}} \cdot Travel\ Time_{MLS} + Toll_{MLS} \right) \quad (8)$$

$$V_{GPLs} = \beta_{Toll} \left( \frac{\beta_{TT}}{\beta_{Toll}} \cdot Travel\ Time_{GPLs} + Toll_{GPLs} \right)$$

The terms in the parenthesis in Equation 8 imply the generalized trip cost of making one trip on the MLs and the GPLs, respectively. Thus, this research defines the generalized trip cost (G) as:

$$G_{MLS} = \frac{\beta_{TT}}{\beta_{Toll}} \cdot Travel\ Time_{MLS} + Toll_{MLS} \quad (9)$$

$$G_{GPLs} = \frac{\beta_{TT}}{\beta_{Toll}} \cdot Travel\ Time_{GPLs} + Toll_{GPLs}$$



### Step 2: Measuring Consumer Surplus

From Equation 9, the probability (P) of an individual choosing alternative i (MLs or GPLs) can be estimated by:

$$P_i = \frac{e^{\beta_{Toll} \cdot G_i}}{\sum_j e^{\beta_{Toll} \cdot G_j}} \quad (10)$$

where:

j = MLs and GPLs.

The denominator in Equation 10 gives the maximum satisfaction that an individual can get from the choice situation. This maximum satisfaction can be used to measure the consumer surplus (CS) to the individual as follows (Small and Rosen 1981):

$$CS = -\frac{1}{\beta_{Toll}} \ln \sum_j e^{\beta_{Toll} \cdot G_j} \quad (11)$$

This way of measuring consumer surplus is called the Small–Rosen log sum method and most experts believe it to be the most comprehensive method to measure consumer surplus in choice situations (Cambridge Systematics, Inc. 1998).

### Step 3: Estimating the Option Value of the MLs

Let  $G_{MLS}$  and  $G_{GPLs}$  be the generalized trip costs on the MLs and the GPLs, respectively. From Equation 11, in the choice situation where both MLs and GPLs are available, consumer surplus per person-trip can be defined as:

$$CS = -\frac{1}{\beta_{Toll}} \ln[e^{\beta_{Toll} G_{MLS}} + e^{\beta_{Toll} G_{GPLs}}] \quad (12)$$

The absence of the MLs would lead to change in  $G_{MLS}$  from the current value  $G_{MLS}$  to an infinitely large value  $G_{MLS}^A$  (i.e., an infinite value is equivalent to services not being available) where the subscript A indicates the case without MLs. The absence of MLs would also change  $G_{GPLs}$  from the current value  $G_{GPLs}$  to  $G_{GPLs}^A$ . In order to estimate the effect of the absence of the MLs on the GPLs, this research assumes that all travelers on the MLs will use the GPLs when the MLs are unavailable. Note that travel times and congestion on the GPLs do not change for this analysis. Therefore, when the MLs are unavailable, consumer surplus per person-trip can be defined as:

$$CS_A = -\frac{1}{\beta_{Toll}} \ln[e^{\beta_{Toll} G_{GPLs}^A}] \quad (13)$$

The difference in consumer surplus per person-trip with and without the MLs is:

$$\Delta CS = CS - CS_A \quad (14)$$

Equation 14 can be broken into two parts:

$$\Delta CS = [CS - CS^0] + [CS^0 - CS_A] \quad (15)$$

where:

$$CS^0 = -\frac{1}{\beta_{Toll}} \ln[e^{\beta_{Toll} G_{GPLs}}]$$

$CS^0$  refers to consumer surplus when GPLs are only available assuming that the performance of GPLs without MLs is the same as the performance of GPLs with MLs. Thus, the first component in Equation 15  $[CS - CS^0]$  estimates change in consumer surplus caused by difference in the availability of the MLs without change in the performance of the GPLs. Conceptually, this translates into the option value of MLs in the user's GPL trip. The second component,  $[CS^0 - CS_A]$ , estimates change in consumer surplus caused by change in the performance of GPLs due to the MLs being available. The first component of Equation 15 implies the option value of the MLs per person-trip and is rewritten as:

$$\begin{aligned}
OV_{MLS} &= -\frac{1}{\beta_{Toll}} [\ln(e^{\beta_{Toll} \cdot G_{MLS}} + e^{\beta_{Toll} \cdot G_{GPLs}}) - \ln(e^{\beta_{Toll} \cdot G_{GPLs}})] \\
&= -\frac{1}{\beta_{Toll}} [\ln(e^{(\beta_{Toll} \cdot G_{MLS} - \beta_{Toll} \cdot G_{GPLs})} + 1)] \\
&= -\frac{1}{\beta_{Toll}} [\ln(e^{-(\beta_{Toll} \cdot G_{GPLs} - \beta_{Toll} \cdot G_{MLS})} + 1)] \\
&= -\frac{1}{\beta_{Toll}} [\ln(1 + e^{(\beta_{Toll} \cdot G_{GPLs} - \beta_{Toll} \cdot G_{MLS})}) - \ln(e^{(\beta_{Toll} \cdot G_{GPLs} - \beta_{Toll} \cdot G_{MLS})})]
\end{aligned} \tag{16}$$

where:  $OV_{MLS}$  = Option Value of MLs.

This research uses Equation 16 to estimate the option value of the MLs per person-trip. An annual option value for each option user is the sum of those option values estimated from all his/her GPL trips in 2012.

## 5.4 Sample Estimations Using the Log Sum Method

This section estimates the option value of the MLs for three travelers using the log sum method. This section first provides three travelers' choice records in 2012 that included actual and alternative choices. Then, based on these trips, the option values are estimated.

Three option users who occasionally used the MLs in 2012 (the option users in Category b in Table 18) were selected at random from the dataset. These users are included in both Cases 1 and 2 (Case 1 only includes the users who occasionally used the MLs in 2012 as the option users and Case 2 includes the users who occasionally or never used the MLs in 2012 as the option users). Table 22 provides choice records in 2012 for the three travelers. The lane choice column indicates whether the record is an actual choice (Lane Choice=1) or an alternative choice that was available but was not chosen (Lane Choice=0). The total trips in 2012 column indicates the total number of trips (ML and GPL trips) and the total ML trips in 2012 column indicates the total number of ML trips on the freeway in 2012 by the user. From Table 22, Traveler A traveled the freeway twice, once on the GPLs and once on the MLs. Traveler B traveled the freeway three times, twice on the GPLs and once on the MLs, and Traveler C traveled the freeway five times, three times on the GPLs and twice on the MLs, in 2012.

**Table 22** Three Option Users' Choice Records in 2012

Traveler	Option User Category	Trip		Lane Choice	Travel Time (minutes)	Toll (\$)	Total Trips in 2012	Total ML Trips in 2012
		No.	Type					
A	b*	1	GPL	1	12.22	0	2	1
		1	ML	0	10.11	1.4	2	1
		2	GPL	0	24.97	0	2	1
		2	ML	1	11.42	4	2	1
B	b*	1	GPL	1	20.92	0	3	1
		1	ML	0	20.82	1	3	1
		2	GPL	1	13.12	0	3	1
		2	ML	0	11.37	0.7	3	1
		3	GPL	0	22.28	0	3	1
		3	ML	1	20.13	1	3	1
C	b*	1	GPL	1	7.62	0	5	2
		1	ML	0	6.32	2	5	2
		2	GPL	1	8.57	0	5	2
		2	ML	0	6.57	2	5	2
		3	GPL	1	8.72	0	5	2
		3	ML	0	6.38	2	5	2
		4	GPL	0	23.94	0	5	2
		4	ML	1	16.78	3.4	5	2
		5	GPL	0	14.48	0	5	2
		5	ML	1	7.98	4	5	2

\*: The option user category b indicates the users who occasionally used the MLs in 2012 in Table 18.

The option values of the MLs for the three option users in Table 22 are then estimated using the log sum method. As mentioned, the option value is estimated based only on the GPL trips. For example, for Traveler A, only Trip 1 is used to estimate the option value of the MLs (see Table 22).

To determine the option value using the log sum method, the first task is to estimate the generalized trip cost on each lane (either the MLs or the GPLs) using Equation 9. For example, the generalized trip cost (G) on each lane in Trip 1 of Traveler A is estimated in Equation 17. Note that since all option users in Table 22 are those who

occasionally used MLs in 2012, the utility functions of Case 1 in Table 20 are used for all sample estimations in this section.

$$G_{MLS} = \frac{\beta_{TT}}{\beta_{Toll}} \cdot Travel\ Time_{MLS} + Toll_{MLS} = \frac{-0.2489}{-1.0677} \cdot 10.11 + 1.4 = \$3.759 \quad (17)$$

$$G_{GPLs} = \frac{\beta_{TT}}{\beta_{Toll}} \cdot Travel\ Time_{GPLs} + Toll_{GPLs} = \frac{-0.2489}{-1.0677} \cdot 12.22 + 0 = \$2.851$$

Then, those generalized trip costs are applied to Equation 16 to estimate the option values of MLs. Hence, for Trip 1 of Traveler A, the option value of the MLs is given by Equation 18:

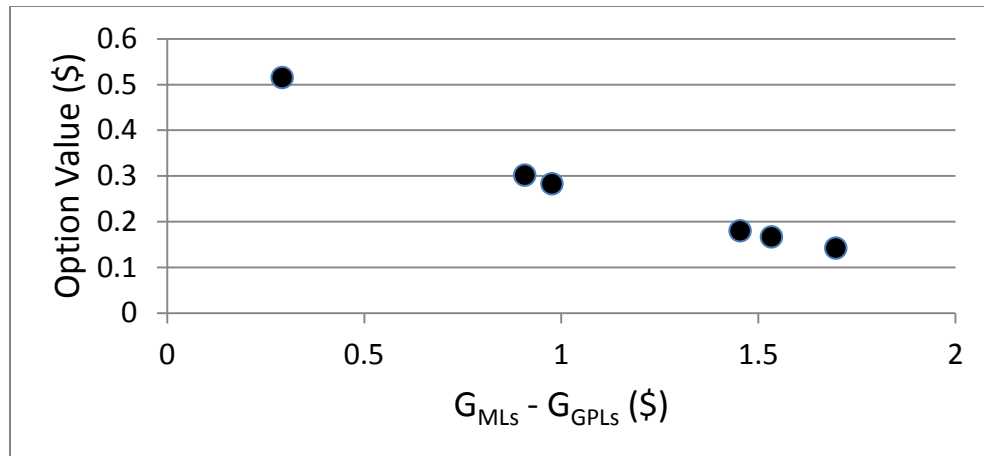
$$\begin{aligned} OV_{MLS} &= -\frac{1}{\beta_{Toll}} [\ln(e^{\beta_{Toll} \cdot G_{MLS}} + e^{\beta_{Toll} \cdot G_{GPLs}}) - \ln(e^{\beta_{Toll} \cdot G_{GPLs}})] \quad (18) \\ &= -\frac{1}{-1.0677} [\ln(e^{-1.0677 \cdot 3.759} + e^{-1.0677 \cdot 2.851}) - \ln(e^{-1.0677 \cdot 2.851})] \\ &= \$0.302 \end{aligned}$$

Using the same procedures in Equations 17 and 18, the option values of the other travelers' GPL choices (trips) were also estimated. Table 23 summarizes these estimation results for each GPL choice (trip). Note that the annual option value for each traveler in 2012 is calculated by summing all option values estimated from each traveler's GPL trips. This summation appears reasonable because travelers generally place a higher value on the option of having the MLs available the more frequently they travel the freeway. Estimates of the annual option values for Travelers A, B, and C were

\$0.302, \$0.790, and \$0.489, respectively. The relationship between the option value using the log sum method and the difference between the generalized trip costs on the MLs and the GPLs is important. The larger the increase in generalized trip costs on the MLs over the GPLs, the lower the option value of the MLs (see fifth and sixth columns in Table 23 and Figure 13). This relationship corresponds with our expectation. If travelers can easily use MLs with a small increase in trip cost, the MLs are more attractive as a backup for their travel. A similar relationship was also expected in the TCRP 78 report (ECO Northwest, Ltd. 2002) as mentioned in Section 3.2.3.

**Table 23** Option Value for Each GPL Choice (Trip) Using the Log Sum Method

Traveler	Trip No.	$G_{MLS}$ (\$)	$G_{GPLs}$ (\$)	$G_{MLS} - G_{GPLs}$ (\$)	$OV_{MLS}$ (\$)	Annual Option Value for the Traveler (\$)
A	1	3.759	2.851	0.908	0.302	0.302
B	1	5.858	4.881	0.977	0.283	0.799
	2	3.353	3.061	0.292	0.515	
C	1	3.475	1.778	1.697	0.142	0.489
	2	3.533	2.000	1.533	0.167	
	3	3.489	2.035	1.454	0.180	



**Figure 13** Relationship between the option value using the log sum method and the difference between the generalized trip costs on the MLs and GPLs

### 5.5 Methodology to Estimate the Option Value based on the Black-Scholes Method

The concept of option value is not limited to transportation economics. Other types of services in the economy involve contingency events: i.e., call and put options, insurance policies, mortgage contracts, etc. Because of the prevalence of option value concepts in the economy, economists have developed mathematical procedures in order to estimate option values (ECO Northwest, Ltd. 2002). Thus, the methodology in this section was adopted from the mathematical procedure to value a call-option in stock markets.

TCRP 78 report (ECO Northwest, Ltd. 2002) realized that the Black-Scholes call-option pricing method (Black and Scholes 1973) can be used to estimate the option value of transit. A call-option in the stock markets gives its holder an opportunity to buy a stock at the exercise/strike price on or before the expiration date. The option to buy the underlying stock is exercised if the market value of the stock is greater than the



exercise/strike price and the holder gains a profit from the call-option. Thus, the call-option provides an alternative way to buy the underlying stock if the price of the stock unexpectedly increases. Public transit value to an automobile user depends on factors which are similar to factors for the call-option. As the price of stocks is volatile, the cost of automobile trips is volatile because of the risk of weather, breakdowns, crashes, and other random factors. Public transit provides an alternative way to travel in such cases with a lower trip cost. This relationship between automobile and transit services is similar to the relationship between a stock and its call-option. Furthermore, if auto users were to perceive a strong possibility that the automobile trip cost will move in such a way that the transit trip cost will become more attractive, the transit will have value as a backup option. This is similar to the motive of purchasing a call-option on a stock.

This research proposes that the relationship between a GPL trip and an ML trip is similar to the relationship between an automobile trip and a transit trip. This is because the trip cost on the GPLs is volatile compared to the trip cost on the MLs due to unexpected situations on the GPLs, such as severe traffic congestion and incidents. The MLs provide an alternative way to travel in such situations with a lower trip cost. Therefore, the Black-Scholes method can also be applied to estimate the option value of the MLs. This Black-Scholes method is provided in Equation 19, and Table 24 provides a description of its variables used to estimate the value of a call-option, the option value of transit service, and the option value of the MLs, respectively. Note that this research modified the variables in the Black-Scholes method corresponding to the purpose of this

research to estimate the option value of the MLs with assumptions discussed in the material that follows:

$$C = S \cdot N(d_1) - X \cdot e^{-rT} \cdot N(d_2) \quad (19)$$

where:

$$d_1 = \frac{\ln(S/X) + (r + \sigma^2/2) \cdot T}{\sigma\sqrt{T}}, d_2 = d_1 - \sigma\sqrt{T}$$

**Table 24** Descriptions of Variables in the Black-Scholes Method

Variable	For Stock and Call-Option	For Auto Trip and Transit Trip	For Trip on GPLs and Trip on MLs
C	Value of Call-Option	Option Value of Transit	Option Value of MLs
S	Stock Price	Auto Trip Cost	Generalized Trip Cost on GPLs
X	Exercise/Strike Price	Transit Trip Cost	Generalized Trip Cost on MLs
T (year)	Time until Expiration of the Call-Option	1/(Frequency with which transit is optionally used in a year)	1/(Frequency with which MLs are optionally used in a year)
r	Risk-Free Return	Risk-Free Return	Risk-Free Return
N	Normal Distribution	Normal Distribution	Normal Distribution
$\sigma$	Volatility of $S^2$	Standard Deviation of S	Standard Deviation of S

To estimate the option value of the MLs using the Black-Scholes method, this research assumes that the generalized trip cost on the MLs (the variable X) is always

---

<sup>2</sup> Volatility refers to the amount of uncertainty about the size of changes in a stock's price ("Volatility Definition" 2015). In the Black-Scholes method, volatility was defined as the standard deviation of the rate of return on the stock (Black and Scholes 1973). TCRP 78 report (ECO Northwest, Ltd. 2002) defined the volatility as the standard deviation of auto trip costs. This research follows the definition of volatility in TCRP 78 report — as discussed on the next page.

stable as the exercise price (much like the transit trip cost) because the MLs generally provide reliable travel.

For the stock and call-option, the variable  $T$  (called time until expiration in years) indicates the amount of time between the purchase date and the expiration date of the call-option. On the purchase date, the call-option buyer can choose a range of expiration dates when s/he wants to exercise the call-option, such as one month from now, even though the call-option is not exercised if the exercise price is greater than the market value of the stock on or before the expiration date. Thus, the call-option buyer can decide a specific time point in the future as the expiration date to cover unexpected changes in the stock price on or before the expiration date. Based on the specific time point, the time until expiration (the variable  $T$ ) of the call-option is determined. However, the variable  $T$  for the trips on the GPLs and on the MLs is unknown because the ML option is not specifically traded. Thus, this research needs to find an analogy for the variable  $T$  to estimate the option value of the MLs using the Black-Scholes method. Since the definition of the variable  $T$  for the stock and call-option does not fit for this research, this research adopted the definition of the variable  $T$  in the TCRP 78 report (ECO Northwest, Ltd. 2002) that provides the methodology to estimate the option value of transit. The option of having transit available will only be valuable to auto users at some point during the year. For example, people who drive themselves to work will likely place some value on having transit as an option on those days when their car is being repaired. Thus, a good number to use for the variable  $T$  would be the time period over which the auto user expects to take advantage of the option the first time. The

frequency with which transit might be optionally used in a year could capture this time period. For example, if the auto user expects to use transit 10 times a year, the variable  $T$  can be equal to 0.1 year ( $=1 \text{ year}/10 \text{ times}$ ). This 0.1 year represents the average amount of time over which the auto user takes advantage of the transit option and, in the course of a year, one would expect the auto user to be willing to buy 10 such options of which the time until expiration is 0.1 year (ECO Northwest, Ltd. 2002). Similarly, the ML option will only be valuable to the ML option users at specific points during the year. For example, people who primarily use the GPLs will likely place some value on having the MLs as an option when they encounter unexpected congestion or are in a hurry. Thus, a good number to use for the variable  $T$  would also be the average time period over which the ML option user expects to take advantage of the ML option. Our dataset could estimate this time period as the amount of time between the date when the GPL trip was made and the date when the next ML trip was made for each option user. However, if a traveler traveled once on the freeway in 2012, this estimation cannot be used. Also, the frequency with which the MLs are optionally used in a year can be used as in the TCRP 78 report (ECO Northwest, Ltd. 2002). From the frequency, the average amount of time over which the traveler takes advantage of the ML option can be estimated. Thus, this research uses  $[1/\text{frequency with which MLs are optionally used in a year}]$  as the variable  $T$ . For the frequency, the total number of ML trips in 2012 of each option user was used. In addition, for the option users who never used the MLs in 2012 (included in Case 2), this research assumes the frequency with which MLs are optionally used in a year as one. If their frequency is considered as zero, this implies that they do not expect to use the

MLs at any time. In this case, it is unreasonable to consider them as an ML option user because users who do not expect use the MLs would not value the ML option. These users, who never used MLs, were excluded in Case 1.

The volatility (uncertainty) of a stock price (the variable  $\sigma$  in the last row of Table 24) implies the volatility which will occur over the life of the option. Clearly, this volatility is unknown because it will occur in the future. Thus, in the stock markets, there are various methods to estimate the volatility of the stock price. One of the representative methods is to calculate historical volatility. Historical volatility depends on the historical period over which the volatility is to be calculated. The historical period may be ten days, six months, or any period that the trader chooses. Longer periods tend to yield an average volatility, while shorter periods may explain unusual extremes in the volatility. Thus, if option traders are selling long-term options, the long-term data is generally used. If the option traders are selling short-term options, the short-term data is generally used (Natenberg 1994).

Table 25 provides the number of the option users (who occasionally used the MLs in 2012) by their total number of ML trips in 2012. 84.7 percent of the option users used the MLs 10 or fewer times in 2012 (less than once per month). Furthermore, 42.4 percent of the option users used the MLs only once in 2012. These results indicate that using the long-term GPL volatility may be appropriate because the average period over which the ML option was used is greater than one month. Thus, the variable  $\sigma$  is defined as the standard deviation of the generalized trip costs on the GPLs of all users who traveled the same GPL section at the same 10-minute interval of every day in 2012 as

the option user did (the long-term volatility). However, to estimate the standard deviation of the generalized trip costs on the GPLs, weekend and weekday trip costs on the GPLs are separated. Though this research primarily uses the standard deviation of the generalized trip costs on the GPLs for one year as the long-term volatility, the option value using the standard deviation of the generalized trip costs on the GPLs during the month in which the option user's trip was made is also estimated (in Section 6.5.3) to identify the effect of using a short-term volatility on the option value.

**Table 25** The Number Option Users by their Total Number of ML Trips in 2012

The Total Number of ML Trips in 2012	The Number of the Option Users (included in Category b*)	Percentage (%)
1	91,078	42.4
2	33,479	15.6
3	17,816	8.3
4	11,383	5.3
5	7,891	3.7
6	5,884	2.7
7	4,532	2.1
8	3,806	1.8
9	3,107	1.5
10	2,671	1.2
10+	33,212	15.3
Total	214,859	100.0

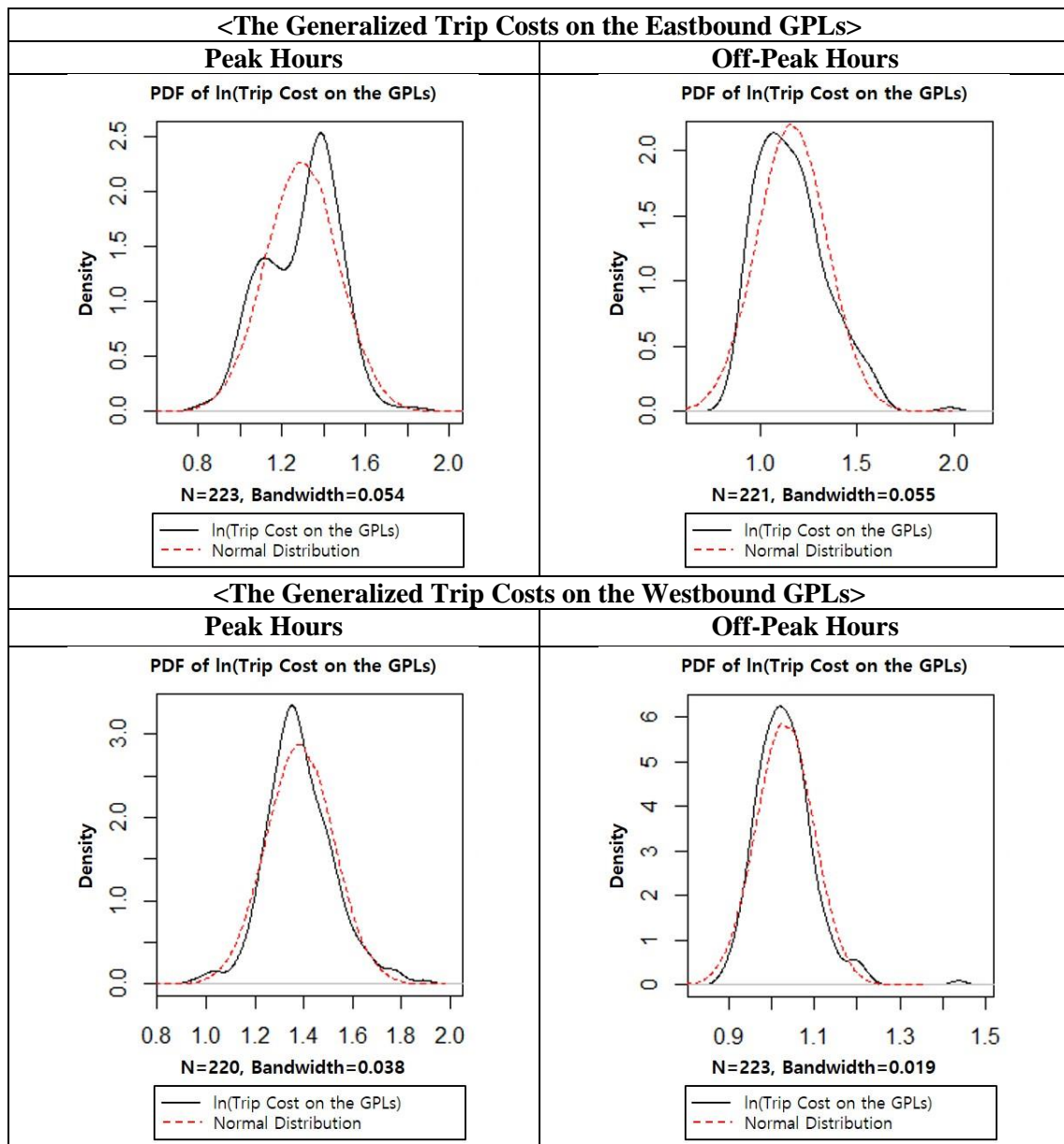
\*: The option user category b indicates the users who occasionally used the MLs in 2012 in Table 18.

Another important assumption of the Black-Scholes method is that the distribution of possible stock prices at the end of any finite interval is log-normal. From this assumption, the natural log in  $d_1$  and the normal distribution (the operator N) in

Equation 19 were derived. In order to determine whether this assumption is also appropriate for the managed lane option value estimation using the Black-Scholes method, the distribution of the generalized trip costs on the GPLs was examined. Since travel time on the GPLs was the main component of the generalized trip cost on the GPLs, this research first estimated travel times on a specific GPL section in each direction at the specific time period of every day. The section between the sensors 427 and 426 on the eastbound freeway and the section between the sensors 425 and 396 on the westbound freeway were selected to estimate the travel times (see Figure 5 for the locations). Then, on the eastbound GPL section, the average travel time between 7:20 AM and 7:40 AM and the average travel time between 17:20 PM and 17:40 PM were estimated for every day in 2012. On the westbound GPL section, the average travel time between 16:20 PM and 16:40 PM and the average travel time between 9:20 AM and 9:40 AM were estimated for every day in 2012. The generalized trip costs on both GPL sections were estimated by multiplying the estimated average travel times and the value of travel time parameter in Case 1 (\$14/hour) to identify the distributions of the generalized trip costs on both GPL sections for peak and off-peak hours, respectively. The generalized trip costs for weekdays in 2012 were only included and the generalized trip costs for holidays and weekends were excluded because, in general, weekday trips show different trip characteristics from holiday/weekend trips. Finally, the generalized trip costs of 223 weekdays for the peak period and the generalized trip costs of 221 weekdays for the off-peak period were used to plot the distributions of the generalized trip costs on the eastbound GPL section. The generalized trip costs of 220 weekdays for

the peak period and the average travel times of 223 weekdays for the off-peak period were used for the westbound GPL section. Figure 14 provides the distribution of the natural logarithm generalized GPL trip costs and a normal distribution. In the distribution of the natural logarithm generalized GPL trip costs, the dotted line indicates the normal distribution. On the westbound section, the distributions of the natural logarithm generalized GPL trip costs for both peak and off-peak hours are very close to the normal distribution. In addition, the normality test using the Shapiro-Wilk test (1965) provided the p-values of 0.04, 0.30, 0.25, and 0.66 for the natural logarithm generalized GPL trip costs of the eastbound peak hours, the eastbound off-peak hours, the westbound peak hours, and the westbound off-peak hours, respectively. The null hypothesis of the Shapiro-Wilk test (1965) is that the population is normally distributed. If the p-value is less than the chosen significance level, the null hypothesis is rejected and there is evidence that the data tested are not from a normally distributed population. Thus, those p-values imply that the natural logarithm generalized GPL trip costs of the eastbound peak hours, the eastbound off-peak hours, the westbound peak hours, and the westbound off-peak hours came from a normally distributed population at a significance level of 0.01, respectively. Thus, the assumption that the distribution of the generalized trip costs on the GPLs is log-normal can also be appropriate for the managed lane option value estimation.





**Figure 14** Distribution of the Generalized Trip Costs on the GPLs and a Normal Distribution

The generalized trip costs on the GPLs and the MLs (the variables S and X) are estimated using Equation 9 in Section 5.3 and, for the risk-free return (the variable R), 0.20 percent which was the interest rate on a twelve-month U.S. Treasury bill in 2012

(U.S. Department of the Treasury 2015) is used (see the last column in Table 24 for descriptions of the variables) because, in practice, the interest rate on a U.S. Treasury bill is commonly considered as the risk-free return in financial markets (“Risk-Free Return” 2015). Finally, this research uses Equation 19 to estimate the option value of the MLs per person-trip based on those defined variables. An annual option value for each option user is the sum of those option values estimated from all his/her GPL trips in 2012.

### **5.6 Sample Estimations Using the Black-Scholes Method**

This section provides sample estimations using the Black-Scholes method for the option value of MLs. As in Section 5.4, the same trip records in 2012 of the option users in Table 22 were used for the sample estimations. Table 26 provides values of the variables in the Black-Scholes method (Equation 19) for each choice record. The generalized trip costs on the GPLs and the MLs (the variables  $S$  and  $X$  in the Black-Scholes method) were calculated in the same way as in the log sum method. Travelers A and B traveled the MLs once in 2012, and Traveler C traveled the MLs twice in 2012. Thus, from the total numbers of ML trips in 2012, the variable  $T$  (1/frequency with which the MLs are optionally used in a year) was estimated for each traveler. The variable  $r$  (risk-free return) were set based on assumptions in Section 5.5. The standard deviation of the generalized trip costs on the GPLs (the variable  $\sigma$  or  $\sigma_m$ ) was estimated from generalized trip costs on the GPLs of all users who traveled the same GPL section at the same 10-minute interval of every day either in 2012 (for variable  $\sigma$ ) or during the same month (for variable  $\sigma_m$ ) as the option user did. However, to estimate the standard

deviation of the generalized trip costs on the GPLs, weekend and weekday trip costs on the GPLs were separated because, in general, weekday trips show different trip characteristics from weekend trips, as mentioned. Note that since Traveler C traveled the same GPL section in Trips 1 and 2, the variable  $\sigma$  (or  $\sigma_m$ ) has the same value.

**Table 26** Values of the Variables in the Black-Scholes Method for Each Choice (Trip)

Variable	A-1**	B-1**	B-2**	C-1**	C-2**	C-3**
S (\$)	2.851	4.881	3.061	1.778	2.000	2.035
X (\$)	3.759	5.858	3.353	3.475	3.533	3.489
T (year)	1	1	1	1/2	1/2	1/2
R (%)	0.20	0.20	0.20	0.20	0.20	0.20
$\sigma^*$ (\$)	0.764	0.455	0.304	0.223	0.223	0.237
$\sigma_m^*$ (\$)	1.015	0.457	0.257	0.208	0.208	0.233

Note: The variables S, X, T, and r indicate the generalized trip cost on the GPLs, the generalized trip cost on the MLs, 1/Frequency with which the MLs are optionally used in a year, and risk-free return, respectively (see Section 5.5 for more details).

\*:  $\sigma$  indicates the standard deviation of S in 2012, and  $\sigma_m$  indicates the standard deviation of S during the month when the trip was made (see Section 5.5 for more details).

\*\* : A, B and C indicate Travelers A, B, and C, respectively, and 1, 2, and 3 indicate Trips 1, 2 and 3, respectively.

Except for the variable  $\sigma_m$ , Table 26 values were first applied to Equation 19 (the Black-Scholes method) to estimate the option value of the MLs for each GPL trip. Hence, in Trip 1 of Traveler A, the variables  $d_1$  and  $d_2$  in the Black-Scholes method were first calculated by Equation 20:

$$d_1 = \frac{\ln(S/X) + (r + \sigma^2/2) \cdot T}{\sigma\sqrt{T}} = \frac{\ln(2.851/3.759) + (0.002 + 0.764^2/2) \cdot 1}{0.764\sqrt{1}} \quad (20)$$

$$= 0.023$$

$$d_2 = d_1 - \sigma\sqrt{T} = 0.023 - 0.764\sqrt{1} = -0.741$$

Then, the option value of the MLs for Trip 1 of Traveler A is calculated by Equation 21:

$$\begin{aligned} C &= S \cdot N(d_1) - X \cdot e^{-rT} \cdot N(d_2) \\ &= 2.851 \cdot N(0.023) - 3.759 \cdot e^{-(0.002 \cdot 1)} \cdot N(-0.741) = \$0.591 \end{aligned} \quad (21)$$

Using the same procedures in Equations 20 and 21, the option values of the other travelers' GPL trips were also estimated. Table 27 summarizes these estimation results for each GPL choice (trip). The annual option value for each traveler in 2012 was also calculated by summing all option values estimated from each traveler's GPL trips.

Next, the option value using the standard deviation of the generalized trip costs on the GPLs during the month when the GPL trip was made (the variable  $\sigma_m$ ) was also estimated. Except for Trip 1 of Traveler A, the option values using the standard deviation of the generalized trip costs on the GPLs for all of 2012 ( $\sigma$ ) are similar to the option values using the standard deviation of the generalized trip costs on the GPLs during one month ( $\sigma_m$ ) (see Table 27).

For Travelers A, B, and C, the annual option values of \$0.591, \$0.816, and \$0.0001 were estimated using the standard deviation of the generalized trip costs on the GPLs in 2012. Since Travelers A and B had larger generalized trip costs on the GPLs (the variable  $S$ ) than Traveler C and they traveled the GPL sections where the standard deviation (variability) of the generalized trip costs (the variable  $\sigma$ ) was higher than the

GPL sections where Traveler C traveled, the annual option values of Travelers A and B were much greater than that of Traveler C (see Table 26 for the variables  $S$  and  $\sigma$ ).

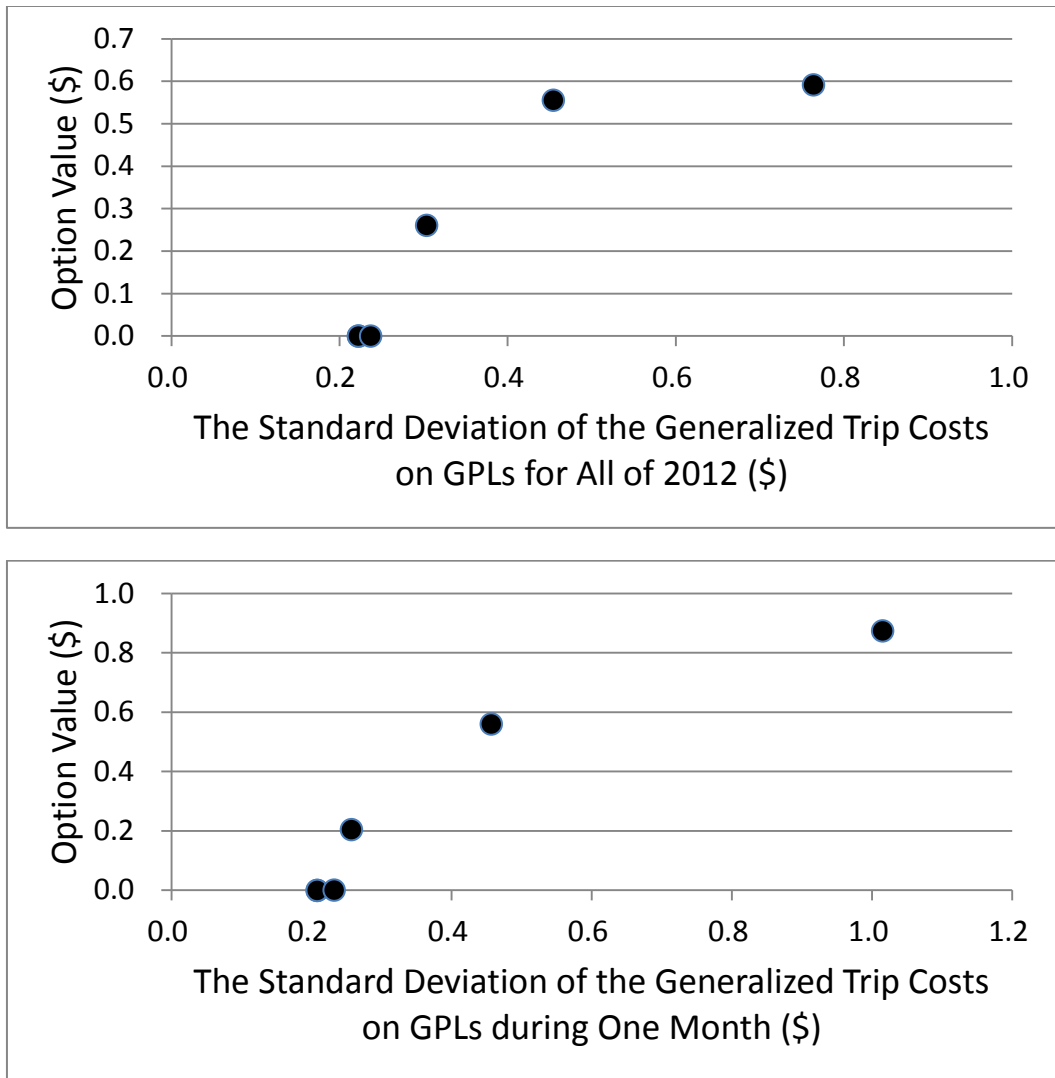
Finally, the relationship between the option value using the Black-Scholes method and the standard deviation of the generalized trip costs needs to be noted. The higher the standard deviation of the generalized trip costs on the GPLs, the higher the option value of the MLs (see second and fourth rows in Table 27 and Figure 15). This relationship corresponds with our expectation because high variability in the trip cost on GPLs increases the chances that MLs may be needed. A similar relationship was also expected in the TCRP 78 report (ECO Northwest, Ltd. 2002) as mentioned in Section 3.2.3.

**Table 27** Option Value for Each GPL Choice (Trip) using the Black-Scholes Method

Value (\$)	A-1	B-1	B-2	C-1	C-2	C-3
$\sigma$ ( $\sigma$ in minutes)	0.764 (3.27)	0.455 (1.95)	0.304 (1.30)	0.223 (0.96)	0.223 (0.96)	0.237 (1.02)
$\sigma_m$ ( $\sigma_m$ in minutes)	1.015 (4.35)	0.457 (1.96)	0.257 (1.10)	0.208 (0.89)	0.208 (0.89)	0.233 (1.00)
Option Value using $\sigma$	0.591	0.555	0.261	0.0000	0.0000	0.0001
Option Value using $\sigma_m$	0.874	0.560	0.205	0.0000	0.0000	0.0001
Annual Option Value using $\sigma$	0.591	0.816		0.0001		
Annual Option Value using $\sigma_m$	0.874	0.765		0.0001		

Note:  $\sigma$  indicates the standard deviation of  $S$  in 2012, and  $\sigma_m$  indicates the standard deviation of  $S$  during the month when the trip was made.

Note: A, B and C indicate Travelers A, B, and C, respectively according to Trips 1, 2, and 3, respectively.



**Figure 15** Relationship between the option value using the Black-Scholes Method and the standard deviation of the generalized trip costs on the GPLs

### 5.7 Comparisons of Sample Estimations Using the Log Sum and the Black-Scholes Methods

This section compares the option value using the log sum method and the option value using the Black-Scholes method in the sample estimations in Sections 5.4 and 5.6 (see Table 28). Both methods result in very different option values. This is amplified by

the option values for the trips of Traveler C. The option values using the Black-Scholes method were close to zero, while the option values using the log sum method were \$0.142, \$0.167, and \$0.180, respectively (see the option values of Traveler C in Table 28). For the trips of Traveler C, the ML options have much higher trip costs than the trip costs on the GPLs (approximately 1.7 to 2 times the trip cost on the GPLs) and the standard deviations (variability) of the generalized trip costs on the GPLs were low, approximately \$0.2 (see  $G_{GPLs}$ ,  $G_{MLS}$ , and  $\sigma$  of Traveler C in Table 28). In these cases, the option value of zero estimated from the Black-Scholes method could be more reasonable because Traveler C could rarely gain benefit from using the ML option. For example, in Trip 1 of Traveler C in Table 28, the generalized trip cost on the GPLs is \$1.778 and the variability (standard deviation) of this cost is \$0.223. Since the generalized trip cost on the MLs, \$3.475, is much greater than the sum of the generalized trip cost on the GPLs and the variability, Traveler C could rarely gain benefit even if Traveler C has the ML option. Whereas, since the log sum method assumes that travelers would have a higher consumer surplus when both MLs and GPLs are available rather than when only GPLs are available, the method appears to provide a worse estimate in the case of Traveler C.

**Table 28** Comparison of the Option Value using the Log Sum Method and the Option Value using the Black-Scholes Method

Value (\$)	A-1	B-1	B-2	C-1	C-2	C-3
$G_{GPLs}$ (S)	2.851	4.881	3.061	1.778	2.000	2.035
$G_{MLs}$ (X)	3.759	5.858	3.353	3.475	3.533	3.489
$G_{MLs} - G_{GPLs}$	0.908	0.977	0.292	1.697	1.533	1.454
$\sigma$	0.764	0.455	0.304	0.223	0.223	0.237
Option Value using the Log Sum Method	0.302	0.283	0.515	0.142	0.167	0.180
Option Value using the Black-Scholes Method	0.591	0.555	0.261	0.0000	0.0000	0.0001
Annual Option Value using the Log Sum Method	0.302	0.799		0.489		
Annual Option Value using the Black-Scholes Method	0.591	0.816		0.0001		

Note: A, B and C indicate Travelers A, B, and C, respectively, and 1, 2, and 3 indicates Trips 1, 2 and 3, respectively.

The log sum method considers the difference in the generalized trip costs on the MLs and the GPLs to determine the option value of the MLs. The option value of the MLs increases as the generalized trip cost on the MLs decreases over the generalized trip cost on the GPLs because travelers can easily use the MLs with a small increase in the trip cost. The Black-Scholes method considers how much option users can reduce their trip cost by using the MLs in a contingency event, such as unexpected traffic congestion, to determine the option value of the MLs. Thus, in estimating the option value of managed lanes, the log sum method could be more desirable when option users place a high value on MLs as an additional option that they can easily use with a small increase in the trip cost. The Black-Scholes method could be more desirable when travelers place a high value on MLs as a backup option for contingency events. However, if the GPLs



provide as reliable trip costs as the MLs do and the trip cost on the GPLs is much less than the trip cost on the MLs, the log sum method could provide the less accurate estimate of the option value of the MLs (as in the trips of Traveler C in Table 28). In those cases, the ML option is not valuable to the option user because the user rarely gains benefit from using the MLs. Thus, the Black-Scholes method could be more appropriate to estimate the option value of the MLs.

The next section estimates the option value of the MLs on the Katy Freeway in 2012 (from January to November) using the entire data set based on the two methodologies. In addition, sensitivity analysis for the option value is provided. Based on these results, this research discusses the two methodologies in detail to identify the more appropriate methodology to estimate the option value of MLs.

## 6. OPTION VALUE OF MANAGED LANES ON KATY FREEWAY

This section estimates the option value of the MLs on the Katy Freeway in 2012 (from January to November) based on the option users' GPL choices (trips) identified in Section 4. This research considered two potential definitions of option users: 1) Case 1 only includes the travelers who used the MLs at least once in 2012 and 2) Case 2 includes all Katy Freeway travelers with transponders – even if they did not use the MLs at all in 2012. Thus, for Case 1, 214,859 option users were selected and their 11,721,575 GPL trips were used to estimate the option value of the MLs in 2012. For Case 2, 1,561,879 option users were selected and their 31,234,266 GPL trips were used to estimate the option value of the MLs in 2012. As shown in Table 20, in Section 5.2, each case has different coefficients in their utility functions. From the option users in Case 1, the value of travel time was estimated as \$14/hour and from the option users in Case 2, the value of travel time was estimated as \$5.97/hour. The option value of the MLs was estimated using two methodologies, the log sum method and the Black-Scholes method.

A sensitivity analysis for four parameters was also performed. The four parameters were 1) the value of travel time, 2) the standard deviation of the generalized trip costs on the GPLs, 3) the time period over which the standard deviation of the generalized trip costs on the GPLs was calculated, and 4) the risk-free return. Finally, the option value of the MLs was compared with the value of travel time savings for the option users. The value of travel time savings refers to the benefits from reduced travel time by using the MLs (Litman 2009).

## **6.1 Option Value of the MLs using the Log Sum Method**

This research first estimated the option value of the MLs in 2012 using the log sum method. A total of \$4,576,625 and \$2,975,032 were estimated as the option value of the MLs in 2012 in Cases 1 and 2, respectively, (see Table 29). Recall, for Case 1, the 214,859 option users who occasionally used MLs in 2012 were selected and their 11,721,575 GPL trips were used to estimate the option value of the MLs in 2012. For Case 2, 1,561,879 option users who occasionally or never used the MLs in 2012 were selected and their 31,234,266 GPL trips were used to estimate the option value of the MLs in 2012. Though the number of GPL trips made by the option users in Case 1 was about one third that of the number of GPL trips in Case 2, Case 1 presented 1.5 times the total option value of the MLs in 2012. This was because the larger value of travel time estimated from the option users in Case 1 was applied to estimate the option value of the MLs (i.e., \$14/hour in Case 1 and \$5.97/hour in Case 2). The average option value per person-trip was estimated as \$0.39 and \$0.10 in Cases 1 and 2, respectively. The average option value for each option user for the entire year was estimated as \$21.30 and \$1.91 in Cases 1 and 2, respectively (see Table 29).

**Table 29** Option Value of the MLs using the Log Sum Method

Statistical Measure	Case 1	Case 2
Total Number of GPL Trips	11,721,575	31,234,266
Total Number of Option Users	214,859	1,561,879
Total Option Value	\$4,576,625	\$2,975,032
Average Option Value Per Person-Trip	\$0.390	\$0.095
Standard Deviation of Option Value Per Person-Trip	\$0.198	\$0.069
Average Option Value Per Person	\$21.301	\$1.905
Standard Deviation of Option Value Per Person	\$29.006	\$4.363

About 99 percent of the option values per person-trip were less than \$1 in both cases (see Table 30). However, a few large option value estimates were observed in Case 1 and the maximum option value per person-trip in Case 1 was \$7.07 (see Table 30). Thus, this research needed to identify why the large option values were estimated from the log sum method. In the trip that had the option value of \$7.07, travel time on the MLs (the alternative choice) was 28.6 minutes with a payment of \$1 for a toll, while travel time on the GPLs (the actual choice) was 63.2 minutes. Using the value of travel time of \$14/hour, these travel times and the toll led to \$7.68 and \$14.75 as the generalized trip costs on the MLs and the GPLs, respectively. The difference between the generalized trip costs on the MLs and the GPLs was -\$7.07 ( $=7.68-14.75$ ). This implies that the ML option had the much smaller trip cost than the GPLs. Conversely, in the trip that resulted in almost zero dollars for the option value per person-trip, the difference between the generalized trip costs on the MLs and the GPLs (the generalized ML trip cost – the generalized GPL trip cost) represented a large positive value (\$6.84),

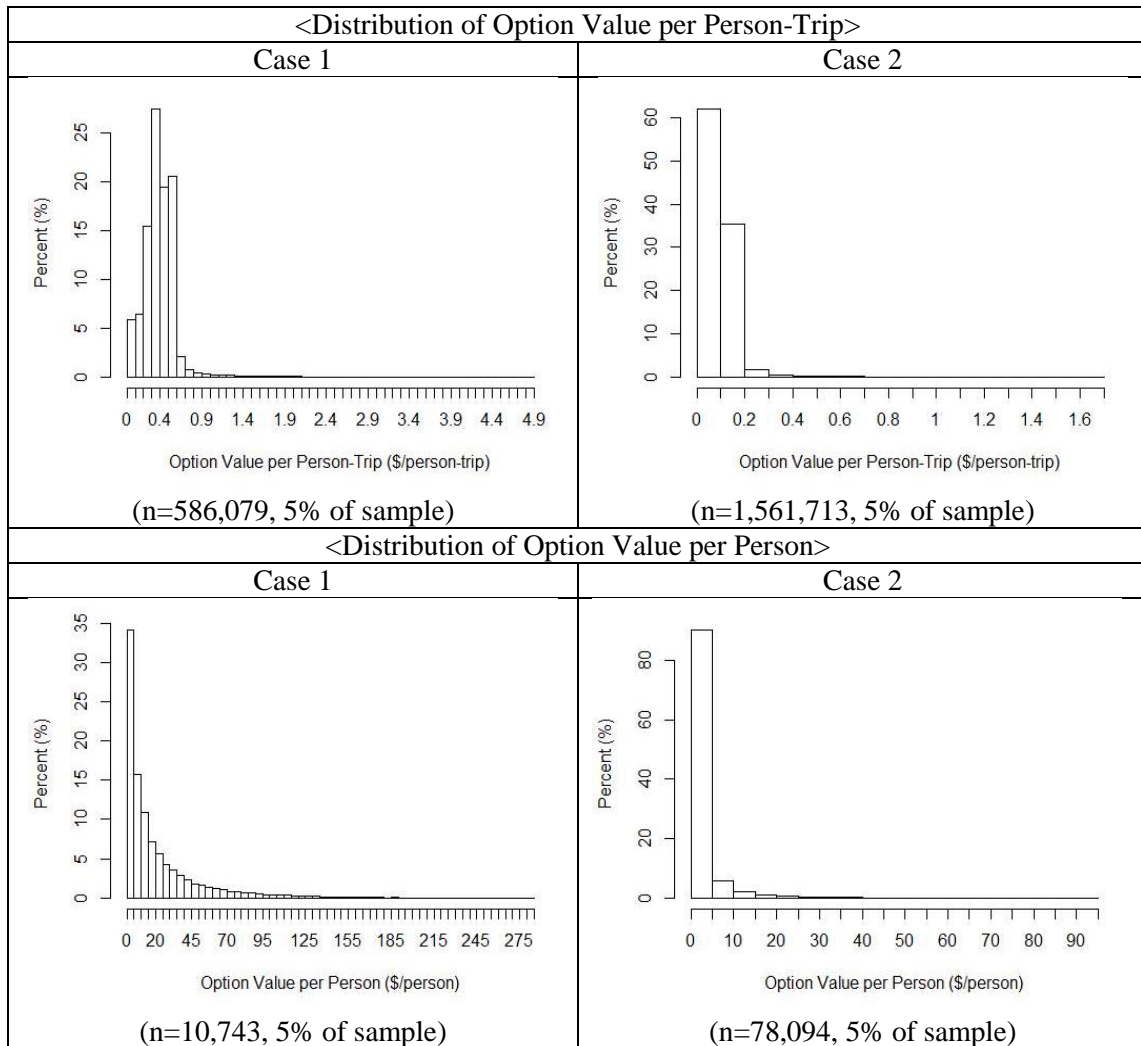
which implies that the ML option had the much larger trip cost than the GPLs. These results coincide with the finding in the sample estimations using the log sum method in Section 5.4: the option value of the MLs decreases as the generalized trip cost on the MLs increases over the generalized trip cost on the GPLs. Specifically, the correlation coefficient between the option value per person-trip and the difference between the generalized trip costs on the MLs and the GPLs (the generalized ML trip cost – the generalized GPL trip cost) was estimated as  $-0.86$ . Thus, those unusually large option value estimates could be reasonable because, in the case where the generalized trip cost on the MLs is much smaller than that on the GPLs, the MLs could be more attractive to travelers on the GPLs as a backup. In addition, the option value of zero dollars could be reasonable in the case where the generalized trip cost on the MLs is much larger than that on the GPLs because the MLs are not useful as a backup in that case.

In both cases, approximately 99% of the option users value the managed lane option less than \$135 for the year. The unusually large option value per person (the maximum option value per person in Case 1 in Table 30) is a result of the large number of GPL trips in 2012, i.e., 1,202 trips in 2012.

**Table 30** Quantiles of the Option Value of the MLs Using the Log Sum Method

Quantiles	Case 1		Case 2	
	Per Person-Trip	Per Person	Per Person-Trip	Per Person
100% Max	\$7.069	\$718.889	\$2.441	\$188.004
99%	\$1.091	\$134.844	\$0.254	\$21.875
95%	\$0.595	\$81.459	\$0.185	\$8.761
90%	\$0.549	\$57.033	\$0.175	\$4.884
75% Q3	\$0.501	\$27.435	\$0.158	\$1.587
50% Median	\$0.383	\$10.161	\$0.079	\$0.446
25% Q1	\$0.283	\$3.102	\$0.041	\$0.166
10%	\$0.173	\$1.027	\$0.015	\$0.073
5%	\$0.084	\$0.532	\$0.002	\$0.048
1%	\$0.022	\$0.273	\$0.000	\$0.015
0% Min	\$0.000	\$0.001	\$0.000	\$0.000

Figure 16 provides the distributions of the option value estimates using the log sum method. Since SAS could not draw the distributions using the more than 10 million observations, this research randomly selected five percent of the option value estimates to display. As shown in Figure 16, most option value estimates for each trip were less than \$1 in Case 1 and most option value estimates for each trip were less than \$0.2 in Case 2. Most option value estimates for the year for each option user were less than \$115 in Case 1, and most option value estimates for option users were less than \$15 in Case 2.



**Figure 16** Distribution of the Option Value using the Log Sum Method

## 6.2 Option Value of the MLs Using the Black-Scholes Method

This research estimated the option value of the MLs in 2012 using the Black-Scholes method. A total of \$1,117,755 and \$221,518 were estimated as the option value of the MLs in 2012 in Cases 1 and 2, respectively (see Table 31). Recall, for Case 1, 214,859 travelers who used the MLs at least once in 2012 were selected and their

11,721,575 GPL trips were used to estimate the option value of MLs in 2012. For Case 2, 1,561,879 option users (any traveler with a transponder who used Katy Freeway in 2012) were selected and their 31,234,266 GPL trips were used to estimate the option value of the MLs in 2012. The total option value estimate in Case 2 was about one fifth of the total option value estimate in Case 1 even though the number of GPL trips made by option users in Case 2 was approximately three times the number of GPL trips in Case 1. This was because the larger value of travel time estimated from Case 1 option users served to estimate the option value of the MLs (i.e., \$14/hour in Case 1 and \$5.97/hour in Case 2). The average option value per person-trip was estimated as \$0.095 and \$0.007 in Cases 1 and 2, respectively. The average option value for each option user for the year was estimated as \$5.20 and \$0.14 in Cases 1 and 2, respectively (see Table 31).

**Table 31** Option Value of the MLs using the Black-Scholes Method

Statistical Measure	Case 1	Case 2
Total Number of GPL Trips	11,721,575	31,234,266
Total Number of Option Users	214,859	1,561,879
Total Option Value	\$1,117,755	\$221,518
Average Option Value Per Person-Trip	\$0.095	\$0.007
Standard Deviation of Option Value Per Person-Trip	\$0.337	\$0.049
Average Option Value Per Person	\$5.202	\$0.142
Standard Deviation of Option Value Per Person	\$18.205	\$0.875

About 99 percent of the option values per person-trip were less than \$1.7 in both cases (see Table 32). However, a few large option value estimates were observed in Case



1, and the maximum option value per person-trip in Case 1 was \$15.02 (see Table 32). Thus, this research needed to identify why the large option values were estimated from the Black-Scholes method. The option value of the MLs partially depends on the standard deviation of the generalized trip costs on the GPLs. For the trip that had an option value of \$15.02, the standard deviation of the generalized trip costs on the GPLs in 2012 was \$2.74. This cost is equivalent to 11.7 minutes of variability (standard deviation) in travel times on GPLs using the value of travel time of \$14/hour in Case 1. Whereas, for the trips that had almost no option value per person-trip, the standard deviation of the generalized trip costs on the GPLs in 2012 was very small (e.g., \$0.02, which is equivalent to 0.1 minutes of the variability in travel times). These results coincide with the finding in the sample estimations using the Black-Scholes method in Section 5.6: the higher the standard deviation of the generalized trip costs on the GPLs in 2012, the higher the option value of MLs. Specifically, from the option value estimates using the Black-Scholes method, the correlation coefficient between the option value per person-trip and the standard deviation of the generalized trip costs on the GPLs in 2012 was estimated as 0.58. Thus, those unusually large option value estimates could be caused by very high variability in the generalized trip cost (travel time) on the GPLs because the high variability increases the chances that the MLs may be needed (valued) by GPL travelers.

About 99% of the option users valued the option of having the MLs less than \$72 for the year in both cases. The unusually large option value per person (the maximum option value per person in Case 1 in Table 32) is due to the large number of GPL trips in

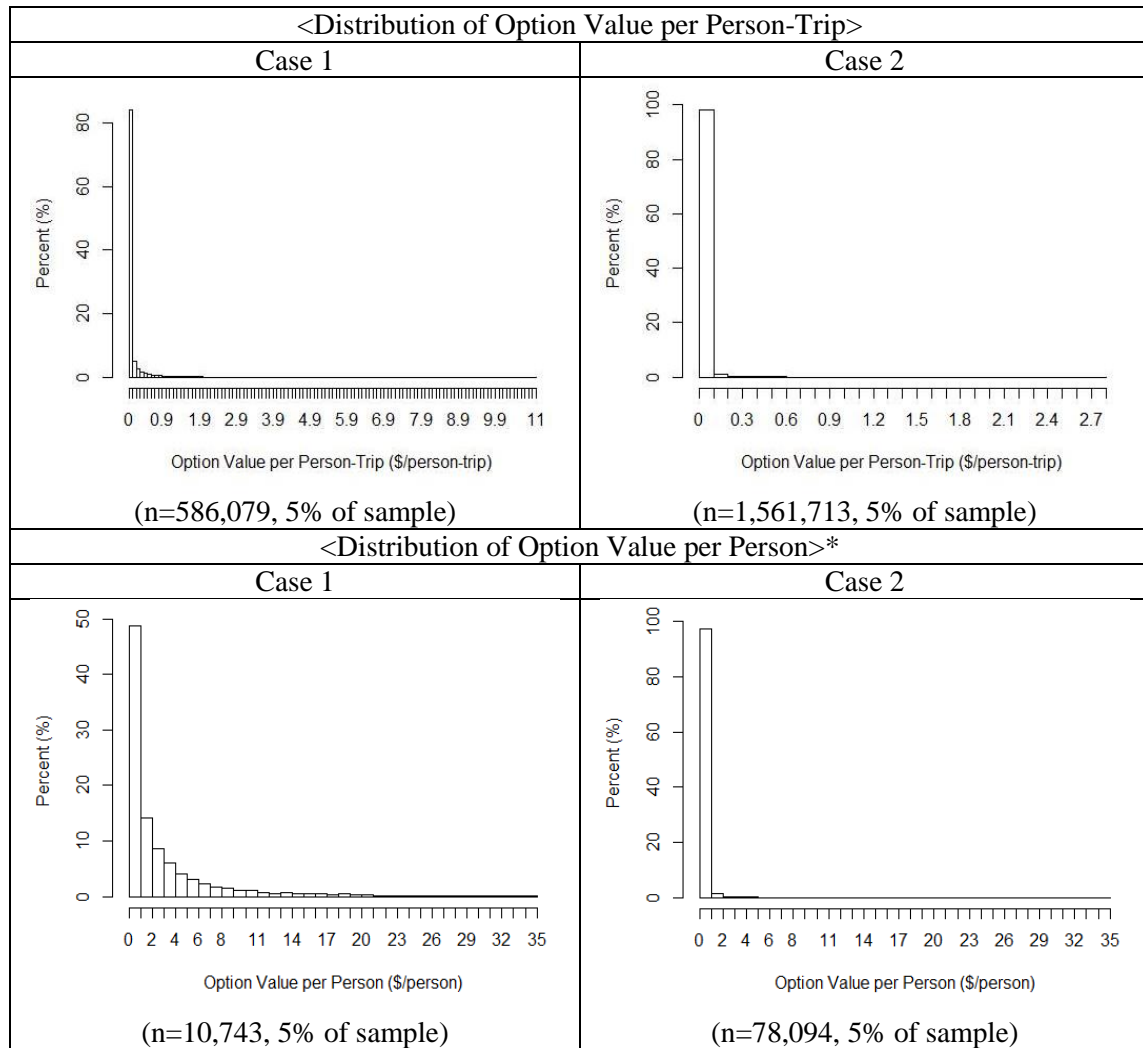
2012, i.e., 434 trips in 2012. In addition, the traveler who had the maximum option value per person in Case 1 frequently traveled on the GPL section where the generalized trip cost is unreliable. Specifically, s/he traveled 86 times (out of 434 times) on the section where the standard deviation of the generalized trip costs on the GPLs was greater than \$2 (which is equivalent to 8.6 minutes of the variability in travel times).

**Table 32** Quantiles of the Option Value of the MLs Using the Black-Scholes Method

Quantiles	Case 1		Case 2	
	Per Person-Trip	Per Person	Per Person-Trip	Per Person
100% Max	\$15.019	\$1165.799	\$3.514	\$204.477
99%	\$1.645	\$71.810	\$0.203	\$2.178
95%	\$0.536	\$20.300	\$0.016	\$0.622
90%	\$0.231	\$10.715	\$0.002	\$0.292
75% Q3	\$0.025	\$3.795	\$0.000	\$0.038
50% Median	\$0.000	\$1.100	\$0.000	\$0.001
25% Q1	\$0.000	\$0.182	\$0.000	\$0.000
10%	\$0.000	\$0.000	\$0.000	\$0.000
5%	\$0.000	\$0.000	\$0.000	\$0.000
1%	\$0.000	\$0.000	\$0.000	\$0.000
0% Min	\$0.000	\$0.000	\$0.000	\$0.000

Figure 17 provides the distributions of the option value estimates using the Black-Scholes method. As in the previous section, this research randomly selected five percent of the option value estimates to draw the distributions. As shown in Figure 17, most option value estimates for each trip were less than \$0.9 in Case 1 and most option value estimates for each trip were less than \$0.2 in Case 2. Most option value estimates for each option user were less than \$35 per year in Case 1 and most option value

estimates for each option user were less than \$2 per year in Case 2. The next section compares the option values of the MLs in 2012 using both methods.



\*: The option value estimate samples that were greater than \$35 per year were excluded in the distributions. Those were 2.4% and 0.01% of the samples in Cases 1 and 2, respectively.

**Figure 17** Distribution of the Option Value Using the Black-Scholes Method

### 6.3 Comparison of the Option Values of the MLs using the Log Sum and the Black-Scholes Methods

This section compares the option values of the MLs in 2012, which were estimated using the log sum method and the Black-Scholes method in Sections 6.1 and 6.2. Furthermore, in Section 5.7, this research identified that the log sum method may provide poor estimates of the option value of the MLs in the case where the GPLs provide as reliable generalized trip costs as the MLs do and the generalized trip cost on the GPLs is much less than the generalized trip cost on the MLs (as in the trips of Traveler C in Table 28). Therefore, those estimates were also examined in this section.

*First, the log sum method estimated larger total option values than the Black-Scholes method in both cases (see Table 33).* In Case 1, the total option value estimated using the log sum method was about four times the total option value estimated using the Black-Scholes method. In Case 2, the total option value estimated using the log sum method was about 13 times the total option value estimated using the Black-Scholes method.

*Second, in the log sum method, the ratio of the Case 1 total option value to the Case 2 total option value was approximately 1.5, whereas, in the Black-Scholes method, the ratio of the Case 1 total option value to the Case 2 total option value was approximately 5.0.* Thus, the option value in the Black-Scholes method was more affected by the value of the travel time estimate than that in the log sum method. As mentioned, in Case 2, the option users who never used MLs were added to those Case 1 option users who used the MLs at least once. The main difference between Cases 1 and 2

is the value of travel time estimate used to estimate the option value in each case (i.e., \$14/hour in Case 1 and \$5.97/hour in Case 2). The sensitivity analysis in Section 6.5.1 considers the effect of the value travel time estimate on the option value in both methods.

**Table 33** Comparison of the Option Values of the MLs in 2012 in Both Methods

Value (\$)	The Log Sum Method		The Black-Scholes Method	
	OV in Case 1	OV in Case 2	OV in Case 1	OV in Case 2
Total in 2012	4,576,625	2,975,032	1,117,755	221,518
Average Per Person-Trip	0.390	0.095	0.095	0.007
Average Per Person	21.301	1.905	5.202	0.142

Note: The OV represents the option value of the MLs.

**Third**, if a traveler travels on the GPL section where the generalized trip cost on the GPLs is much less than that on the MLs and the GPLs provide as reliable trip costs as the MLs do, s/he rarely gains benefit from using the MLs. In this case, the ML option is not valuable to the traveler and therefore the option value of zero could be reasonable for the traveler. However, in the sample estimations in Section 5.7, this research identified that the log sum method provided estimates that were not close to zero in those cases. *It was found that this occurs much more frequently in the log sum method than in the Black-Scholes method (see Table 34).* In order to examine the frequency of these poor estimates, this research used the statistics of the normal distribution in Equation 22. For the normal distribution (Montgomery 2007):

$$P(\mu_Y - 4\sigma_Y < Y < \mu_Y + 4\sigma_Y) = 0.9999 \quad (22)$$

where:

P = probability,

Y = normally distributed random variable,

$\mu_Y$  = mean of Y,

$\sigma_Y$  = standard deviation of Y.

From the statistics of the normal distribution,  $P(Y > \mu_Y + 4\sigma_Y) \leq 0.0001$ . In Section 5.5, this research identified that the distribution of the generalized trip costs on the GPLs is approximately log-normal. Thus, based on the statistics of the normal distribution, it is reasonable that:

$$P(\ln(S) > \mu_{\ln(S)} + 4\sigma_{\ln(S)}) \leq 0.0001 \quad (23)$$

where:

S = the generalized trip cost on the GPLs.

Thus, if the natural log of the generalized trip cost on the MLs ( $\ln(X)$ ) is greater than  $[\mu_{\ln(S)} + 4\sigma_{\ln(S)}]$ , it is also reasonable that:  $P(\ln(S) > \ln(X)) \leq 0.0001$ . This probability indicates that there is almost no chance that (the natural log of) the generalized trip cost on the GPLs is greater than that on the MLs. In this case, since the generalized trip cost on the MLs is almost always greater than that on the GPLs, the traveler could not gain benefit from using the MLs. Thus, the option value of zero could

be reasonable. Conversely, if the estimate of the option value was not close to zero in that case, the estimate of the option value could be the poor estimate. Based on this concept, this research established criteria to determine whether the estimate of the option value was the poor estimate that was not close to zero or not in those cases. These criteria are: 1)  $\ln(X) > \mu_{\ln(S)} + 4\sigma_{\ln(S)}$  and 2) the option value of the MLs  $\neq 0$ . These criteria determine whether the estimate of the option value is zero or not in the case where the generalized trip cost on the MLs is almost always greater than that on the GPLs.

The average of the natural logarithm generalized trip costs on the GPLs ( $\mu_{\ln(S)}$ ) was estimated from all users who traveled the same GPL section at the same 10-minute interval of every day in 2012 as the option user did. However, to estimate the average of the natural logarithm generalized trip costs on the GPLs, weekend and weekday trip costs on the GPLs were separated because, in general, weekday trips show different trip characteristics from weekend trips. The standard deviation of the natural logarithm generalized trip costs on the GPLs ( $\sigma_{\ln(S)}$ ) was also estimated in the same way as the average of the natural logarithm generalized trip costs on the GPLs ( $\mu_{\ln(S)}$ ). In Case 1, 21.5 percent of the total option value estimates using the log sum method were the poor estimate that was not close to zero in the case where the ML option was not valuable to the option user, whereas 1.7 percent of the total option value estimates using the Black-Scholes method were the poor estimate. In Case 2, 69.3 percent of the total option value estimates using the log sum method were the poor estimate, whereas 5.2 percent of the total option value estimates using the Black-Scholes method were the poor estimate (see

Table 34). Based on these results, the log sum method appears to provide the poor estimates of the option value of the MLs in the case where the generalized trip cost on the GPLs is much less than that on the MLs and the GPLs provide as reliable trip costs as the MLs do.

**Table 34** The Number of the Poor Estimates of the Option Value in Both Methods

Method	Case 1		Case 2	
	The Number of the Poor Estimates of OV	Percentage	The Number of the Poor Estimates of OV	Percentage
The Log Sum	2,515,746	21.5%	21,647,717	69.3%
The Black-Scholes	194,074	1.7%	1,611,459	5.2%
Total Number of the OV Estimates in Each Case	11,721,575		31,234,266	

Note: The OV represents the option value of the MLs.

#### 6.4 Comparison of the Option Values for the Option Users in Case 2

Case 2 includes the users who never used the MLs in 2012 (Category a in Table 18 in Section 5.1) and the users who used the MLs at least once in 2012 (Category b in Table 18 in Section 5.1). This section compares the option values for these option users. In the option values estimated from the log sum method, the average option value per person-trip for the option users in Category a was greater than that for the option users in Category b — 10 cents and 8.8 cents in Categories a and b, respectively. However, the average total yearly option value for each option user in Category a was less than that for



each option user in Category b. These results were also similar in the Black-Scholes method (see Table 35). Thus, the option users who occasionally used MLs in Category b will value the option of having the MLs available more than the option users who never used MLs in Category a. This is because option users in Category b traveled more frequently on the freeway than option users in Category a. In 2012, the average GPL trip frequency for Category a option users was 15 times (=19,512,691 GPL trips/1,347,020 users) and the average GPL trip frequency for Category b option users was 55 times (=11,721,575 GPL trips/214,859 users). Thus, it could be reasonable that travelers who frequently travel on the freeway will pay more for the managed lane option than travelers who rarely travel on the freeway.

**Table 35** Comparison of the Option Values of the Option Users in Categories a and b

		The Option Users in Category a*	The Option Users In Category b**	Total (a+b)
Number of the GPL Trips		19,512,691	11,721,575	31,234,266
Number of the Users		1,347,020	214,859	1,561,879
The Log Sum Method	Average OV per Person-Trip	\$0.100	\$0.088	\$0.095
	Average OV per Person	\$1.443	\$4.797	\$1.905
	Total OV in 2012	\$1,944,291	\$1,030,741	\$2,975,032
The Black-Scholes Method	Average OV per Person-Trip	\$0.009	\$0.004	0.007
	Average OV per Person	\$0.130	\$0.219	0.142
	Total OV in 2012	\$174,468	\$47,050	221,518

Note: OV represents option value of MLs.

\*: Category a includes users who never used the MLs in 2012.

\*\*: Category b includes users who occasionally used the MLs in 2012.

## 6.5 Sensitivity Analysis

Because of uncertainty in model inputs, the exact model output is unknown. Sensitivity analysis helps identify the degree to which model outputs are affected by changes in inputs and can provide a range of possible outcomes rather than a single number (Kockelman et al. 2013). The option value estimates from both methods can also be affected by parameter inputs, such as the value of travel time estimate. Thus, it is important to identify the degree to which the estimated option values are affected by changes in those parameter inputs and a range of possible option value estimates in the methods. This section presents a sensitivity analysis with respect to four parameter inputs: 1) the value of travel time, 2) the standard deviation of the generalized trip costs on the GPLs, 3) the time period over which the standard deviation of the generalized trip costs on the GPLs is calculated, and 4) the risk-free return. The sensitivity analysis for the value of travel time was conducted in both methods. The sensitivity analyses for the other parameters were conducted in the Black-Scholes method because they only applied to the Black-Scholes method. Note that the sensitivity analysis in this section was conducted only for Case 1 because it aims to identify the effect of the parameter inputs on the option value of the MLs and a range of possible option value estimates, not the exact option value of the MLs. In addition, Case 1 that included only the option users who used the MLs at least once in 2012 was primarily considered in this research. This is because the option users who never used the MLs in 2012 (included only in Case 2) could continue to exclusively use the GPLs in the future regardless of traffic situations and may not value the ML option.

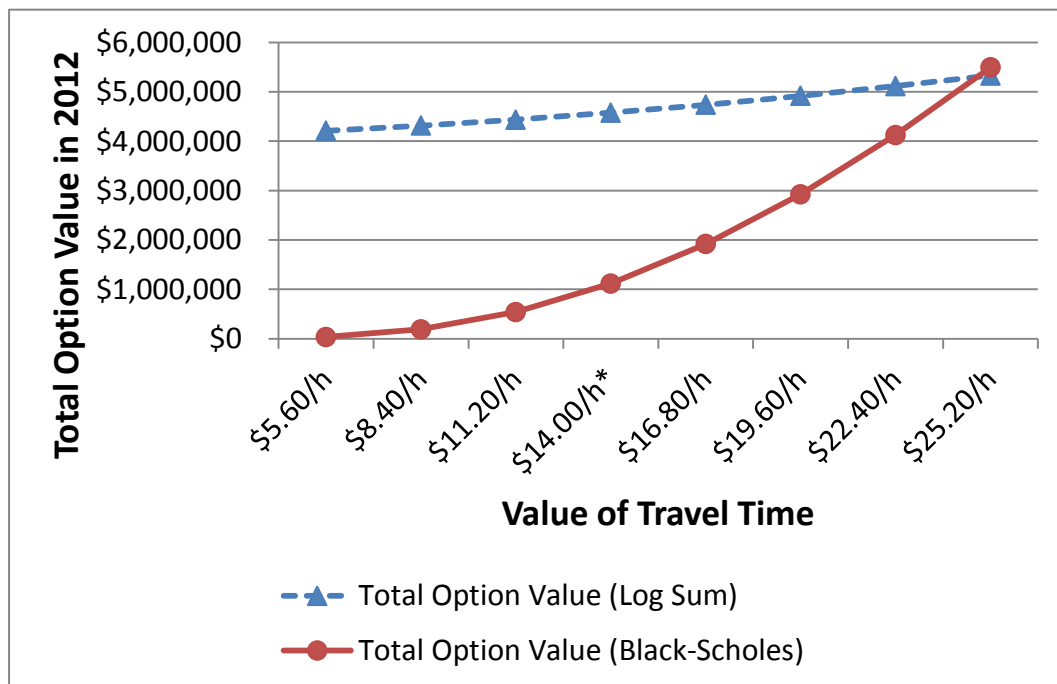
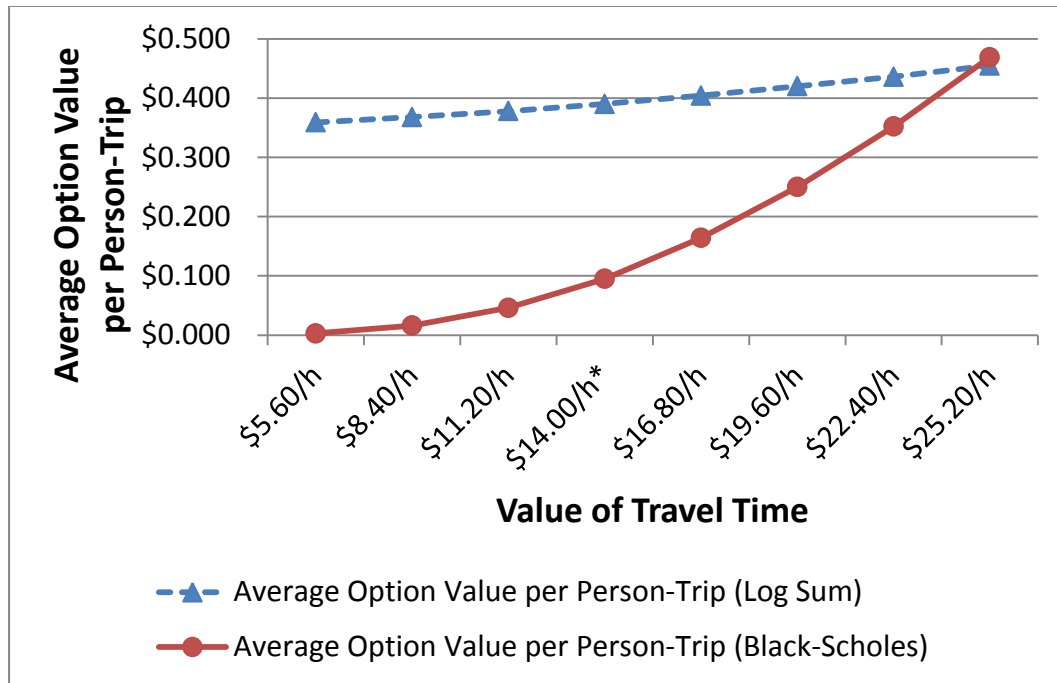
#### *6.5.1 Effect of the Value of Travel Time on the Option Value of the MLs*

While comparing Cases 1 and 2 in Section 6.3, this research recognized that the option value of the MLs in 2012 estimated using the Black-Scholes method largely decreased (from about 1.1 million dollars in Case 1 to 0.2 million dollars in Case 2) due to the change in the value of travel time parameter (from \$14/hour to \$5.97/hour) even though Case 2 also included the option users who never used MLs in 2012 with the option users who occasionally used MLs in 2012 in Case 1. Thus, this section identifies the effect of the value of travel time parameter on the option value of the MLs in Case 1. To identify the effect, the initial value of travel time was both increased and decreased by 20, 40, 60 and 80 percent, respectively. Note that, when the initial value of travel time was decreased by 60 percent, the total option value using the Black-Scholes method was very small (\$35,756). In this case, the initial value of travel time was not decreased by 80 percent. The other parameters in each method were unchanged. The total option values in 2012 estimated using the log sum method ranged from \$4,212,720 to \$5,329,086, while the total option values in 2012 estimated using the Black-Scholes method ranged from \$35,756 to \$5,493,631 (see Table 36). The option values estimated using the Black-Scholes method were greatly affected by the value of travel time. In the Black-Scholes method, the total option value using the value of travel time of \$25.20/hour was about 154 times the total option value using the value of travel time of \$5.60/hour. While in the log sum method, the total option value estimate applying the value of travel time of \$25.20/hour was about 1.3 times the total option value estimate applying the value of travel time of \$5.60/hour (see Figure 18).

**Table 36** Option Value of the MLs depending on the Value of Travel Time

Value of Travel Time (VOT)	Method	Average OV per Person-Trip	Average OV per Person	Total OV in 2012
\$5.60/hour (60% Decrease in VOT)	Log Sum	\$0.359	\$19.607	\$4,212,720
	Black-Scholes	\$0.003	\$0.166	\$35,756
\$8.40/hour (40% Decrease in VOT)	Log Sum	\$0.368	\$20.082	\$4,314,807
	Black-Scholes	\$0.016	\$0.884	\$189,940
\$11.20/hour (20% Decrease in VOT)	Log Sum	\$0.378	\$20.645	\$4,435,805
	Black-Scholes	\$0.046	\$2.519	\$541,320
\$14.00/hour (No Change in VOT)	Log Sum	\$0.390	\$21.301	\$4,576,625
	Black-Scholes	\$0.095	\$5.202	\$1,117,755
\$16.80/hour (20% Increase in VOT)	Log Sum	\$0.404	\$22.049	\$4,737,341
	Black-Scholes	\$0.164	\$8.923	\$1,917,184
\$19.60/hour (40% Increase in VOT)	Log Sum	\$0.420	\$22.886	\$4,917,226
	Black-Scholes	\$0.250	\$13.615	\$2,925,263
\$22.40/hour (60% Increase in VOT)	Log Sum	\$0.436	\$23.806	\$5,114,993
	Black-Scholes	\$0.352	\$19.192	\$4,123,511
\$25.20/hour (80% Increase in VOT)	Log Sum	\$0.455	\$24.803	\$5,329,086
	Black-Scholes	\$0.469	\$25.569	\$5,493,631

Note: The OV represents the option value of the MLs.



\*: No change in the value of travel time (\$14.00/hour).

**Figure 18** Option value of the MLs depending on the value of travel time

### *6.5.2 Effect of Standard Deviation of Generalized GPL Trip Costs on the Option Value of the MLs*

In the sample estimations using the Black-Scholes method in Section 5.6, this research identified that the option value using the Black-Scholes method depends on the standard deviation of the generalized trip costs on the GPLs. This standard deviation of the generalized trip costs on the GPLs is directly affected by the standard deviation (variability) of travel times on the GPLs because the generalized trip cost on the GPLs is estimated by multiplying the travel time on the GPLs and the value of travel time parameter. The standard deviation of travel times changes as traffic conditions on the GPLs change. Thus, this section identifies the effect of the standard deviation of the generalized trip costs on the GPLs on the option value of the MLs in Case 1 in the Black-Scholes method. In order to identify the effect, the initial standard deviation of the generalized trip costs on the GPLs for each trip increases (or decreases) by 20, 40, and 60 percent, respectively. These changes are well within the range observed on the freeway. For example, on the eastbound section between the sensors 427 and 426, the standard deviation of the generalized trip costs on the GPLs could be increased by up to 198% or decreased by up to 96% during the same hour when the trip was made, whereas, on the westbound section between the sensors 425 and 396, the standard deviation of the generalized trip costs on the GPLs could be increased by up to 53% or decreased by up to 41% during the same hour when the trip was made. The other parameters in the Black-Scholes method are unchanged. The total option values in 2012 estimated using the Black-Scholes method ranged from \$385,097 to \$2,075,960 (see Table 37). Due to

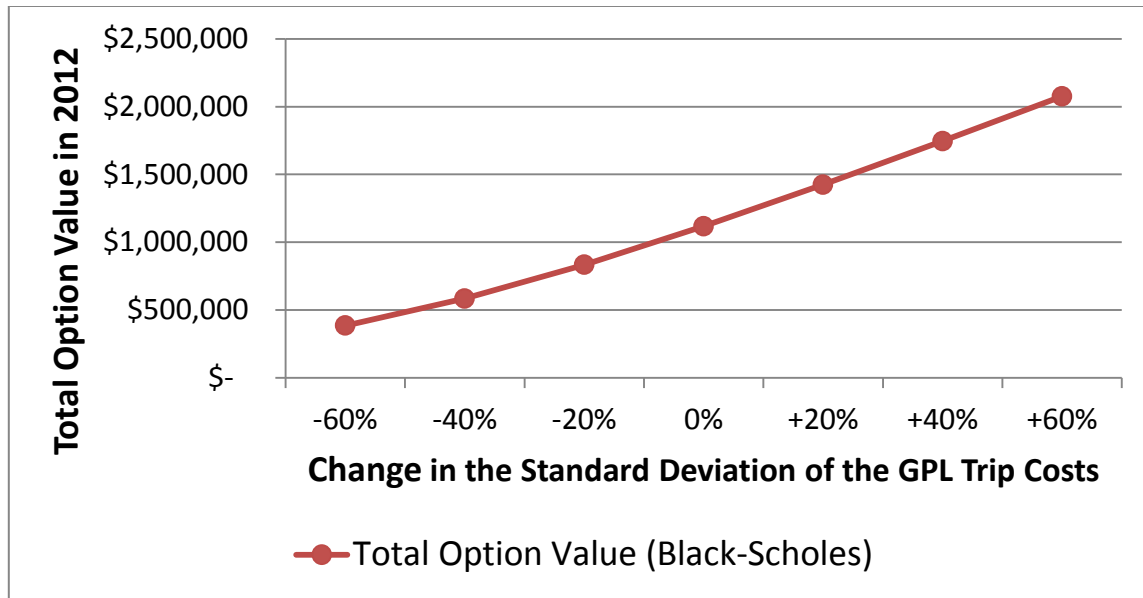
this 60% increase in the standard deviation of the generalized trip costs on the GPLs, the total option value increased by about 86%; furthermore, the 60% decrease in standard deviation of the generalized trip costs on the GPLs caused the total option value to decrease by about 66%.

**Table 37** Option Value of the MLs Depending on the Standard Deviation of the Generalized Trip Costs on the GPLs Using the Black-Scholes Method

The Standard Deviation of the Generalized Trips Costs on the GPLs (SD)	Average OV per Person-Trip	Average OV per Person	Total OV in 2012	Percent Change from the initial OV
60% Decrease in the SD	\$0.033	\$1.792	\$385,097	-65.5%
40% Decrease in the SD	\$0.050	\$2.720	\$584,432	-47.7%
20% Decrease in the SD	\$0.071	\$3.881	\$833,877	-25.4%
No Change in the SD	\$0.095	\$5.202	\$1,117,755	0.0%
20% Increase in the SD	\$0.122	\$6.630	\$1,424,622	27.5%
40% Increase in the SD	\$0.149	\$8.127	\$1,746,079	56.2%
60% Increase in the SD	\$0.177	\$9.662	\$2,075,960	85.7%

Note: The OV represents the option value of the MLs.

Figure 19 shows consistent increases in the total option values of MLs in 2012 as the standard deviation of the generalized trip costs on the GPLs increased.



**Figure 19** Total option value of the MLs depending on change in the standard deviation of the generalized trip costs on the GPLs using the Black-Scholes Method

### 6.5.3 Comparison of Applying Short-Term and Long-Term Standard Deviations of the Generalized Trip Costs on the GPLs to the Black-Scholes Method

As mentioned in Section 5.5, the standard deviation of the generalized trip costs on the GPLs in the Black-Scholes method implies the standard deviation of the generalized trip costs on the GPLs which travelers on the GPLs will encounter in the future. This standard deviation of the generalized trip costs on the GPLs is clearly unknown because it will occur in the future. Thus, this research used the historical standard deviation of the generalized trip costs on the GPLs in 2012 instead. However, the historical standard deviation could depend on the historical period over which the standard deviation of the generalized trip costs on the GPLs is to be calculated.



In the sample estimations using the Black-Scholes method in Section 5.6, Trip 1 of Traveler A in Table 27 presented quite a different standard deviation of the generalized trip costs on the GPLs during the month when the trip was made ( $\sigma_m$ ) from the standard deviation of the generalized trip costs on the GPLs in all of 2012 ( $\sigma$ ) — \$1.015 and \$0.764 for  $\sigma_m$  and  $\sigma$ , respectively. As a result, for Trip 1 of Traveler A, the option value estimate using the standard deviation of the generalized trip costs on the GPLs during one month ( $\sigma_m$ ) was \$0.874, whereas the option value estimate using the standard deviation of the generalized trip costs on the GPLs for all of 2012 ( $\sigma$ ) was \$0.591 (see Table 27 in Section 5.6 for Trip 1 of Traveler A). However, in Trip 1 of Traveler B, the standard deviation of the generalized trip costs on the GPLs during one month ( $\sigma_m$ ) was similar to the standard deviation of the generalized trip costs on the GPLs for all of 2012 ( $\sigma$ ) — \$0.457 and \$0.455 for  $\sigma_m$  and  $\sigma$ , respectively. As a result, for Trip 1 of Traveler B, the option value estimate using the standard deviation of the generalized trip costs on the GPLs during one month ( $\sigma_m$ ) was similar to the option value estimate using the standard deviation of the generalized trip costs on the GPLs for all of 2012 ( $\sigma$ ) — \$0.560 and \$0.555, respectively (see Table 27 in Section 5.6 for Trip 1 of Traveler B). Thus, this section compares the option value of the MLs using the standard deviation of the generalized trip costs on the GPLs during the month when the trip was made (the short-term standard deviation) with the option value of the MLs using the standard deviation of the generalized trip costs on the GPLs in 2012 (the long-term standard deviation) for all option users in Case 1. For the short-term standard deviation ( $\sigma_m$ ), the standard deviation of the generalized GPL trip costs of all users was calculated

for travelers using same GPL section at the same 10-minute interval every day during the month when the trip was made by the option user. However, as the standard deviation of the generalized trip costs on the GPLs in 2012 ( $\sigma$ ), weekend and weekday trip costs on the GPLs are separated to estimate the standard deviation of the generalized trip costs on the GPLs during one month ( $\sigma_m$ ). Applying the short-term standard deviation to the Black-Scholes method without changes in the other parameters decreased the total option value of the MLs in 2012 by \$93,793, which was about 8.4 percent of the initial total option value (see Table 38). Based on these results, the time period over which the standard deviation of the generalized trip costs on the GPLs is to be calculated had a small impact on the option value of the MLs using the Black-Scholes method.

**Table 38** Comparison of Applying the Short-Term and the Long-Term Standard Deviations to the Option Value Estimation in the Black-Scholes Method

The Standard Deviation of the Generalized Trip Costs on the GPLs (SD)	Average OV per Person-Trip	Average OV per Person	Total OV in 2012
Applying the Standard Deviation of the Generalized Trip Costs on the GPLs in 2012 (the Long-Term Standard Deviation) – No Change	\$0.095	\$5.202	\$1,117,755
Applying the Standard Deviation of the Generalized Trip Costs on the GPLs during the Month when the Trip Was Made (the Short-Term Standard Deviation)	\$0.087	\$4.766	\$1,023,962

Note: The OV represents the option value of the MLs.

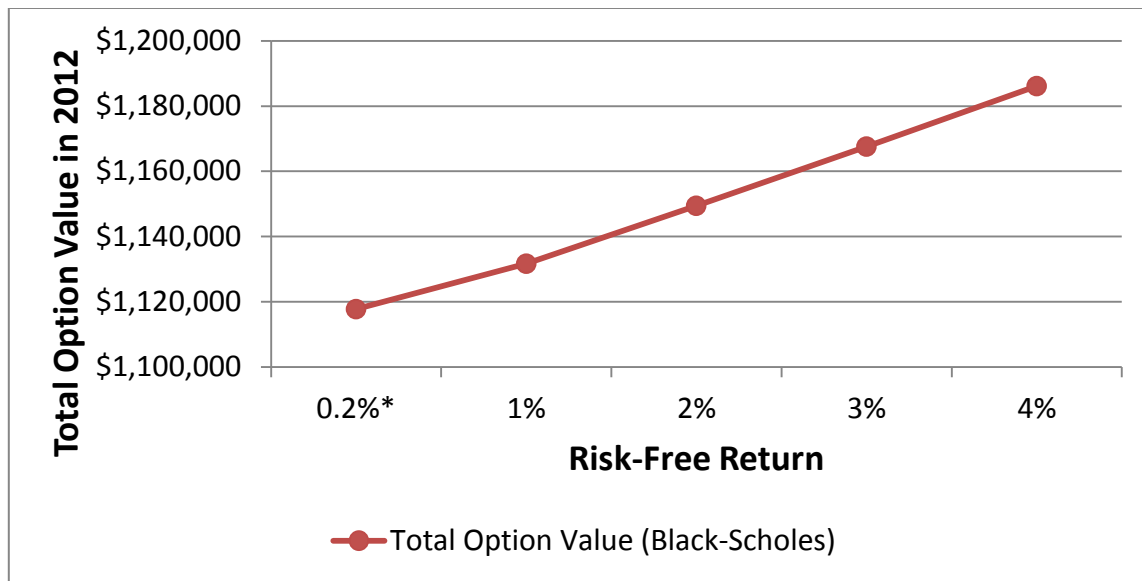
#### 6.5.4 Effect of the Risk-Free Return on the Option Value of the MLs

The risk-free return parameter in the Black-Scholes method was assumed as 0.2% based on the interest rate of a 12-month U.S. Treasury bill in 2012 (U.S. Department of the Treasury 2015). The interest rate could vary depending on economic situations. Thus, this section identifies the effect of the risk-free return on the option value of the MLs for the option users in Case 1 in the Black-Scholes method. In order to identify the effect, the risk-free returns of 1, 2, 3 and 4 percent are applied, respectively. But the other parameters in the Black-Scholes method are unchanged. Applying the risk-free return of 4 percent increased the total option value in 2012 by \$68,347, which was about 6.1 percent of the initial total option value applying the risk-free return of 0.2% (see Table 39). The increase in the risk-free return resulted in a slight increase in the total option value in 2012 as shown in Figure 20. From these results, the risk-free return parameter had a minimal impact on the option value of the MLs when using the Black-Scholes method.

**Table 39** Option Value of the MLs Depending on the Risk-free Return in the Black-Scholes Method

Risk-Free Return	Average OV per Person-Trip	Average OV per Person	Total OV in 2012
0.2% (No Change)	\$0.095	\$5.202	\$1,117,755
1%	\$0.097	\$5.267	\$1,131,679
2%	\$0.098	\$5.350	\$1,149,429
3%	\$0.100	\$5.434	\$1,167,568
4%	\$0.101	\$5.520	\$1,186,102

Note: The OV represents the option value of the MLs.



\*: No change in the risk-free return (0.2%).

**Figure 20** Total option values of the MLs depending on the risk-free return in the Black-Scholes Method

## 6.6 Comparison of the Option Values with the Value of Travel Time Savings from the MLs

This section compares the option value of the MLs with the value of travel time savings for Case 1 option users who used the MLs at least once in 2012. As mentioned, the value of travel time savings refers to the benefits from reduced travel time by using MLs. This comparison provides the relative amount of the option use benefit of the MLs compared to the actual use benefit of the MLs, which is the value of travel time savings. If the option use benefit is very small compared to the actual use benefit of the MLs, the option value estimation in this research only adds complexity to the estimation of ML benefits but has minimal impact on the results. Conversely, if the option value is larger than the value of travel time savings, this would cast doubt on the validity of this option

value analysis as the option value would logically be smaller than the largest actual benefit: the travel time savings.

The number of Case 1 option users who occasionally used MLs in 2012 was 214,859, and they made 1,781,919 ML trips (choices) in 2012. From these ML trips, the value of travel time savings of the MLs was estimated for the option users in Case 1. As estimated in Table 20, the value of travel time of \$14/hour for Case 1 option users was used to estimate the value of travel time savings. The value of travel time savings is calculated by Equation 24:

$$VTTS = \sum^{\text{All ML Trips}} VOT \times (TT \text{ on the GPLs} - TT \text{ on the MLs}) \quad (24)$$

where:

$VTTS$  = the value of travel time savings,

$VOT$  = the value of travel time, which is \$14/hour for the option users in Case 1,

$TT$  = travel time (in hours).

The value of travel time savings for the option users in Case 1 in 2012 was estimated to be \$1,096,258 from 1,781,919 ML trips. The estimate of the value of travel time savings was about one fourth of the total option value using the log sum method and was similar to the total option value using the Black-Scholes method (see Table 40). Based on these results, ML option use benefits were greater than or similar to the value of travel time savings from the MLs for the option users in Case1.

**Table 40** Comparison of the Option Values with the Value of Travel Time Savings from the MLs for the Option Users in Case 1

Value	Average per Person-Trip	Average per Person	Total in 2012	Number of Trips used for Estimation of the Value
The Option Value using the Log Sum Method	\$0.390	\$21.301	\$4,576,625	11,721,575 GPL Trips
The Option Value using the Black-Scholes Method	\$0.095	\$5.202	\$1,117,755	11,721,575 GPL Trips
The Value of Travel Time Savings	\$0.615	\$5.102	\$1,096,258	1,781,919 ML Trips

### 6.7 Summary of the Option Value of the MLs on the Katy Freeway

In Case 1, the total option values of \$4,576,625 and \$1,117,755 were estimated using the log sum method and the Black-Scholes method, respectively. In Case 2, the total option values of \$2,975,032 and \$221,518 were estimated using the log sum method and the Black-Scholes method, respectively (see Table 41). However, this research identified that the log sum method frequently provided a poor estimate of the option value in the case where the generalized trip cost on the GPLs is much less than that on the MLs and the GPLs provide as reliable trip costs as the MLs do. Since, in this case, ML option users could not benefit from using the MLs, the option value of zero may be reasonable. However, in Case 1, 21.5 percent of the total estimates from the log sum method were not close to zero in those cases, whereas 1.7 percent of the total estimates from the Black-Scholes method were not close to zero in those cases. These results were amplified in Case 2 (see Table 41). In addition, in Section 6.2, this research found that the option value of the MLs using the Black-Scholes method partially depended on the volatility (the standard deviation) in the GPL trip cost: the higher the

volatility in the GPL trip cost, the higher the option value of MLs. This volatility directly considers when the GPL trip cost unexpectedly increases. The main benefit of the ML option is the reduced trip cost by using the MLs when the GPL trip cost unexpectedly increases. However, the log sum method does not consider the volatility in the GPL trip cost to determine the option value of the MLs. Lastly, in Section 6.5.1, this research identified that the option value using the log sum method was not sensitive to the value of travel time parameter, whereas the option value using Black-Scholes method increases as the value of travel time parameter increases. In general, travelers who have high values of travel time are likely to pay more to save their travel time than travelers who have low values of travel time. Thus, it is reasonable that an increase in the value of travel time should increase the option value of the MLs. However, in the log sum method, as the value of travel time parameter increased by 350%, the option value increased by only 26%. Thus, this research recommends the Black-Scholes method to estimate the option value of MLs.

**Table 41** Summary of the Results

Summary of the Results	Case 1	Case 2
Option Users*	b)	a) + b)
Total Number of Option Users	214,859	1,561,879
Total Number of GPL Trips	11,721,575	31,234,266
Total Option Value using the Log Sum Method	\$4,576,625	\$2,975,032
Total Option Value using the Black-Scholes Method	\$1,117,755	\$221,518
The Number of the Poor Estimates of Option Value using the Log Sum Method (Percentage)	2,515,746 (21.5%)	21,647,717 (69.3%)
The Number of the Poor Estimates of Option Value using the Black-Scholes Method (Percentage)	194,074 (1.7%)	1,611,459 (5.2%)
Total Value of Travel Time Savings**	\$1,096,258	NA

\*: a) indicates the users who never used the MLs in 2012, and b) indicates the users who occasionally used the MLs in 2012.

\*\*: The total value of travel time savings was estimated from 1,781,919 ML trips of the option users in Case 1.

This research examined two potential definitions of option users who valued the MLs as a travel option. These definitions included Case 1) the option users who used the MLs at least once in 2012 and Case 2) the option users who occasionally or never used the MLs in 2012. However, the fact that a traveler never used the MLs in 2012 would likely indicate that the traveler will not use the MLs in the future. In order to identify whether this inference is appropriate or not, this research traced travel records until September 30, 2014 of the users who never used the MLs in 2012. 93.3 percent of the users who never used the MLs in 2012 did not use the MLs in 2013 and 90.3 percent of them did not use the MLs in all of the data (see Table 42). In addition, they made a total of 18,736,229 trips (including both ML and GPL trips) and 1.6 percent of these trips were the ML trip in 2013. Also, they made a total of 10,645,575 trips (including both ML and GPL trips) and 2.8 percent of these trips were the ML trip in 2014 (from January to September). Whereas, ML trip percentages of all travelers with a transponder



were 6.7 percent and 8.1 percent in 2013 and 2014, respectively. Based on these results, the users who never used the MLs in 2012 are those users who very rarely use the MLs. Thus, this research recommends ML option users to be defined as only those users who occasionally used MLs and the total option value of \$1,117,755 using the Black-Scholes method in Case 1 for the option value of the MLs in 2012. Note that this research excluded travelers who did not have a transponder in their vehicle from the ML option users. This is because, in order to use the MLs on the freeway as a toll paying SOV, vehicles have to have a transponder and this indicates that travelers who did not install the transponder had no intention of using the MLs as a toll paying SOV, even occasionally. However, there is a chance that some of those travelers who did not install the transponder could value the ML option. Therefore, the total option value of the MLs could increase due to the travelers excluded in this research. However, we expect very little increase in the total option value because the total option value was already estimated using 1,993,347 different transponders in the data set. It is likely that the vast majority of travelers who wanted to use the MLs already had a transponder.

**Table 42** Number of the Option Users who never used the MLs in 2012 by their ML Trip Frequency until 2013 and 2014

ML Trip Frequency	until Dec 31, 2013		until Sep 30, 2014	
	Number of the option users*	Percentage	Number of the option users*	Percentage
0	1,257,412	93.3%	1,215,844	90.3%
1	52,705	3.9%	67,523	5.0%
2	15,415	1.1%	23,136	1.7%
3	6,662	0.5%	11,120	0.8%
4	3,628	0.3%	6,556	0.5%
5	2,251	0.2%	4,254	0.3%
5+	8,947	0.7%	18,587	1.4%
Total	1,347,020	100%	1,347,020	100%

\*: The option users indicate those users who never used the MLs in 2012.

From the total option value of \$1,117,755, the average option value per person-trip was estimated to be \$0.095 ( $=\$1,117,755/11,721,575$  trips). The average generalized trip cost on the GPLs for each trip of the ML option users in Case 1 was \$1.921. Thus, the average option value per person-trip was 5.0 percent of the average generalized trip cost on the GPLs.

In practice, the option value of MLs is not directly charged to ML option users. However, the purchase of a transponder (toll tag) by the ML option users can be considered as a part of the willingness to pay for having the ML option. This is because some may have purchased the transponder to occasionally use the MLs. However, many likely purchased the transponder to use other toll roads in the Houston area considering only 7 percent of our sample used the Katy MLs. In order to activate the transponder, each option user needs to pay \$15 (“EZ TAG” 2014). Theoretically, the transponder can be used permanently. However, in practice, since the transponder is installed on the

windshield of a car, it is sometimes broken when it is removed from the windshield due to replacement of the windshield or purchase of a new car. Thus, if we consider the average length of car ownership as the average duration of the transponder, the average yearly cost of the transponder can be estimated. In 2012, the average length of car ownership was 4.9 years (Statista, Inc. 2015). Thus, the average yearly cost of the transponder is estimated as \$3.09 ( $=\$15/4.9$  years). The average option value using the Black-Scholes method for each option user in Case 1 was \$5.20 ( $=\$1,117,755/214,859$  users) and is greater than the average yearly cost of the transponder. Therefore, the option benefit was greater, on average, than the cost to obtain that benefit.

This research recommends the Black-Scholes method to estimate the option value of MLs. To apply the Black-Scholes method in Equation 19 in Section 5.5 to estimate the option value of MLs, this research suggests that:

- 1) the generalized trip costs on the GPLs and the MLs (the variables S and X) should be estimated using utility functions of GPL and ML choices for the option users who occasionally used MLs,
- 2) the variable T should equal 1/frequency with which MLs are used in a year) because it provides the average time period over which the ML option user expects to take advantage of the ML option,
- 3) the risk-free return (the variable r) should be adopted from the interest rate on a U.S. Treasury bill at the time of analysis because, in practice, the interest rate on a U.S. Treasury bill is commonly considered as the risk-free return in financial markets (“Risk-Free Return” 2015). Even though this research

adopted the interest rate from the financial market, the risk-free return had a minimal impact on the option value of the MLs. Applying the risk-free return of 4 percent increased the total option value by 6.1 percent of the initial total option value applying the risk-free return of 0.2%,

- 4) the standard deviation (volatility) of the generalized trip costs on the GPLs (the variable  $\sigma$ ) should be estimated from the historical standard deviation of the generalized trip costs on the GPLs for one year (the long-term volatility). This is because 42.4 percent of the option users who occasionally used the MLs used the MLs once in 2012. In addition, applying the standard deviation of the generalized trip costs on the GPLs during one month (the short-term volatility) instead of the long-term volatility had a small impact (8.4 percent decrease in the total option value in Case 1) on the option value of the MLs.

However, from the sensitivity analysis in Section 6.5, this research identified that the option value of the MLs using the Black-Scholes method is largely affected by the value of travel time parameter and the standard deviation of the generalized trip costs on the GPLs in 2012. Thus, when estimating the option value of MLs using the Black-Scholes method, those parameters should carefully be examined.

Last, this research compared the option value of the MLs with the value of travel time savings from using the MLs for the option users who occasionally used the MLs in 2012. The total option value of the MLs estimated using the Black-Scholes method in Case 1 was similar to the value of travel time savings from using the MLs for those option users (see Table 41). Thus, considering ML option use benefit in the benefit-cost

analysis (BCA) is important for accurate decision making when choosing whether or not to invest in MLs. The next section concludes this research and provides recommendations to estimate the option value of the MLs.

## 7. CONCLUSIONS

The option value of MLs refers to travelers' willingness to pay (WTP) for the option of having MLs available for possible use at some time in the future, even if the option may never be used. MLs could provide a choice opportunity as a backup (which can be interpreted as the option value) for travelers even when they do not use MLs. MLs offer added flexibility in travel decisions, such as time of departure. Additionally, travelers have the option to use MLs depending on their situation such as when they encounter unexpected congestion and are in a hurry. Despite these potential benefits of MLs, earlier studies of MLs have only considered the actual use benefits, such as travel time savings and vehicle operating cost savings. This may present a potential problem as the value of MLs may be underestimated.

This research estimated the option value of the MLs on the Katy Freeway (I-10) in Houston. After opening the Katy Freeway MLs in 2008, the performance of the Freeway, including both MLs and GPLs, has been continuously monitored by the Texas Department of Transportation (TxDOT) and the Harris County Toll Road Authority (HCTRA) using automatic vehicle identification (AVI) sensors. This monitoring has resulted in a massive quantity of detailed travel data on the freeway. The data gives a unique opportunity to recognize lane-choice decisions between MLs and GPLs of all travelers with a transponder. This research used these data for approximately one year (from January to November in 2012). Using these data, this research identified revealed preferences (RP) of the travelers on the Katy Freeway in 2012. These RP data might be

suitable to estimate the option value of MLs and provide more precise results as it provides actual choice results depending on traffic situations, not latent preferences or stated preferences in a survey. Since these data only provided actual choices of the travelers and their attributes, such as travel time and the amount of toll paid, this research generated alternative choices that the travelers could have chosen at the time of travel but did not. By combining the actual and alternative choices and their attributes, each traveler's revealed preference was identified.

This research examined two potential definitions of option users who valued the MLs as a travel option. These definitions included case 1) the travelers who used the MLs at least once in 2012 and case 2) travelers who occasionally or never used the MLs in 2012. However, this research identified that the option users who never used the MLs in 2012 extremely rarely used the MLs even in 2013 and 2014. 90.3 percent of the option users who never used the MLs in 2012 did not use the MLs in 2013 or in the first nine months of 2014 (the extent of the data available). Thus, this research recommends ML option users to be defined as only those users who used the MLs at least once in a year.

This research used a modified log sum method and a modified Black-Scholes method to estimate the option value of the MLs. The log sum method estimates the option value of the MLs by measuring change in consumer surplus between the situation where both MLs and GPLs are available and the situation where only the GPLs are available. The Black-Scholes method considers how much option users can reduce their trip cost by using the MLs when the GPL trip cost unexpectedly increases to determine

the option value of the MLs. However, this research identified that the log sum method frequently provided a poor estimate of the option value in the case where the generalized trip cost on the GPLs is much less than that on the MLs and the GPLs provide as reliable trip costs as the MLs do. Since, in this case, ML option users could not benefit from using the MLs, the option value of zero may be reasonable. However, in Case 1, 21.5 percent of the total estimates from the log sum method were not close to zero in those cases, whereas 1.7 percent of the total estimates from the Black-Scholes method were not close to zero in those cases.

In addition, this research found that the option value of the MLs using the Black-Scholes method partially depended on the volatility (the standard deviation) in the GPL trip cost: the higher the volatility in the GPL trip cost, the higher the option value of MLs. This volatility considers when the GPL trip cost unexpectedly increases. An important benefit of the ML option is the reduced trip cost by using the MLs when the GPL trip cost unexpectedly increases. However, the log sum method does not consider the volatility in the GPL trip cost to determine the option value of the MLs. Lastly, the option value using the log sum method was not sensitive to the value of travel time parameter, whereas the option value using Black-Scholes method increases as the value of travel time parameter increases. In general, travelers who have high values of travel time are likely to pay more to save their travel time than travelers who have low values of travel time. Thus, it is reasonable that an increase in the value of travel time should increase the option value of the MLs. However, in the log sum method, as the value of travel time parameter increased by 350%, the option value increased by only 26%. Thus,



this research recommends the Black-Scholes method to estimate the option value of MLs (see Sections 5.5 and 6.7 for details of the Black-Scholes method). Note that, from the sensitivity analysis, this research identified that the option value of the MLs using the Black-Scholes method is largely affected by the value of travel time parameter and the standard deviation of the generalized trip costs on the GPLs in 2012. Thus, when estimating the option value of MLs using the Black-Scholes method, those parameters should carefully be examined.

Using the revealed preference data for the option users, this research examined the option value of the MLs in 2012. The total option value of \$1,117,755 was estimated using the Black-Scholes method for the option users who occasionally used the MLs in 2012. Thus, this research recommends the option value of \$1,117,755 for the option value of the MLs in 2012. This amount was similar to the value of travel time savings from the MLs (one of the important actual benefits of the MLs) for the option users who occasionally used the MLs in 2012. Thus, the estimation of ML option use benefit is also important to accurately identify the societal benefits of MLs and will provide a new dimension to our understanding of the value of MLs.

To validate the option value of MLs, there much still remains to be addressed in the future work. First, ML option valuation study that adopts the stated preference (SP) data needs to be conducted. The option value of MLs refers to the willingness to pay for the potential use of MLs, not the actual use. Travelers on freeways could differently recognize the meaning of the potential use of MLs depending on their purpose of travel. For example, a traveler who frequently travels for urgent situations (e.g. attending an

important appointment/event) has high probability for the potential use of MLs and could value more for the ML option not to be late. Whereas, a traveler who only makes an ordinary trip (a typical trip in the week) has low probability for the potential use of MLs and could value less for the ML option. In addition, it was known that the value of travel time could be different depending on the purpose of travel, the ordinary trip versus a trip under the urgent situations (Patil et al. 2011). As shown in the sensitivity analysis in Section 6.5, the value of travel time parameter is an important factor that affects the option value of the MLs. The purpose of travel could not be considered from the RP data in this research and can be revealed from the SP data by directly asking their purpose of travel. Knowing the purpose of travelers on freeways could improve accuracy of the ML option value estimation.

Second, option users of MLs need to be defined. This research primarily considered the travelers who occasionally used the MLs in 2012 as the option user of MLs. These occasional ML users include travelers who have different percentages of ML use. For example, the occasional ML users include the traveler whose percentage of ML use is 5% and the traveler whose percentage of ML use is 95% without considering any weight for these travelers depending on the percentage of ML use. However, it could be possible that the traveler who frequently uses MLs values the ML option more than the traveler who less frequently uses MLs because the traveler who frequently uses MLs has high probability for the ML option use. In addition, this research assumed the option users of the MLs as the travelers who used the Katy Freeway in 2012 and excluded the travelers who did not use the Katy Freeway in 2012. This is because there could be very

few option users of the MLs among the travelers who did not use the Katy Freeway in 2012. But, for accurate estimation of the option value, whether the travelers who did not use the Katy Freeway value the MLs or not needs to be identified. These limitations of defining the option user of MLs need further considerations for the future work.

Last, the parameters, the value of travel time and the standard deviation (volatility) of the generalized trip costs on the GPLs, in the Black-Scholes method need to be carefully studied. This research recommends the Black-Scholes method to estimate the option value of MLs. In addition, in the sensitivity analysis in Section 6.5, this research identified that the estimate of the option value of MLs was largely affected by those parameters. However, those parameters could change. The value of travel time parameter changed depending on the option users of the MLs that were included in the estimation of the option value. Case 1 included only the users who occasionally used the MLs in 2012 and the value of travel time was estimated as \$14/hour in Case 1. Case 2 included the users who occasionally or never used the MLs in 2012 and the value of travel time was estimated as \$5.97/hour in Case 2. In addition, depending on attributes of each alternative (MLs or GPLs) included in utility functions, the value of travel time parameter might change. This research also identified that the standard deviation of the generalized trip costs on the GPLs changed even during the same hour when the trip was made. For example, on the westbound section between the sensors 425 and 396, the standard deviation of the generalized trip costs on the GPLs could be increased by up to 53% or decreased by up to 41% during the same hour when the trip was made. Thus, further considerations for the value of travel time for the option users of MLs and how

the option users recognize the volatility of the generalized trip cost on the GPLs could be desirable for the future work.

## REFERENCES

- Anderson, Simon P., and Andre de Palma. 1992. "The Logit as a Model of Product Differentiation." *Oxford Economic Papers* 44 (1): 51–67.
- Bateman, Ian J., Richard T. Carson, Brett Day, Michael Hanemann, Nick Hanley, Tannis Hett, and Michael Jones Lee. 2002. *Economic Valuation with Stated Preference Techniques: A Manual*. Cheltenham, UK ; Northampton, MA: Edward Elgar.
- Black, Fischer, and Myron Scholes. 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81 (3): 637–54.
- Bristow, A. L., P. G. Hopkinson, C. A. Nash, and M. R. Wardman. 1991. "Evaluation of the Use and Non-Use Benefits of Public Transport: Report Number 2 – Application of the Method." Monograph. January. <http://www.its.leeds.ac.uk/>.
- Burris, Mark, and Santosh Rao Danda. 2014. "Empirical Measurement of Travelers' Value of Reliability." <http://ntl.bts.gov/lib/45000/45900/45986/161304-1.pdf>.
- Burris, Mark W., and Sunil Patil. 2009. "Estimating the Benefits of Managed Lanes," September. <http://trid.trb.org/view.aspx?id=902316>.
- Burris, Mark W., and Bill R. Stockton. 2004. "HOT Lanes in Houston-Six Years of Experience." *Journal of Public Transportation* 7: 1–22.
- Burris, M., and E. Sullivan. 2006. "Benefit-Cost Analysis of Variable Pricing Projects: QuickRide HOT Lanes." *Journal of Transportation Engineering* 132 (3): 183–90.
- Cambridge Systematics, Inc., Transit Cooperative Research Program, National Research Council (U.S.). 1998. "Economic Impact Analysis of Transit Investments: Guidebook for Practitioners." Washington, D.C.: National Academy Press.
- Chang, Justin S. 2010. "Estimation of Option and Non-Use Values for Intercity Passenger Rail Services." *Journal of Transport Geography* 18 (2): 259–65.
- Chang, Justin S., Soo Yon Cho, Beom Shin Lee, Yonghwan Kim, and Suk Kang Yun. 2012. "A Dichotomous Choice Survey for Quantifying Option and Non-Use Values of Bus Services in Korea." *Transportation* 39 (1): 33–54.
- Chu, X., and S. E. Polzin. 1998. "The Value of Having a Public Transit Travel Choice." *Journal of Public Transportation* 2 (1): 91–116.

- Collier, Tina, and Ginger Daniels Goodin. 2002. *Managed Lanes: More Efficient Use of the Freeway System*. Texas Transportation Institute.
- Crockett, D. 1992. "Should Non-Use Benefits Be Included in Social Cost Benefit Analysis." *Institute for Transport Studies, University of Leeds, Leeds*.
- "Economics Basics: Elasticity." 2013. *Investopedia*. November 14.  
<http://www.investopedia.com/university/economics/economics4.asp>.
- ECO Northwest, Ltd, Parsons, Brinckerhoff, Quade & Douglas, United States, Transit Cooperative Research Program, Transit Development Corporation, and National Research Council (U.S.). 2002. *Estimating the Benefits and Costs of Public Transit Projects: A Guidebook for Practitioners*. TCRP Report 78. Washington, D.C: National Academy Press.
- "EZ TAG." 2014. *Wikipedia, the Free Encyclopedia*.  
[http://en.wikipedia.org/w/index.php?title=EZ\\_TAG&oldid=619641206](http://en.wikipedia.org/w/index.php?title=EZ_TAG&oldid=619641206).
- Fujiwara, Daniel, Ross Campbell, Great Britain, Department for Work and Pensions, Great Britain, and Treasury. 2011. *Valuation Techniques for Social Cost-Benefit Analysis Stated Preference, Revealed Preference and Subjective Well-Being Approaches: A Discussion of the Current Issues*. London: HM Treasury.  
[http://research.dwp.gov.uk/asd/asd5/greenbook\\_valuationtechniques.pdf](http://research.dwp.gov.uk/asd/asd5/greenbook_valuationtechniques.pdf).
- Geurs, Karst, Rinus Haaijer, and Bert Van Wee. 2006. "Option Value of Public Transport: Methodology for Measurement and Case Study for Regional Rail Links in the Netherlands." *Transport Reviews* 26 (5): 613–43.
- Halcrow. 2006. "Appraisal of Proposed Re-Opening of Station at Conon Bridge: Transport Economic Efficiency Analysis." Highland Rail Partnership and HITRANS.
- Hanley, Nick, Susana Mourato, and Robert E. Wright. 2001. "Choice Modelling Approaches: A Superior Alternative for Environmental Valuation?" *Journal of Economic Surveys* 15 (3): 435–62.
- Highland Rail Developments. 2000. "Beaulay Station Re-Opening Rail Passenger Partnership Funding Bid."
- Highland Rail Partnership. 2003. "Invernet Revenue and Cost Projections [Excel Spreadsheet]."

- Humphreys, Martin, and A. S. Fowkes. 2006. "The Significance of Indirect Use and Non-Use Values in Transport Appraisal." *International Journal of Transport Economics= Rivista Internazionale de Economia Dei Trasporti* 33 (1).
- Jacobs. 2006. "Airdrie-Bathgate Railway and Linked Improvements Bill Design Development Appraisal Report." Network Rail.
- Kockelman, Kara, Donna T. Chen, Katie Larsen, and Brice Nichols. 2013. *The Economics of Transportation Systems: A Reference for Practitioners (No. 0-6628-P1)*. The University of Texas at Austin: Center for Transportation Research.
- Laird, James, Karst Geurs, and Chris Nash. 2009. "Option and Non-Use Values and Rail Project Appraisal." *Transport Policy* 16 (4): 173–82.
- Litman, Todd. 2009. "Transportation Cost and Benefit Analysis." *Victoria Transport Policy Institute*. <http://www.vtpi.org/tca/tca01.pdf>.
- Montgomery, Douglas C. 2007. *Applied Statistics and Probability for Engineers*. 4th ed. Hoboken, NJ: Wiley.
- Natenberg, Sheldon. 1994. *Option Volatility & Pricing: Advanced Trading Strategies and Techniques*. Updated edition. McGraw-Hill.
- Painter, Kathleen M., Robert Douglas Scott, Philip R. Wandschneider, and Kenneth L. Casavant. 2002. "Using Contingent Valuation to Measure User and Nonuser Benefits: An Application to Public Transit." *Review of Agricultural Economics* 24 (2): 394–409.
- Parsons Brinckerhoff, Inc. SEH, and Texas Transportation Institute. 2010. "I-94 Managed Lanes Study." Minnesota Department of Transportation. <http://www.sehinc.com/files/online/I-94-Managed-Lanes-Study-Final-Report.pdf>.
- Patil, Sunil, Mark Burris, W. Douglass Shaw, and Sisinnio Concas. 2011. "Variation in the Value of Travel Time Savings and Its Impact on the Benefits of Managed Lanes." *Transportation Planning and Technology* 34 (6): 547–67.
- Perloff, Jeffrey M., and Steven C. Salop. 1985. "Equilibrium with Product Differentiation." *The Review of Economic Studies* 52 (1): 107–20.
- "Risk-Free Return." 2015. *Investopedia*. March 14. <http://www.investopedia.com/terms/r/risk-freereturn.asp>.

- Roson, Roberto. 2000. "Social Cost Pricing When Public Transport Is an Option Value." *Innovation: The European Journal of Social Science Research* 13 (1): 81–94.
- Sattinger, Michael. 1984. "Value of an Additional Firm in Monopolistic Competition." *The Review of Economic Studies* 51 (2): 321–32.
- Shapiro, S. S., and M. B. Wilk. 1965. "An Analysis of Variance Test for Normality (complete Samples)." *Biometrika* 52 (3-4): 591–611.
- Sisiopiku, Virginia P., Andrew Sullivan, Ozge Cavusoglu, Saiyid Sikder, Kyriacos Mouskos, and Curtis Barrett. 2009. "Managed Lanes: Current Status and Future Opportunities," August. <http://trid.trb.org/view/2009/M/1086124>.
- Small, Kenneth A. 2007. *The Economics of Urban Transportation*. New York: Routledge.
- Small, Kenneth A., and Harvey S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica* 49 (1): 105–30.
- SRF Consulting Group, INC., Inc. Kimley-Horn & Associates, and Inc. ZAN Associates. 2013. "I-35W North Managed Lanes Corridor Study." Minnesota Department of Transportation. <http://www.dot.state.mn.us/metro/projects/i35wstudy/pdf/finalreport.pdf>.
- Statista, Inc. 2015. "U.S. Vehicles: Average Length of Ownership in 2001 and 2012 (in Months)." <http://www.statista.com/statistics/276537/length-of-vehicle-ownership-in-the-united-states/>.
- Suen, Wing. 1991. "The Value of Product Diversity." *Oxford Economic Papers* 43 (2): 217–23.
- Sullivan, E., and M. Burris. 2006. "Benefit-Cost Analysis of Variable Pricing Projects: SR-91 Express Lanes." *Journal of Transportation Engineering* 132 (3): 191–98.
- Sullivan, Edward. 2000. "Continuation Study to Evaluate the Impacts of the SR 91 Value-Priced Express Lanes: Final Report." *Submitted to State of California Department of Transportation Traffic Operations Program HOV Systems Branch, Sacramento CA 94273*.



[http://ops.fhwa.dot.gov/tolling\\_pricing/value\\_pricing/pubs\\_reports/projectreports/pdfs/sr91\\_expresslanes.pdf](http://ops.fhwa.dot.gov/tolling_pricing/value_pricing/pubs_reports/projectreports/pdfs/sr91_expresslanes.pdf).

Train, Kenneth E. 1994. "Self-Selecting Tariffs under Pure Preferences among Tariffs." *Journal of Regulatory Economics* 6 (3): 247–64.

Treasury, H. M. 2003. *The Green Book-Appraisal and Evaluation in Central Government TSO*. London.

UK Department for Transport. 2006. "Consultation Document on the Implementation of the Railways Act 2005 Provisions on Closures and Minor Modifications." Consultation. January 26.  
<http://www.scotland.gov.uk/Publications/2006/01/26094938/0>.

UK Department for Transport, Great Minster House. 2009. "Department for Transport - Transport Analysis Guidance - WebTAG." Web page. August 17.  
<http://www.dft.gov.uk/webtag/documents/index.php>.

U.S. Department of the Treasury. 2015. "Daily Treasury Bill Rates."  
<http://www.treasury.gov/resource-center/data-chart-center/interest-rates/Pages/TextView.aspx?data=billrates>.

"Volatility Definition." 2015. *Investopedia*. March 8.  
<http://www.investopedia.com/terms/v/volatility.asp>.

Wallis, Ian, D Wignall, and NZ Transport Agency. 2012. *The Benefits of Public Transport: Option Values and Non-Use Values*. Wellington, N.Z.: NZ Transport Agency.

Weitzman, Martin L. 1992. "On Diversity." *The Quarterly Journal of Economics* 107 (2): 363–405.

## APPENDIX A

### A.1 Descriptions of Transportation Cost Categories (Litman 2009)

Cost	Description
Vehicle ownership	Fixed costs of owning a vehicle.
Vehicle operating	Variable vehicle costs, including fuel, oil, tires, tolls and short-term parking fees.
Internal: parking	Off-street residential parking and long-term leased parking paid by users.
Travel time	The value of time used for travel.
Internal: crash	Crash costs borne directly by travelers.
Internal: activity	Health benefits of active transportation to travelers (a cost where foregone).
Operating subsidies	Financial subsidies for public transit services.
External: parking	Off-street parking costs not borne directly by users
Road facilities	Roadway facility construction and operating expenses not paid by user fees.
Roadway land value	The value of land used in public road rights-of-way.
Traffic services	Costs of providing traffic services such as traffic policing, and emergency services.
Transportation diversity	The value to society of a diverse transport system, particularly for non-drivers.
Air pollution	Costs of vehicle air pollution emissions.
Noise	Costs of vehicle noise pollution emissions.
Barrier effect	Delays that roads and traffic cause to non-motorized travel.
Water pollution	Water pollution and hydrologic impacts caused by transport facilities and vehicles
Waste disposal	External costs associated with disposal of vehicle wastes.
Congestion	Congestion costs imposed on other road users.
Greenhouse gas	Lifecycle costs of greenhouse gases that contribute to climate change.
Resource consumption	External costs of resource consumption, particularly petroleum.
Land use impacts	Increased costs of sprawled, automobile-oriented land use.
External: crash	Crash costs a traveler imposes on others.
Healthful activity	Health benefits of active transportation to society

## APPENDIX B

### B.1 SAS Program Code to Generate RP Data on Katy Freeway

```
/* -----
- Program Name: RP Data on Katy Freeway -----
- Date Created: 02/19/2014 -----
- Author: Sunghoon Lee -----
- Inputs: 1. raw2012.sas7bdat, 2. toll files, 3. distance.sas7bdat ----
- Outputs: 1. trip12.sas7bdat, 2. tripend.sas7bdat,
           3. tripchoice12.sas7bdat, 4. trip12_seg_tt_sd.sas7bdat,
           5. trip12stddev.sas7bdat
-----
----- */

libname katy 'D:\Thesis Data\My Work';
libname katy2 'D:\Backup\Backup_1022';
libname katy100 'D:\Thesis Data\My Work\3.Analysis_V2\Backup_0501_Case2
Sensitivity';

/* PART 1: Random ID -> Result: raw2012.sas7bdat -----*/

/* PART 2 -----*/
/* 2.1 Create Trip12 Table */
proc sql noprint;
    create table katy.trip12_2 as select * from katy.raw2012;
quit;

proc sql noprint;
    Create Table Katy.trip12 as select distinct * from Katy.trip12_2;
quit;
/* trip12: all trips in 2012 */

/* 2.2 Delete ML Sensor (451, 454, 455, 458, 459) Records */
Data katy.trip12;
    set katy.trip12;
    if ((sensor=451) or (sensor=454) or (sensor=455) or (sensor=458)
        or (sensor=459)) then delete;
run;

/* 2.3 Sort Trip12 according to Vehicle ID, Datetime Order */
proc sort data=katy.trip12;
    by descending randid datetime; /* randid 1st, datetime 2nd */
run;

data katy.trip12;
    set katy.trip12;
    obs= _N_; /* making order */
run;
```

```

/* 2.4 Add Corresponding Sensors for DUMMY Trip */
/* sensor= real trip, sensor1= dummy trip */
data katy.trip12;
    set katy.trip12;
        if sensor=468 then sensor1=466;
    else if sensor=466 then sensor1=468;
    else if sensor=469 then sensor1=467;
    else if sensor=467 then sensor1=469;
    else if sensor=453 then sensor1=444;
    else if sensor=444 then sensor1=453;
    else if sensor=456 then sensor1=445;
    else if sensor=445 then sensor1=456;
    else if sensor=460 then sensor1=425;
    else if sensor=425 then sensor1=460;
    else if sensor=101 then sensor1=465;
    else if sensor=102 then sensor1=465;
    else if sensor=465 then sensor1=101;
    else if sensor=106 then sensor1=440;
    else if sensor=105 then sensor1=440;
    else if sensor=440 then sensor1=105;
    else if lag(randid)= randid and lag(sensor)=443 and sensor=103
        then sensor1=466;
    else if sensor=103 then sensor1=443;
    else if lag(randid)= randid and lag(sensor)=443 and sensor=104
        then sensor1=466;
    else if sensor=104 then sensor1=443;
    else if sensor=443 then sensor1=103;
    else if lag(randid)= randid and lag(sensor)=441 and sensor=107
        then sensor1=445;
    else if sensor=107 then sensor1=441;
    else if lag(randid)= randid and lag(sensor)=441 and sensor=108
        then sensor1=445;
    else if sensor=108 then sensor1=441;
    else if sensor=441 then sensor1=107;
    else if lag(randid)= randid and lag(sensor)=109 and sensor=442
        then sensor1=442;
    else if lag(randid)= randid and lag(sensor)=110 and sensor=442
        then sensor1=442;
    else if sensor=442 then sensor1=109;
    else if sensor=109 then sensor1=442;
    else if sensor=110 then sensor1=442;
    else if lag(randid)= randid and lag(sensor)=111 and sensor=396
        then sensor1=396;
    else if lag(randid)= randid and lag(sensor)=112 and sensor=396
        then sensor1=396;
    else if sensor=396 then sensor1=111;
    else if sensor=111 then sensor1=396;
    else if sensor=112 then sensor1=396;
    else sensor1=sensor;

run;

proc sort data=katy.trip12;
    by decending obs;

```

```

run;

data katy.trip12;
  set katy.trip12;
  if lag(randid)= randid and lag(sensor)=103 and sensor=443
  then sensor1=443;
  else if lag(randid)= randid and lag(sensor)=104 and sensor=443
  then sensor1=443;
  else if lag(randid)= randid and lag(sensor)=107 and sensor=441
  then sensor1=441;
  else if lag(randid)= randid and lag(sensor)=108 and sensor=441
  then sensor1=441;
  else if lag(randid)= randid and lag(sensor)=442 and sensor=109
  then sensor1=467;
  else if lag(randid)= randid and lag(sensor)=442 and sensor=110
  then sensor1=467;
  else if lag(randid)= randid and lag(sensor)=396 and sensor=111
  then sensor1=442;
  else if lag(randid)= randid and lag(sensor)=396 and sensor=112
  then sensor1=442;

run;

proc sort data=katy.trip12;
  by obs;
run;

/* 2.5 Add Trip Direction */
/* 0= East Bound, 1= West Bound */
data katy.trip12;
  set katy.trip12;
  if (sensor = 107 or sensor = 108 or sensor = 109 or sensor = 110
  or sensor = 111 or sensor = 112 or sensor = 416 or sensor =
  415 or sensor = 414 or sensor = 369 or sensor = 396 or sensor
  = 442 or sensor = 467 or sensor = 445 or sensor = 441 or
  sensor = 425 or sensor = 6 or sensor = 272 or sensor = 199 or
  sensor = 459 or sensor = 458 or sensor = 469 or sensor = 456
  or sensor = 455 or sensor = 460)
  then direction = 1;
  else direction = 0;

run;

/* 2.6 Add Sensor Type for Real Trip and Dummy Trip */
/* M= Managed L, T= Toll L, H=HOV L, G=General Purpose L, O=Outside
Managed L */
/* location= Real Trip, location1=Dummy Trip */
/* 2.6.1 Real Trip */
data katy.trip12;
  set katy.trip12;
  if (sensor = 459 or sensor = 458 or sensor = 469 or sensor = 456
  or sensor = 455 or sensor = 460 or sensor = 449 or sensor =
  451 or sensor = 468 or sensor = 453 or sensor = 454)
  then location = 'M';
  else if (sensor = 101 or sensor = 103 or sensor = 105 or sensor =

```

```

        107 or sensor = 109 or sensor = 111) then location =
'T';
    else if (sensor = 102 or sensor = 104 or sensor = 106 or sensor =
        108 or sensor = 110 or sensor = 112) then location =
'H';
    else if (sensor = 396 or sensor = 442 or sensor = 467 or sensor =
        445 or sensor = 441 or sensor = 425 or sensor = 440 or
        sensor = 444 or sensor = 466 or sensor = 443 or sensor =
        427 or sensor = 465 or sensor = 426) then location =
'G';
    else location = 'O';
run;

/* 2.6.2 Dummy Trip */
data katy.tripl2;
    set katy.tripl2;
    if (sensor1 = 459 or sensor1 = 458 or sensor1 = 469 or sensor1 =
        456 or sensor1 = 455 or sensor1 = 460 or sensor1 = 449 or
        sensor1 = 451 or sensor1 = 468 or sensor1 = 453 or sensor1 =
        454) then location1 = 'M';
    else if (sensor1 = 101 or sensor1 = 103 or sensor1 = 105 or
        sensor1 = 107 or sensor1 = 109 or sensor1 = 111)
        then location1 = 'T';
    else if (sensor1 = 102 or sensor1 = 104 or sensor1 = 106 or
        sensor1 = 108 or sensor1 = 110 or sensor1 = 112)
        then location1 = 'H';
    else if (sensor1 = 396 or sensor1 = 442 or sensor1 = 467 or
        sensor1 = 445 or sensor1 = 441 or sensor1 = 425 or
        sensor1 = 440 or sensor1 = 444 or sensor1 = 466 or
        sensor1 = 443 or sensor1 = 427 or sensor1 = 465 or
        sensor1 = 426)
        then location1 = 'G';
    else location1 = 'O';
run;

/* 2.7 Identify Trip Start and Trip End (Trip Ends) */
/* Start, End =1, otherwise = 0 */
/* 2.7.1 Trip Start */
data katy.tripl2;
    set katy.tripl2;
    start = 0;
    end = 0;
    timediff = datetime-lag(datetime);
    if randid ne lag(randid) then start=1;
    if (timediff > hms(0,15,0)) then start=1;
    if (direction ne lag(direction)) then start=1;
run;

/*2.7.2 Trip End */
data katy.tripl2;
    set katy.tripl2;
    obs= _N_; /* making order */
run;

```

```

proc sort data= katy.trip12;
    by descending obs; /* reverse */
run;

data katy.trip12;
    set katy.trip12;
    Desc_obs= _N_; /* making order */
run;

data katy.trip12;
    set katy.trip12;
    if lag(start)=1 then end=1;
    if desc_obs=1 then end=1;
    drop desc_obs;
run;

proc sort data= katy.trip12;
    by obs; /* return to original */
run;

/* 2.8 Delete Trips which are only detected at one Sensor */
data katy.trip12;
    set katy.trip12;
    if ((start=1) and (end=1)) then delete;
run;

/* 2.9 Delete Error Detections-Repeated Detections at one sensor */
data katy.trip12;
    set katy.trip12;
    if (randid=468736080 and Sensor=445 and
        datepart(datetime)='23APR2012'd) then delete;
run;

data katy.trip12;
    set katy.trip12;
    if ((randid=928115562 and Sensor=6 and
        datepart(datetime)='12SEP2012'd) and (hour(datetime)=13 or
        hour(datetime)=14)) then delete;
run;

data katy.trip12;
    set katy.trip12;
    if (randid=677801105 and Sensor=271 and
        datepart(datetime)='16AUG2012'd) then delete;
run;

/* 2.10 Create variable with all sensors for a trip */
data Katy.trip12;
    set Katy.trip12;
    length Allsensor Allsensor1 $100;
    retain INTER INTER1;
    if Start = 1 then

```

```

do;
    Allsensor = catx(",", sensor);
    INTER = Allsensor;
    Allsensor1 = catx(",", sensor1);
    INTER1 = Allsensor1;
end;
else
do;
    Allsensor = catx(",", INTER, sensor );
    INTER = Allsensor;
    Allsensor1 = catx(",", INTER1, sensor1 );
    INTER1 = Allsensor1;
end;
drop INTER INTER1;
run;

/* 2.11 Create variable with all sensor types for a trip */
Data Katy.trip12;
    length Weave Weave1 $50;
    length INTER INTER1 $50;
    set Katy.trip12;
    retain INTER INTER1;
    if Start = 1 then
        do;
            Weave = catx(" ", location);
            INTER = Weave;
            Weave1 = catx(" ", location1);
            INTER1 = Weave1;
        end;
    else
        do;
            Weave = catx(" ", INTER, location);
            INTER = Weave;
            Weave1 = catx(" ", INTER1, location1);
            INTER1 = Weave1;
        end;
    drop INTER INTER1;
run;

/* 2.12 Determine trip start sensor, trip start time, trip end, and
trip end time */
/* 2.12.1 For Real Trip */
Data Katy.trip12;
    set Katy.trip12;
    retain Inter Inter2;
    if start=1 then
        do;
            Start_Sensor= Sensor;
            Inter=Start_Sensor;
            Starttime=datetime;
            Inter2= Starttime;
            Time=0;
            End_Sensor=0;

```



```

        Endtime=0;
    end;

    if end=1 then
        do;
            Start_Sensor= Inter;
            Starttime = Inter2;
            End_Sensor=Sensor;
            Endtime=datetime;
            Time= datetime-Inter2;
        end;

    if ((start=0) and (end=0)) then
        do;
            Start_Sensor= Inter;
            Inter= Start_Sensor;
            Starttime=Inter2;
            Inter2=Starttime;
            Time=0;
            End_Sensor=0;
            Endtime=0;
        end;
    drop inter inter2;
run;

/* 2.12.2 For Dummy Trip: Identifying only Start and End Sensors */
Data Katy.trip12;
set Katy.trip12;
retain Inter;
if start=1 then
    do;
        Start_Sensor1= Sensor1;
        Inter=Start_Sensor1;
        End_Sensor1=0;
    end;

if end=1 then
    do;
        Start_Sensor1= Inter;
        End_Sensor1=Sensor1;
    end;

if ((start=0) and (end=0)) then
    do;
        Start_Sensor1= Inter;
        Inter= Start_Sensor1;
        End_Sensor1=0;
    end;
drop inter;
run;

/* 2.13 Input trip end sensor and trip end time in all the trip
observations */

```

```

proc sort data= Katy.trip12;
    by descending obs;
run;
Data Katy.trip12;
    set Katy.trip12;
    retain Inter Inter2 Inter3;
    if end=1 then
        do;
            Inter= Time;
            Inter2= End_Sensor;
            Inter3= Endtime;
        end;
    if (((start=0) and (end=0)) or (Start=1)) then
        do;
            Time= Inter;
            Inter= Time;
            End_Sensor= Inter2;
            Inter2= End_Sensor;
            Endtime= Inter3;
            Inter3= Endtime;
        end;
    drop Inter Inter2 Inter3;
run;

Data Katy.trip12;
    set Katy.trip12;
    retain Inter2;
    if end=1 then
        do;
            Inter2= End_Sensor1;
        end;
    if (((start=0) and (end=0)) or (Start=1)) then
        do;
            End_Sensor1= Inter2;
            Inter2= End_Sensor1;
        end;
    drop Inter2;
run;

proc sort data= Katy.trip12;
    by obs;
run;

/* PART 3 -----*/
/* 3.1 Input incident (lane closure) information */
/* 3.1.1 Import the lane closure data file */
proc import
    datafile = "D:\Thesis Data\My
               Work\Incidents\Laneclosure2012_brad.xlsx"
    out = Katy.laneclosure
    dbms = xlsx;
    sheet = "2012_I-10Katy_AllIncidents";
    getnames = yes;

```

```

run;

Proc sort data=Katy.laneclosure;
    by cleared_date_time;
run;

data Katy.laneclosure;
    set Katy.laneclosure;
        incidentstart=detection_date_time;
        incidentend=cleared_date_time;
/*Delete incidents in which start or end time is not properly record.*/
    if incidentend=253717919999 then delete;
run;

data Katy.trip12;
    set Katy.trip12;
    detectiontime=datetime;
run;

/* 3.1.2 Left joining table based on time of incident and nearest
sensor criteria */
proc sql noprint;
    create table Katy.xyz as
    select A.*, B.mainlanes_blocked, B.frontage_lanes_blocked,
        B.ramp_lanes_blocked, B.HOV_lanes_blocked,
        B.shoulder_lanes_blocked,
        B.detection_date_time, B.cleared_date_time
    from Katy.trip12 as A left join Katy.laneclosure as B
        on (B.incidentstart < A.detectiontime < B.incidentend) and
            (A.sensor= B.nearest_avi_sensor);

quit;

data Katy.trip12;
    set Katy.xyz;
run;

proc sort data= Katy.trip12;
    by obs;
run;

/* 3.1.3 Check for the highest number of lanes blocked are recorded for
each trip */
Data Katy.trip12;
    set Katy.trip12;
    retain mainlanes_blocked frontage_lanes_blocked
        ramp_lanes_blocked HOV_lanes_blocked
        Shoulder_lanes_blocked;
    inter1= lag1(mainlanes_blocked);
    inter2= lag1(frontage_lanes_blocked);
    inter3= lag1(ramp_lanes_blocked);
    inter4= lag1(HOV_lanes_blocked);
    inter5= lag1(Shoulder_lanes_blocked);
    if ((start=0) and (end=1)) or ((start=0) and (end=0)) and

```

```

        (inter1 ne . or inter2 ne . or inter3 ne . or inter4 ne .
        or inter5 ne .)
    then do;
        mainlanes_blocked= max(mainlanes_blocked, inter1);
        frontage_lanes_blocked= max(frontage_lanes_blocked,
inter2);
        ramp_lanes_blocked= max(ramp_lanes_blocked, inter3);
        HOV_lanes_blocked= max(HOV_lanes_blocked, inter4);
        Shoulder_lanes_blocked= max(Shoulder_lanes_blocked,
inter5);
    end;
run;

data Katy.trip12;
    set Katy.trip12;
    drop inter1 inter2 inter3 inter4 inter5  detection_date_time
        cleared_date_time ;
run;

Data Katy.trip12_1;
    set Katy.trip12;
run;

Data Katy.trip12;
    set Katy.trip12_1;
run;

Data Katy.trip12;
    set Katy.trip12;
    if ((randid=lag(randid)) and (sensor=lag(sensor)) and
        (datetime=lag(datetime))) then delete;
run;

/* 3.2 Input total trip segments */
Data Katy.trip12;
    set Katy.trip12;
    if start=1 then
        do;
            segmentno=0;
        end;
    if start=0 then
        do;
            segmentno+1;
        end;
run;

/* PART 4 -----*/
/* 4.1 Assign Trip Length */
Proc sort data= katy.distance;
    by sensor;
run;

Proc sort data= Katy.trip12;

```

```

        by sensor;
run;

data Katy.trip12;
    merge Katy.trip12 katy.distance;
    by sensor;
run;

/* 4.2 Delete Data that doesn't have datetime information */
Data Katy.trip12;
    set Katy.trip12;
    if datetime=. then delete;
run;

proc sort data= Katy.trip12;
    by obs;
run;

/* 4.3 Calculate segment length and check negative lengths */
Data Katy.trip12;
    set Katy.trip12;
    retain inter;
    if (start=1) then
        do;
            Inter = coord;
            seglen=0;
        end;

    if (start=0) then
        do;
            seglen=(coord-inter);
            Inter= coord;
            if seglen<0 then flag=1;
        end;
    drop inter;
run;

/* 4.4 Input complete weave information to all trips */
proc sort data= Katy.trip12;
    by descending obs;
run;

Data Katy.trip12;
    set Katy.trip12;
    retain Inter Inter1;

    if end=1 then
        do;
            Inter= Weave;
            Inter1= Weave1;
        end;

    if (((start=0) and (end=0)) or (Start=1)) then

```

```

        do;
            Weave= Inter;
            Inter= Weave;
            Weave1= Inter1;
            Inter1= Weave1;
        end;
    drop Inter Inter1;
run;

proc sort data= Katy.trip12;
    by obs;
run;

/* 4.5 Delete data that has negative segment length and data affected
by the negative segment length */
Data Katy.trip12;
    set Katy.trip12;
    retain inter;
    if ((start=0) and (inter=1)) then flag=1;
    inter=flag;
    drop inter;
run;

proc sort data= Katy.trip12;
    by descending obs;
run;

Data Katy.trip12;
    set Katy.trip12;
    retain inter;
    if end=0 then
        do;
            flag=inter;
        end;
    inter=flag;
    drop inter;
run;

proc sort data= Katy.trip12;
    by obs;
run;

Data Katy.trip12;
    set Katy.trip12;
    if flag=1 then delete;
    drop flag;
run;

/* 4.6 Calculate distance to travel (length) */
Data Katy.trip12;
    set Katy.trip12;
    retain inter disttravel;

```

```

        if (start=1) then
            do;
                length = 0;
                Inter = coord;
                disttravel=0;
            end;

        if (start=0) then
            do;
                length= (coord-inter)+disttravel;
                Inter= coord;
                disttravel=length;
            end;
        drop inter disttravel;
run;

Data katy.distance;
    set katy.distance(rename=(sensor= sensor1 coord=coord1));
run;

proc sort data= Katy.trip12;
    by sensor1;
run;

data Katy.trip12;
    merge Katy.trip12 katy.distance;
    by sensor1;
run;

proc sort data= Katy.trip12;
    by obs;
run;

data Katy.trip12;
    set Katy.trip12;
    retain inter disttravel;

    if (start=1) then
        do;
            length1 = 0;
            Inter = coord1;
            disttravel=0;
        end;

    if (start=0) then
        do;
            length1= (coord1-inter)+disttravel;
            Inter= coord1;
            disttravel=length1;
        end;
    drop inter disttravel;
run;

```

```

Data katy.distance;
    set katy.distance(rename=(sensor1= sensor coord1=coord));
run;

/*Additional step to delete data that does not have datetime info. */
Data Katy.trip12;
    set Katy.trip12;
    if datetime=. then delete;
run;

/* 4.7 Assign Toll, Toll1, Totaltoll and Totaltoll1 */
/* 4.7.1 Assign Toll */
/* Hour and Weekday */
data katy.trip12;
    set katy.trip12;
    hour=hour(datetime);
    weekday=weekday(datepart(datetime));
run;

/* Before and After Sep. 08 */
/* Before Sep. 08 = Presep08=1, After Sep. 08 = Presep08=0 */
data katy.trip12;
    set katy.trip12;
    presep08=0;
    if datepart(datetime) < '08sep2012'd then presep08=1;
run;

/* Peak (peak1=1 and peak2=1), Shoulder Peak (peak1=0 and peak2=1) and
off peak (peak1=0 and peak2=0) */
/* Before Sep. 08
    Peak1= EB:7-9 WB: 17-19,
    Peak2= Including shoulder peak EB: 6-10 WB:16-20 */
/* After Sep. 08
    Peak1= EB:7-9 WB: 16-18,
    Peak2= Including shoulder peak EB: 6-10 WB:15-19 */
Data katy.trip12;
    set katy.trip12;
    peak1=0;
    peak2=0;
    if ((presep08=1) and (weekday ne 1) and (weekday ne 7)) then
        do;
            if ((direction=0) and (7<=hour<=8)) or ((direction=1)
                and (17<=hour<=18)) then peak1=1;
            if ((direction=0) and (6<=hour<=9)) or ((direction=1)
                and (16<=hour<=19)) then peak2=1;
        end;
    else if ((presep08=0) and (weekday ne 1) and (weekday ne 7)) then
        do;
            if ((direction=0) and (7<=hour<=8)) or ((direction=1)
                and (16<=hour<=17)) then peak1=1;
            if ((direction=0) and (6<=hour<=9)) or ((direction=1)
                and (15<=hour<=18)) then peak2=1;
        end;

```



```

end;

run;

/* HOV Hour */
/* HOV=1 (Free) HOV=0 (Toll) */
Data katy.trip12;
  set katy.trip12;
  HOV=0;
  if (location = 'H') and (weekday ne 1 and weekday ne 7) and
    ((5<=hour<=10) or (14<=hour<=19)) then HOV=1;

run;

/* Assign Toll Based on Sensor */
Data katy.trip12;
  set katy.trip12;
  toll=0.0;
  if sensor=101 or sensor=111 then
    do;
      if peak1=1 and peak2=1 and presep08=1 then toll=1.6;
      if peak1=1 and peak2=1 and presep08=0 then toll=2.2;
      if peak1=0 and peak2=1 and presep08=1 then toll=0.8;
      if peak1=0 and peak2=1 and presep08=0 then toll=1.1;
      if peak1=0 and peak2=0 and presep08=1 then toll=0.4;
      if peak1=0 and peak2=0 and presep08=0 then toll=0.4;
    end;
  else if sensor=102 or sensor=112 then
    do;
      if HOV=1 then toll=0.0;
      if HOV=0 then toll=0.4;
    end;
  else if sensor=103 or sensor=109 or sensor=105 or sensor=107 then
    do;
      if peak1=1 and peak2=1 and presep08=1 then toll=1.2;
      if peak1=1 and peak2=1 and presep08=0 then toll=1.4;
      if peak1=0 and peak2=1 and presep08=1 then toll=0.6;
      if peak1=0 and peak2=1 and presep08=0 then toll=0.7;
      if peak1=0 and peak2=0 and presep08=1 then toll=0.3;
      if peak1=0 and peak2=0 and presep08=0 then toll=0.3;
    end;
  else if sensor=104 or sensor=110 or sensor=106 or sensor=108 then
    do;
      if HOV=1 then toll=0.0;
      if HOV=0 then toll=0.3;
    end;

run;

/* 4.7.2 Assign Toll1 */
/* Assign Toll1 Based on Sensor1 */
Data katy.trip12;
  set katy.trip12;
  toll1=0.0;

```

```

    if sensor1=101 or sensor1=111 then
        do;
            if peak1=1 and peak2=1 and presep08=1 then toll1=1.6;
            if peak1=1 and peak2=1 and presep08=0 then toll1=2.2;
            if peak1=0 and peak2=1 and presep08=1 then toll1=0.8;
            if peak1=0 and peak2=1 and presep08=0 then toll1=1.1;
            if peak1=0 and peak2=0 and presep08=1 then toll1=0.4;
            if peak1=0 and peak2=0 and presep08=0 then toll1=0.4;
        end;
    else if sensor1=103 or sensor1=109 or sensor1=105 or sensor1=107
        then
            do;
                if peak1=1 and peak2=1 and presep08=1 then toll1=1.2;
                if peak1=1 and peak2=1 and presep08=0 then toll1=1.4;
                if peak1=0 and peak2=1 and presep08=1 then toll1=0.6;
                if peak1=0 and peak2=1 and presep08=0 then toll1=0.7;
                if peak1=0 and peak2=0 and presep08=1 then toll1=0.3;
                if peak1=0 and peak2=0 and presep08=0 then toll1=0.3;
            end;
        end;
run;

/*Summing up the tolls along a trip*/
data Katy.trip12;
    set Katy.trip12;
    retain inter inter1;
    if (start=1) then
        do;
            Totaltoll = toll;
            Inter = Totaltoll;
            Totaltoll1 = toll1;
            Inter1 = Totaltoll1;
        end;
    if (start=0) then
        do;
            Totaltoll= toll + Inter;
            Inter= Totaltoll;
            Totaltoll1= toll1 + Inter1;
            Inter1= Totaltoll1;
        end;
    drop inter inter1;
run;

/* 4.8 Delete Records that have zero length (detected at the same
sensor twice with certain time interval */
data katy.trip12;
    set katy.trip12;
    if start=0 and length = lag(length) then flag=1;
run;

Data Katy.trip12;
    set katy.trip12;
    retain inter;
    if ((start=0) and (inter=1)) then flag=1;

```

```

        inter=flag;
        drop inter;
run;

proc sort data= Katy.trip12;
    by descending obs;
run;

Data Katy.trip12;
    set Katy.trip12;
    retain inter;
    if end=0 then
        do;
            flag=inter;
        end;
    inter=flag;
    drop inter;
run;

proc sort data= Katy.trip12;
    by obs;
run;

Data Katy.trip12;
    set Katy.trip12;
    if flag=1 then delete;
    drop flag;
run;

/* 4.9 Delete Records that have negative length in dummy trip records*/
data katy.trip12;
    set katy.trip12;
    if start=0 and length1 < lag(length1) then flag=1;
run;

Data Katy.trip12;
    set katy.trip12;
    retain inter;
    if ((start=0) and (inter=1)) then flag=1;
    inter=flag;
    drop inter;
run;

proc sort data= Katy.trip12;
    by descending obs;
run;

Data Katy.trip12;
    set Katy.trip12;
    retain inter;
    if end=0 then
        do;

```

```

                                flag=inter;
                                end;
                                inter=flag;
                                drop inter;
run;
proc sort data= Katy.trip12;
    by obs;
run;

Data Katy.trip12;
    set Katy.trip12;
    if flag=1 then delete;
    drop flag;
run;

/* PART 5 -----*/
/* 5.1 Calculate TT and STD */
/* 5.1.1 Create Dataset (trip12_seg_TT_SD) that has TT and STD */
data Katy.trip12;
    set Katy.trip12;
    startdate=datepart(datetime);
    starttime_sec = timepart(starttime);
    marker10= floor(starttime_sec/600);
    starttime_10min=marker10*600;
run;

Data Katy.trip12;
    set Katy.trip12;
    retain sensor sensor1;
    startsensor1=lag(sensor);
    endsensor1=sensor;
    startsensor11=lag(sensor1);
    endsensor11=sensor1;
    count=1;
    if start=1 then timediff=.;
run;

Data abcl;
    set Katy.trip12;
    if timediff=. then delete;
run;

Proc sql noprint;
    Create Table Katy.trip12_seg_TT_SD
    as select startdate, startsensor1, endsensor1, starttime_10min,
            sum(count) as Count_10_segment, avg(timediff) as
            TT_ave_10min_segment, std(timediff) as
            TT_std_10min_segment
    from abcl
    group by startdate, startsensor1, endsensor1, starttime_10min;
quit;

/* 5.1.2 Merging data (Trip12 + Trip12_seg_TT_SD) */

```

```

/*For real trip*/
proc sort data= Katy.trip12;
    by startdate startsensor1 endsensor1 starttime_10min;
run;

Data Katy.trip12;
    merge Katy.trip12 Katy.trip12_seg_TT_SD;
    by startdate startsensor1 endsensor1 starttime_10min;
run;

proc sort data= Katy.trip12;
    by descending randid datetime;
run;

/*Deleting obs. that do not have corresponding obs. */
Data Katy.trip12;
    set Katy.trip12;
    if randid=. then delete;
run;

/*For dummy trip*/
Data Katy.trip12_seg_TT_SD;
    set Katy.trip12_seg_TT_SD (rename=(startsensor1=startsensor11
    endsensor1=endsensor11 Count_10_segment=Count_10_segment1
    TT_ave_10min_segment=TT_ave_10min_segment1
    TT_std_10min_segment=TT_std_10min_segment1 ));
run;

proc sort data= Katy.trip12;
    by startdate startsensor11 endsensor11 starttime_10min;
run;

Data Katy.trip12;
    merge Katy.trip12 Katy.trip12_seg_TT_SD;
    by startdate startsensor11 endsensor11 starttime_10min;
run;

proc sort data= Katy.trip12;
    by descending randid datetime;
run;

/*Deleting obs. that do not have corresponding obs. */
Data Katy.trip12;
    set Katy.trip12;
    if randid=. then delete;
run;

/*Renaming variables back to original name*/
Data Katy.trip12_seg_TT_SD;
    set Katy.trip12_seg_TT_SD(rename=(startsensor11=startsensor1
    endsensor11=endsensor1 Count_10_segment1=Count_10_segment
    TT_ave_10min_segment1=TT_ave_10min_segment
    TT_std_10min_segment1=TT_std_10min_segment));

```

```

run;

/* 5.1.3 Assigning segment lengths to all segments */
Data Katy.trip12;
    set Katy.trip12;
    retain inter inter1;

    if (start=1) then
        do;
            segmentlength = 0;
            Inter = length;
            segmentlength1 = 0;
            Inter1 = length1;
        end;

    if (start=0) then
        do;
            segmentlength= (length-Inter);
            Inter= length;
            segmentlength1= (length1-Inter1);
            Inter1= length1;
        end;
    drop inter inter1;
run;

/* 5.1.4 Additional step to include better speed approximations */
/* type=1 -> managed lane travel, type=2 -> GPL travel */
data Katy.trip12;
    set Katy.trip12(drop = startdate starttime_sec marker10
                    starttime_10min);

    g= countc(Weave, 'G');
    m= countc(Weave, 'M');
    o= countc(Weave, 'O');
    h= countc(Weave, 'H');
    t= countc(Weave, 'T');
    if (g=m=h=t=0) then delete;
run;

data Katy.trip12;
    set Katy.trip12;
    if ((m>0) or (h>0) or (t>0)) and (g>0) then delete;
    if ((m>0) or (h>0) or (t>0)) then type=1;
    if (g>0 and m=h=t=0) then type=2;
    drop g m o h t;
run;

/*Assigning segment travel times*/
Data Katy.trip12;
    set Katy.trip12;
    if ((start ne 1) and (TT_ave_10min_segment=.) ) then
        do;
            if type=1 then
                do;

```

```

        if ((peak1=1) and (peak2=1)) then
            TT_ave_10min_segment=
(segmentlength/53.1)*3600;
        else if ((peak1=0) and (peak2=1)) then
            TT_ave_10min_segment=
(segmentlength/61.8)*3600;
        else if ((peak1=0) and (peak2=0)) then
            TT_ave_10min_segment=
(segmentlength/70.9)*3600;
        end;
    else if type=2 then
        do;
            if ((peak1=1) and (peak2=1)) then
                TT_ave_10min_segment=
(segmentlength/40.8)*3600;
            else if ((peak1=0) and (peak2=1)) then
                TT_ave_10min_segment=
(segmentlength/55.1)*3600;
            else if ((peak1=0) and (peak2=0)) then
                TT_ave_10min_segment=
(segmentlength/65.3)*3600;
            end;
        end;
    if ((start ne 1) and (TT_ave_10min_segment1=.) ) then
        do;
            if type=1 then
                do;
                    if ((peak1=1) and (peak2=1)) then
                        TT_ave_10min_segment1=
(segmentlength1/53.1)*3600;
                    else if ((peak1=0) and (peak2=1)) then
                        TT_ave_10min_segment1=
(segmentlength1/61.8)*3600;
                    else if ((peak1=0) and (peak2=0)) then
                        TT_ave_10min_segment1=
(segmentlength1/70.9)*3600;
                    end;
                else if type=2 then
                    do;
                        if ((peak1=1) and (peak2=1)) then
                            TT_ave_10min_segment1=
(segmentlength1/40.8)*3600;
                        else if ((peak1=0) and (peak2=1)) then
                            TT_ave_10min_segment1=
(segmentlength1/55.1)*3600;
                        else if ((peak1=0) and (peak2=0)) then
                            TT_ave_10min_segment1=
(segmentlength1/65.3)*3600;
                        end;
                    end;
            end;
        run;
Data Katy.trip12;

```

```

set Katy.trip12;
if ((start ne 1) and (TT_std_10min_segment=.) ) then
do;
TT_std_10min_segment= 0;
end;
if ((start ne 1) and (TT_std_10min_segment1=.) ) then
do;
TT_std_10min_segment1= 0;
end;
run;

Data Katy.trip12;
set Katy.trip12;
drop peak1 peak2 type;
run;

/*Calculating standard deviation by summing up all the std of
individual segments*/
Data Katy.trip12;
set Katy.trip12;
if start=1 then
do;
calcstd = 0;
calcstd1= 0;
calcave= 0;
calcave1= 0;

end;

if start=0 then
do;
calcstd + TT_std_10min_segment;
calcstd1 + TT_std_10min_segment1;
calcave + TT_ave_10min_segment;
calcave1 + TT_ave_10min_segment1;

end;
run;

/* PART 6 -----*/
/* 6.1 Creating new dataset with complete trips */
Data Katy.tripend;
set Katy.trip12 (drop = startsensor1 endsensor1 startsensor11
endsensor11 Count_10_segment Count_10_segment1
TT_std_10min_segment TT_std_10min_segment1
TT_ave_10min_segment TT_ave_10min_segment1
segmentlength segmentlength1);

if end=1;
run;

/* 6.2 Computing average travel times and travel time std over the
complete trip*/
data Katy.tripend;
set Katy.tripend;
startdate=datepart(datetime);

```



```

        starttime_sec = timepart(starttime);
        marker10= floor(starttime_sec/600);
        starttime_10min=marker10*600;
run;

Proc sql noprint;
    Create Table Katy.trip12stddev
    as select startdate, start_sensor, end_sensor, starttime_10min,
            sum(count) as Count_10, avg(time) as TT_ave_10min,
            std(time) as TT_std_10min
    from Katy.tripend
    group by startdate, start_sensor, end_sensor, starttime_10min;
quit;

/*Merging data*/
proc sort data= Katy.tripend;
    by startdate start_sensor end_sensor starttime_10min;
run;

Data Katy.tripend;
    merge Katy.tripend Katy.trip12stddev;
    by startdate start_sensor end_sensor starttime_10min;
run;

proc sort data= Katy.tripend;
    by descending randid datetime;
run;

Data Katy.tripend;
    set Katy.tripend;
    if randid=. then delete;
run;

/*Renaming variable to enable merging according to dummy trip*/
Data Katy.trip12stddev;
    set Katy.trip12stddev(rename=(start_sensor=start_sensor1
    end_sensor=end_sensor1 Count_10=Count_101
    TT_ave_10min=TT_ave_10min1 TT_std_10min=TT_std_10min1 ));
run;

proc sort data= Katy.tripend;
    by startdate start_sensor1 end_sensor1 starttime_10min;
run;

Data Katy.tripend;
    merge Katy.tripend Katy.trip12stddev;
    by startdate start_sensor1 end_sensor1 starttime_10min;
run;

proc sort data= Katy.tripend;
    by descending randid datetime;
run;

```

```

/*Deleting observations that do not have corresponding observations
from trip12 dataset*/
Data Katy.tripend;
    set Katy.tripend;
    if randid=. then delete;
run;

/*Renaming variables back to original name*/
Data Katy.trip12stddev;
    set Katy.trip12stddev(rename=(start_sensor1=start_sensor
    end_sensor1=end_sensor Count_101=Count_10
    TT_ave_10min1=TT_ave_10min TT_std_10min1=TT_std_10min));
run;

/* Delete erroneous detections: timedifference = 0 */
Data katy.tripend;
    set katy.tripend;
    if timediff=0 then delete;
run;

/* PART 7 -----*/
/* 7.1 Assigning trip number to each trip of a specific randid */
Data Katy.tripend;
    set Katy.tripend;
    if randid ne lag(randid) then tripno=1;
    if randid=lag(randid) then tripno+1;
run;

/* 7.2 Assigning total trips made by each randid*/
Proc sort data=Katy.tripend;
    by descending obs;
run;

Data Katy.tripend;
    set Katy.tripend;
    retain inter;
    if (randid ne lag(randid)) then totaltrips=tripno;
    if (randid=lag(randid)) then totaltrips=inter;
    inter=totaltrips;
    drop inter;
run;

Proc sort data=Katy.tripend;
    by obs;
run;

/* 7.3 Deleting trips where the trip or the dummy trip goes through the
managed lanes but is not detected at the toll sensors*/
Data Katy.tripend;
    set Katy.tripend;
    if totaltoll=0 and totaltoll1=0 then delete;

```

```

run;

/* 7.3.1 Deleting trips where actual trip and its dummy trip go through
a toll lane - important for note */
Data Katy.tripend;
    set Katy.tripend;
    if totaltoll>0 and totaltoll1>0 then delete;
run;

/* 7.4 Assigning time and std*/
/* time2=actual travel time, time1=dummy travel time,
std=std of actual travel time, std1=std of dummy travel time */
Data Katy.tripend;
    set Katy.tripend;
    if count_10>2 then
        do;
            std=TT_std_10min;
            time2=TT_ave_10min;
        end;
    else do;
        std=6.37 + 0.48*calcstd+2.20*segmentno;
        time2=calcave;
    end;
    if count_101>2 then
        do;
            time1=TT_ave_10min1;
            std1=TT_std_10min1;
        end;
    else do;
        time1=calcave1;
        std1=6.37+ 0.48*calcstd1+2.20*segmentno;
    end;
    if calcstd=0 then std=0;
    if calcstd1=0 then std1=0;
run;

/* 7.5 Determining the trips made during the peak hours and the trips
made during the peak + shoulder hours*/
/* Hour and Weekday */
data katy.tripend;
    set katy.tripend;
    cov=std/time2;
    cov1=std1/time1;
    hour=hour(starttime);
    weekday=weekday(datepart(starttime));
run;

/* Peak (peak1=1 and peak2=1), Shoulder Peak (peak1=0 and peak2=1) and
off peak (peak1=0 and peak2=0) */
/* Before Sep. 08
Peak1= EB:7-9 WB: 17-19,
Peak2= Including shoulder peak EB: 6-10 WB:16-20 */
/* After Sep. 08

```

```

Peak1= EB:7-9 WB: 16-18,
Peak2= Including shoulder peak EB: 6-10 WB:15-19 */
Data katy.tripend;
    set katy.tripend;
    peak1=0;
    peak2=0;
    if ((presep08=1) and (weekday ne 1) and (weekday ne 7)) then
        do;
            if ((direction=0) and (7<=hour<=8)) or ((direction=1) and
                (17<=hour<=18)) then peak1=1;
            if ((direction=0) and (6<=hour<=9)) or ((direction=1) and
                (16<=hour<=19)) then peak2=1;
        end;
    else if ((presep08=0) and (weekday ne 1) and (weekday ne 7)) then
        do;
            if ((direction=0) and (7<=hour<=8)) or ((direction=1) and
                (16<=hour<=17)) then peak1=1;
            if ((direction=0) and (6<=hour<=9)) or ((direction=1) and
                (15<=hour<=18)) then peak2=1;
        end;
run;

/*Eliminating trips with 0 length*/
Data katy.tripend;
    set katy.tripend;
    if ((length=0) or (length1=0)) then delete;
run;

/* 7.6 Additional step to eliminate unusual detection sequence such as
101 followed by 102*/
Data Katy.tripend;
    set Katy.tripend;
    if time=0 or time1=0 then delete;
run;

/* 7.7 Additional step to eliminate records of which travel speed is
greater than 110 mph */
Data katy.tripend;
    set katy.tripend;
    speed=(length/(time/3600));
run;

Data katy.tripend;
    set katy.tripend;
    if speed > 110 then delete;
run;

/* 7.8 Additional step to eliminate records of which travel speed1
(dummy trip speed) is greater than 110 mph */
Data katy.tripend;
    set katy.tripend;
    speed1=(length1/(time1/3600));
run;

```

```

Data katy.tripend;
    set katy.tripend;
    if speed1 > 110 then delete;
run;

/* PART 8 ----- */
/* 8.1 Assigning trip id to each of the trips */
Data Katy.tripend;
    set Katy.tripend;
    id+1;
run;

/* 8.2 Including omitted toll1 to totaltol11 (for dummy trips)*/
/* 8.2.1 Identify Omitted Toll */
/* 8.2.1.1 East bound - direction=0 */
/* 101 omission */
data katy.tripend;
    set katy.tripend;
    if (find(Allsensor, "411")>0 or find(Allsensor, "412")>0 or
        find(Allsensor, "413")>0 or find(Allsensor, "368")>0 or
        find(Allsensor, "427")>0 or find(Allsensor, "449")>0) and
        (find(Allsensor, "443")>0 or find(Allsensor, "466")>0 or
        find(Allsensor, "444")>0 or find(Allsensor, "440")>0 or
        find(Allsensor, "426")>0 or find(Allsensor, ",5")>0
        or find(Allsensor, "271")>0 or find(Allsensor, "468")>0 or
        find(Allsensor, "453")>0 or find(Allsensor, "103")>0 or
        find(Allsensor, "105")>0) and
        (find(Allsensor, "101")=0) and (find(Allsensor, "102")=0) and
        (find(Allsensor, "465")=0) and direction=0
    then case1=101;
run;

/* 103 omission */
data katy.tripend;
    set katy.tripend;
    if (find(Allsensor, "411")>0 or find(Allsensor, "412")>0 or
        find(Allsensor, "413")>0 or find(Allsensor, "368")>0 or
        find(Allsensor, "427")>0 or find(Allsensor, "465")>0 or
        find(Allsensor, "449")>0 or find(Allsensor, "101")>0) and
        (find(Allsensor, "466")>0 or find(Allsensor, "444")>0 or
        find(Allsensor, "440")>0 or find(Allsensor, "426")>0 or
        find(Allsensor, ",5")>0 or find(Allsensor, "271")>0 or
        find(Allsensor, "468")>0 or find(Allsensor, "453")>0 or
        find(Allsensor, "105")>0) and
        (find(Allsensor, "103")=0) and (find(Allsensor, "104")=0) and
        (find(Allsensor, "443")=0) and direction=0
    then case2=103;
run;

/* 105 omission */
data katy.tripend;
    set katy.tripend;

```

```

    if (find(Allsensor, "411")>0 or find(Allsensor, "412")>0 or
        find(Allsensor, "413")>0 or find(Allsensor, "368")>0 or
        find(Allsensor, "427")>0 or find(Allsensor, "465")>0 or
        find(Allsensor, "443")>0 or find(Allsensor, "466")>0 or
        find(Allsensor, "444")>0 or find(Allsensor, "449")>0 or
        find(Allsensor, "468")>0 or find(Allsensor, "453")>0 or
        find(Allsensor, "101")>0 or find(Allsensor, "103")>0) and
        (find(Allsensor, "426")>0 or find(Allsensor, ",5")>0 or
        find(Allsensor, "271")>0) and
        (find(Allsensor, "105")=0) and (find(Allsensor, "106")=0) and
        (find(Allsensor, "440")=0) and direction=0
    then case3=105;
run;

/* 8.2.1.2 West bound - direction=1 */
data katy.tripend;
    set katy.tripend;
    Allsensor2 = '^'!!Allsensor; /* in order to recognize sensor 6 */
run;

/* 107 omission */
data katy.tripend;
    set katy.tripend;
    if (find(Allsensor2, "199")>0 or find(Allsensor2, "272")>0 or
        find(Allsensor2, ",6")>0 or find(Allsensor2, "^6,")>0 or
        find(Allsensor2, "425")>0 or find(Allsensor2, "460")>0) and
        (find(Allsensor2, "445")>0 or find(Allsensor2, "467")>0 or
        find(Allsensor2, "442")>0 or find(Allsensor2, "396")>0 or
        find(Allsensor2, "369")>0 or find(Allsensor2, "414")>0 or
        find(Allsensor2, "415")>0 or find(Allsensor2, "416")>0 or
        find(Allsensor2, "456")>0 or find(Allsensor2, "469")>0 or
        find(Allsensor2, "109")>0 or find(Allsensor2, "111")>0) and
        (find(Allsensor2, "107")=0) and (find(Allsensor2, "108")=0) and
        (find(Allsensor2, "441")=0) and direction=1
    then case1=107;
run;

/* 109 omission */
data katy.tripend;
    set katy.tripend;
    if (find(Allsensor2, "199")>0 or find(Allsensor2, "272")>0 or
        find(Allsensor2, ",6")>0 or find(Allsensor2, "^6,")>0 or
        find(Allsensor2, "425")>0 or find(Allsensor2, "441")>0 or
        find(Allsensor2, "445")>0 or find(Allsensor2, "467")>0 or
        find(Allsensor2, "460")>0 or find(Allsensor2, "456")>0 or
        find(Allsensor2, "469")>0 or find(Allsensor2, "107")>0) and
        (find(Allsensor2, "396")>0 or find(Allsensor2, "369")>0 or
        find(Allsensor2, "414")>0 or find(Allsensor2, "415")>0 or
        find(Allsensor2, "416")>0 or find(Allsensor2, "111")>0) and
        (find(Allsensor2, "109")=0) and (find(Allsensor2, "110")=0) and
        (find(Allsensor2, "442")=0) and direction=1
    then case2=109;
run;

```

```

/* 111 omission */
data katy.tripend;
    set katy.tripend;
    if (find(Allsensor2, "199")>0 or find(Allsensor2, "272")>0 or
        find(Allsensor2, ",6")>0 or find(Allsensor2, "^6,")>0 or
        find(Allsensor2, "425")>0 or find(Allsensor2, "441")>0 or
        find(Allsensor2, "445")>0 or find(Allsensor2, "467")>0 or
        find(Allsensor2, "442")>0 or find(Allsensor2, "460")>0 or
        find(Allsensor2, "456")>0 or find(Allsensor2, "469")>0 or
        find(Allsensor2, "107")>0 or find(Allsensor2, "109")>0) and
        (find(Allsensor2, "369")>0 or find(Allsensor2, "414")>0 or
        find(Allsensor2, "415")>0 or find(Allsensor2, "416")>0) and
        (find(Allsensor2, "111")=0) and (find(Allsensor2, "112")=0)
        and (find(Allsensor2, "396")=0) and direction=1
    then case3=111;
run;

data katy.tripend;
    set katy.tripend;
    drop Allsensor2;
run;

/* 8.2.2 Assign Omitted Toll */
Data katy.tripend;
    set katy.tripend;
    toll_c1=0.0;
    if case1=101 then
        do;
            if peak1=1 and peak2=1 and presep08=1 then
                toll_c1=1.6;
            if peak1=1 and peak2=1 and presep08=0 then
                toll_c1=2.2;
            if peak1=0 and peak2=1 and presep08=1 then
                toll_c1=0.8;
            if peak1=0 and peak2=1 and presep08=0 then
                toll_c1=1.1;
            if peak1=0 and peak2=0 and presep08=1 then
                toll_c1=0.4;
            if peak1=0 and peak2=0 and presep08=0 then
                toll_c1=0.4;
        end;
    else if case1=107 then
        do;
            if peak1=1 and peak2=1 and presep08=1 then
                toll_c1=1.2;
            if peak1=1 and peak2=1 and presep08=0 then
                toll_c1=1.4;
            if peak1=0 and peak2=1 and presep08=1 then
                toll_c1=0.6;
            if peak1=0 and peak2=1 and presep08=0 then
                toll_c1=0.7;
            if peak1=0 and peak2=0 and presep08=1 then

```

```

        toll_c1=0.3;
    if peak1=0 and peak2=0 and presep08=0 then
        toll_c1=0.3;
    end;
run;

Data katy.tripend;
    set katy.tripend;
    toll_c2=0.0;
    if case2=103 or case2=109 then
        do;
            if peak1=1 and peak2=1 and presep08=1 then
                toll_c2=1.2;
            if peak1=1 and peak2=1 and presep08=0 then
                toll_c2=1.4;
            if peak1=0 and peak2=1 and presep08=1 then
                toll_c2=0.6;
            if peak1=0 and peak2=1 and presep08=0 then
                toll_c2=0.7;
            if peak1=0 and peak2=0 and presep08=1 then
                toll_c2=0.3;
            if peak1=0 and peak2=0 and presep08=0 then
                toll_c2=0.3;
        end;
    end;
run;

Data katy.tripend;
    set katy.tripend;
    toll_c3=0.0;
    if case3=111 then
        do;
            if peak1=1 and peak2=1 and presep08=1 then
                toll_c3=1.6;
            if peak1=1 and peak2=1 and presep08=0 then
                toll_c3=2.2;
            if peak1=0 and peak2=1 and presep08=1 then
                toll_c3=0.8;
            if peak1=0 and peak2=1 and presep08=0 then
                toll_c3=1.1;
            if peak1=0 and peak2=0 and presep08=1 then
                toll_c3=0.4;
            if peak1=0 and peak2=0 and presep08=0 then
                toll_c3=0.4;
        end;
    else if case3=105 then
        do;
            if peak1=1 and peak2=1 and presep08=1 then
                toll_c3=1.2;
            if peak1=1 and peak2=1 and presep08=0 then
                toll_c3=1.4;
            if peak1=0 and peak2=1 and presep08=1 then
                toll_c3=0.6;
            if peak1=0 and peak2=1 and presep08=0 then

```



```

        toll_c3=0.7;
    if peak1=0 and peak2=0 and presep08=1 then
        toll_c3=0.3;
    if peak1=0 and peak2=0 and presep08=0 then
        toll_c3=0.3;
    end;
run;

/*8.2.3 Recalculate totaltoll1 for dummy trips */
/* rename totaltoll1 as totaltoll2 */
data Katy.tripend;
    set Katy.tripend (rename=(totaltoll1= totaltoll2));
run;

data Katy.tripend;
    set Katy.tripend;
    totaltoll1=totaltoll2+Toll_C1+Toll_C2+Toll_C3;
run;

/* 8.2.4 Deleting trips where actual trip and its dummy trip go through
a toll lane */
Data Katy.tripend;
    set Katy.tripend;
    if totaltoll>0 and totaltoll1>0 then delete;
run;

/* 8.3 Deleting trips where travelers cannot make choice decision. That
is, travelers have to use GPLs */
/* 466-444-440, 445-467 */
Data Katy.tripend;
    set Katy.tripend;
    if find(allsensor, "466,444")>0 and length > 2.30 and length
        < 2.40 then delete;
    else if find(allsensor, "444,440")>0 and length > 1.60 and length
        < 1.70 then delete;
    else if find(allsensor, "466,440")>0 and length > 4.00 and length
        < 4.10 then delete;
    else if find(allsensor, "466,444,440")>0 and length > 4.00 and
        length < 4.10 then delete;
    else if find(allsensor, "445,467")>0 and length > 2.30 and length
        < 2.40 then delete;
run;

/* 8.4 Include Heavy Rain Variable */
/* If precipitation was 0.4 inches/hour, rain=1, otherwise rain=0 */
/* The startdate and hour indicates the hour when it rained more that
0.4 inches in 2012 */
data katy.tripend;
    set katy.tripend;
    rain=0;
    if (startdate='25JAN2012'd and (10<=hour<=12)) or
        (startdate='31JAN2012'd and (hour=12)) or
        (startdate='03FEB2012'd and (hour=15)) or

```

```

        (startdate='04FEB2012'd and (hour=7)) or
        (startdate='18FEB2012'd and (hour=5)) or
        (startdate='09MAR2012'd and (hour=16)) or
        (startdate='11MAR2012'd and (6<=hour<=7)) or
        (startdate='11MAR2012'd and (hour=11)) or
        (startdate='20MAR2012'd and (hour=7)) or
        (startdate='16APR2012'd and (3<=hour<=4)) or
        (startdate='16APR2012'd and (hour=12)) or
        (startdate='20APR2012'd and (hour=14)) or
        (startdate='12JUN2012'd and (hour=16)) or
        (startdate='30JUN2012'd and (hour=12)) or
        (startdate='12JUL2012'd and (hour=6)) or
        (startdate='15JUL2012'd and (hour=11)) or
        (startdate='18AUG2012'd and (hour=14)) or
        (startdate='19AUG2012'd and (hour=2)) or
        (startdate='31DEC2012'd and (hour=21)) then rain=1;
run;

/* 8.5 Include Trip Frequency Variable=Total Number of Trips in 2012 */
proc sort data=katy.tripend;
    by descending randid datetime;
run;

data katy.tripend;
    set katy.tripend;
    obs= _N_; /* making order */
run;

Data katy.tripend;
    set katy.tripend;
    drop tripno totaltrips;
run;

Data katy.tripend;
    set katy.tripend;
    if randid ne lag(randid) then tripno=1;
    if randid=lag(randid) then tripno+1;
run;

Proc sort data=katy.tripend;
    by descending obs;
run;

Data katy.tripend;
    set katy.tripend;
    retain inter;
    if (randid ne lag(randid)) then totaltrips=tripno;
    if (randid=lag(randid)) then totaltrips=inter;
    inter=totaltrips;
    drop inter;
run;

Proc sort data=katy.tripend;

```

```

        by obs;
run;

/* month and day of the month */
data katy.tripend;
    set katy.tripend;
    month=month(startdate);
    day=day(startdate);
run;

/* 8.6 Include Paid Trip Variable = Number of Paid Trips in 2012 */
/* mltrip=1, paid trip */
/* mltripno = paid trip number */
/* totalmltrips = total paid trips of each traveler */
/* pay_per = mltotaltrips/totaltrips */

/* 8.6.1 Classification of each transponder */
data katy.tripend;
    set katy.tripend;
    pstart=0;
    pend=0;
    if randid ne lag(randid) then pstart=1;
run;

proc sort data=katy.tripend;
    by descending obs;
run;

data katy.tripend;
    set katy.tripend;
    if lag(pstart)=1 or lag(pstart)=. then pend=1;
run;

proc sort data=katy.tripend;
    by obs;
run;

/* 8.6.2 Counting Each Transponder's Paid Trips in 2012 */
Data katy.tripend;
    set katy.tripend;
    mltrip=0;
    if Totaltoll gt 0 then mltrip=1;
run;

Data katy.tripend;
    set katy.tripend;
    retain inter;
    if randid ne lag(randid) then mltripno=mltrip;
    if randid=lag(randid) then mltripno=inter+mltrip;
    inter=mltripno;
    drop inter;
run;

```

```

Proc sort data=katy.tripend;
    by descending obs;
run;

Data katy.tripend;
    set katy.tripend;
    retain inter;
    if (randid ne lag(randid)) then totalmltrips=mltripno;
    if (randid=lag(randid)) then totalmltrips=inter;
    inter=totalmltrips;
    drop inter;
run;

Proc sort data=katy.tripend;
    by obs;
run;
/* 8.6.3 Adjustment of Alternate Trip TT due to difference in Travel
Length */
data katy.tripend;
    set katy.tripend;
    diff_leng=length-length1;
run;

data katy.tripend;
    set katy.tripend (rename=(time1=time1_old));
run;

data katy.tripend;
    set katy.tripend;
    time1=time1_old+(diff_leng/(length1/time1_old));
run;

/*8.7 Converting each line of data into 2 lines for discrete choice
analysys */
Data Katy.tripchoice12 (keep= randid start_sensor end_sensor starttime
    Endtime weave obs allsensor direction time
    segmentno length Totaltoll count_10 std tripno
    mltripno totaltrips totalmltrips lanechoice
    month day weekday hour presep08 peak1 peak2 id
    mainlanes_blocked frontage_lanes_blocked
    ramp_lanes_blocked hov_lanes_blocked
    shoulder_lanes_blocked cov rain);

    set Katy.tripend;
    lanechoice=1;
    output;
    lanechoice=0;
    start_sensor=start_sensor1;
    end_sensor=end_sensor1;
    weave=weave1;
    allsensor=allsensor1;
    time=time1;
    length=length1;
    totaltoll=totaltoll1;

```

```

        count_10=count_101;
        std=std1;
        cov=cov1;
        output;
run;

/* 8.8 Assigning trip type. 1 for ML and 2 for GPL*/
Data Katy.tripchoice12;
    retain obs id randid tripno mltripno totaltrips totalmltrips type
           lanechoice time std Totaltoll count_10 length segmentno
           start_sensor end_sensor weave allsensor direction month
           day weekday hour presep08 peak1 peak2 starttime endtime
           mainlanes_blocked frontage_lanes_blocked
           ramp_lanes_blocked hov_lanes_blocked
           shoulder_lanes_blocked rain cov;
    set Katy.tripchoice12;
    if totaltoll>0 then type=1;
    else type=2;
run;

Proc sort data=Katy.tripchoice12;
    by obs type;
run;

/* 8.9 Add dummy values for ASC_ML and ASC_GPL*/
Data Katy.tripchoice12;
    set Katy.tripchoice12;
    MLdum=0;
    GPLdum=0;
    if type=1 then
        do;
            MLdum=1;
            GPLdum=0;
        end;
    if type=2 then
        do;
            MLdum=0;
            GPLdum=1;
        end;
run;

/* 8.10 Conversion travel time and std of TT to minute unit */
Data Katy.tripchoice12;
    set Katy.tripchoice12;
    time=time/60;
    std=std/60;
run;
/*CODE COMPLETE*/

```