

PROACTIVE AND ROBUST DYNAMIC PRICING STRATEGIES FOR HIGH
OCCUPANCY/TOLL LANES

By

DIMITRA MICHALAKA

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2009

© 2009 Dimitra Michalaka

To my family for supporting me during my studies

ACKNOWLEDGMENTS

I would like to thank my graduate advisor, Dr. Yafeng Yin of the University of Florida for his insights and guidance throughout this thesis and his valuable support. I, also, wish to thank the remaining members of the thesis committee, Dr. Siriphong Lawphonpanich and Dr. Lily Elefteriadou, for their assistance and their advices. Finally, I express my sincere thanks to graduate students, Yingyan Lou for her help with the coding and to Alexandra Kondyli for her valuable comments.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS	4
LIST OF TABLES	7
LIST OF FIGURES	8
ABSTRACT	10
CHAPTER	
1 INTRODUCTION	12
1.1 Background	12
1.2 Problem Statement	13
1.3 Research Objective, Supporting Tasks and Validation	14
1.4 Document Organization	14
2 LITERATURE REVIEW	15
2.1 Introduction	15
2.2 Congestion Pricing	15
2.3 HOT Lanes	18
2.3.1 HOT Lane Concept	18
2.3.2 Benefits of HOT Lanes	19
2.3.2.1 Benefits of HOT lanes on transportation networks	19
2.3.2.2 Benefits of HOT lanes to society	20
2.3.2.3 Implementation difficulties of HOT lanes	21
2.3.3 Implementation of HOT Lanes in the U.S.	21
2.4 Determination of Dynamic Pricing Strategies	22
2.4.1 Modeling Approaches	22
2.4.1.1 Bottleneck models	22
2.4.1.2 Network models	28
2.4.2 Current Practice of Toll Determination in HOT Lanes Operations	30
2.4.2.1 I-394 HOT lanes in Minnesota	30
2.4.2.2 I-95 HOT lane project in South Florida	32
2.4.3 Self-Learning Control Approaches for Dynamic Tolling	33
2.5 Conclusions That Can Be Drawn from the Literature Review	34
3 RESEARCH APPROACH	36
3.1 Introduction	36
3.2 Methodology Overview	36
3.3 Modeling Preparation	37
3.3.1 Willingness-to-Pay Learning	38

3.3.1.1	Basic concept.....	38
3.3.1.2	Calibration of willingness-to-pay	39
3.3.2	Demand Learning.....	42
3.3.3	Stochastic Capacity Determination.....	44
3.4	Robust Toll optimization	45
3.4.1	Deterministic Case	46
3.4.1.1	Modeling traffic dynamics.....	46
3.4.1.2	Model formulation.....	48
3.4.2	Scenario-based optimization	53
4	DEMAND LEARNING RESULTS	57
4.1	Introduction.....	57
4.2	Empirical Analysis.....	57
4.2.1	Analysis Using the Data of the Entire Period	58
4.2.2	Analysis Using the Data of the Peak Hour.....	64
5	PROACTIVE AND ROBUST PRICING STRATEGIES	69
5.1	Introduction.....	69
5.2	Model Solution and Simulation Study	69
5.3	Numerical Results	71
5.3.1	Low-High-Low Demand Case	71
5.3.1.1	Robust toll optimization	71
5.3.1.2	Deterministic toll optimization.....	75
5.3.1.3	Comparison of the robust versus the deterministic toll optimization	78
5.3.2	Low-Medium-High-Medium-Low demand case	82
5.3.2.1	Robust toll optimization	82
5.3.2.2	Deterministic toll optimization.....	85
5.3.2.3	Comparison of the robust versus the deterministic toll optimization	89
5.4	Conclusions.....	94
	LIST OF REFERENCES.....	95
	BIOGRAPHICAL SKETCH	99

LIST OF TABLES

<u>Table</u>	<u>page</u>
4-1 Data from SR-91	58
4-2 Estimated flows.....	61
4-3 Actual and predicted flows	62
4-4 Data from SR-91	65
4-5 Estimated flows.....	66
4-6 Actual and predicted flows	67

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
2-1 Optimal congestion toll (Morrison,1986).	16
2-2 Diamond lane design (Scott, 2007).....	31
2-3 I-394 toll rates on May 18, 2005 (source: www.MnPASS.net).....	32
3-1 Toll determination procedure.....	37
3-2 System configuration	39
3-3 Representation of the road with cells.....	46
3-4 Flow-density relationship.....	47
3-5 A loss function (Yin, 2007)	54
4-1 SR-91	57
4.2 Actual flows and weaker bounds.	64
4-3 Actual flows and weaker bounds during the peak hour.	68
5-1 Flow-Density Curves	70
5-2 Traffic demand profile.	72
5-3 Optimal Toll Rates.....	72
5-4 Freeway, HOT and GP throughputs.....	73
5-5 Average densities along HOT and GP lanes.....	74
5-6 Queue length upstream of the bottleneck at HOT and GP lane.....	75
5-7 Optimal Toll Rates.....	76
5-8 Freeway, HOT and GP throughputs.....	76
5-9 Average densities along HOT and GP lanes.....	77
5-10 Queue length upstream of the bottleneck at HOT and GP lane.....	78
5-11 Optimal Toll Rates.....	79
5-12 HOT throughput.....	80

5-13	Average density along HOT lane.....	80
5-14	Queue length upstream of the bottleneck at HOT lane.....	81
5-15	Traffic demand profile.....	82
5-16	Optimal Toll Rates.....	83
5-17	Freeway, HOT and GP throughputs.....	83
5-18	Average densities along HOT and GP lanes.....	84
5-19	Queue length upstream of the bottleneck at HOT and GP lane.....	85
5-20	Optimal Toll Rates.....	86
5-21	Freeway, HOT and GP throughputs.....	87
5-22	Average densities along HOT and GP lanes.....	88
5-23	Queue length upstream of the bottleneck at HOT and GP lane.....	89
5-24	Optimal Toll Rates.....	90
5-25	HOT throughput.....	91
5-26	Average density along HOT lane.....	92
5-27	Queue length upstream of the bottleneck at HOT lane.....	93

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

PROACTIVE AND ROBUST DYNAMIC PRICING STRATEGIES FOR HIGH
OCCUPANCY/ TOLL LANES

By

Dimitra Michalaka

May 2009

Chair: Yafeng Yin

Major: Civil Engineering

Congestion pricing is to reduce congestion in transportation infrastructure by charging motorists a certain amount of money—known as a toll—for the use of the roads. Congestion pricing has been promoted by economists and transportation researchers as one of the most efficient means to mitigate traffic congestion because it employs the price mechanism with almost all the advantages of efficiency, universality and clarity. When tolls implemented on highway lanes vary by the time of day, with higher values charged during peak traffic periods, it is called as dynamic tolling. The tolled lanes are High Occupancy/ Toll Lanes (HOT) if the high occupancy vehicles are allowed to use the lanes toll-free.

As the literature review indicates, many studies have been conducted to determine optimal dynamic tolls than can be implemented to roads with high congestion levels. However, most of these studies take into consideration idealized and hypothetical situations in order to derive solutions. For instance, the travel demand is assumed to be known as well as motorists' willingness to pay, i.e., how much money they are likely to pay for using the managed facility. In addition, there is not any model that takes into consideration uncertainty of demand or capacity for the determination of the toll values. Therefore, this thesis develops a more robust and proactive approach to determine time-varying tolls for HOT lanes in response to real-time traffic

conditions. The toll rates are optimized to provide free-flow conditions to managed lanes while maximizing freeway's throughput. The approach consists of several key components, including demand learning and scenario-based robust toll optimization. Simulation experiments are conducted to validate and demonstrate the proposed approach.

CHAPTER 1 INTRODUCTION

1.1 Background

Traffic congestion has become one of the most important problems of modern life due to the increase in movement of both people and goods on highway facilities. People are demanding more goods and are engaging in more activities. Congestion levels have risen in cities of all sizes since 1982, indicating that even the smaller areas are not able to keep pace with the rising demand (FHWA, 2004). According to Texas Transportation Institute (2002), the average urban driver spent 62 hours sitting in traffic in 2000, compared to 16 in 1982, which is an increase of 288%. Furthermore, in 75 urban areas in 2000, rush hours lasted longer and were more extensive than the previous year and cost the country \$68 billion a year. These costs were due to 3.6 billion hours of delay and 5.7 billion gallons of wasted fuel. More than half of major roads are crowded during rush hour, up from a third in 1982. Two out of every five urban interstate miles are congested with traffic at volumes that result in significant delays. The proportion of urban interstate miles that are considered congested increased from 33 to 41 percent from 1996 to 2001 (FHWA, 2003).

As the travel demand increases, existing roads become congested and solutions are required to manage the flow of traffic. The construction of new facilities is very difficult due to environmental constraints and the limited public funding. This has led transportation agencies to explore other alternatives to manage traffic flow. Typically, this has been done by using lane management strategies that regulate demand, separate traffic streams to reduce turbulence, and utilize available and unused capacity on existing transportation facilities. In recent years, such operational policies are called “managed lanes”. The Federal Highway Administration (FHWA) defines managed lanes as highway facilities or a set of lanes in which operational strategies are

implemented and managed in real time in response to changing conditions. Managed lanes include high-occupancy vehicle (HOV) lanes, high-occupancy/toll (HOT) lanes, priced and special use lanes such as express, bus-only, or truck-only lanes (Obenberger, 2004). HOT lanes are facilities that combine pricing and vehicle eligibility to maintain free-flow conditions while maximizing the freeway's throughput. People driving alone in a vehicle can buy their way into the HOT lanes while the HOVs can use the HOT lanes for free. HOV and special use lanes have been used for decades while HOT lanes are much newer innovation. The first HOT express lanes opened at California State Route 91 in December 1995. Currently, there are six HOT lanes in the U.S.

1.2 Problem Statement

Congestion pricing has been promoted by economists and transportation researchers as one of the most efficient means to mitigate congestion in highway facilities and in urban areas because it employs the price mechanism with almost all the advantages of efficiency, universality and clarity. Although the basic theory to congestion pricing has been essentially unchanged since Pigou (1920) and Knight (1924) who were the first to analyze this policy, many extensions have been developed over the years.

As the literature review indicates, many studies have been conducted to determine optimal dynamic tolls than can be implemented to roads with high congestion levels. However, most of these studies take into consideration idealized and hypothetical situations in order to derive solutions. For instance, in many papers the travel demand is assumed to be known and do not consider how much money users are likely to pay for using the managed facility. In addition, no model takes into consideration uncertainty of demand or capacity for the determination of the toll values. Therefore, there is a need to develop a more proactive approach that will give appropriate

toll rates with respect to incoming flow to maintain free-flow conditions on managed lane while maximizing road's throughput.

1.3 Research Objective, Supporting Tasks and Validation

The objective of this research is to develop a procedure that determines time-varying tolls for HOT lanes in response to real-time traffic conditions. The toll values must be appropriate in order to provide free-flow conditions for the users of the managed lanes while maximizing the freeway's throughput.

The tasks that will be conducted to achieve the objective are as follows:

- Review thoroughly the literature to identify existing methods and procedures for managed lanes.
- Learn the real-time demand based on Bayesian inference.
- Use the California SR-91 data for the empirical investigation of demand learning.
- Develop a scenario-based robust toll optimization model.
- Validate the proposed approach via simulation.

1.4 Document Organization

Chapter 2 presents the literature review on managed lanes, dynamic tolling and the methods and procedures used previously to determine tolls. Chapter 3 describes the research approach including willingness-to-pay learning, real-time demand learning, stochastic capacity determination and robust toll optimization. Chapter 4 reports some results of demand learning by using empirical data from SR-91 and Chapter 5 presents robust and proactive pricing strategies for different demand scenarios.

CHAPTER 2 LITERATURE REVIEW

2.1 Introduction

This chapter reviews current practice of HOT lane operations and existing methodologies for determining dynamic pricing strategies.

2.2 Congestion Pricing

Road pricing is not a new innovation. The first toll roads in the United States and in many other countries date back to the late 18th century. In these cases, toll rates were fixed and everybody paid the same price for using the same road. The purpose of tolling was to recover the construction cost or simply to raise revenue. In the early 1920s, economists and transportation researchers started to consider tolling as a policy measure to manage demand and address congestion that has started to increase in many cities throughout the world (Morrison, 1986).

The basic theoretical approach to congestion pricing has been developed by Pigou (1920) and Knight (1924) who were the first persons who analyzed this policy as a measure to alleviate congestion. Morrison (1986) further developed the theory of optimal congestion tolls based on the works of Pigou (1920) and Knight (1924). He used the speed-flow curve that is very important in highway engineering to derive the relationship between flow and cost per user. The most significant effect of congestion was considered as the cost associated with increased trip times with an assumed value of travel time. If speed is inverted at the speed-flow curve, time per mile is obtained. Multiplying the value of time by the time per mile and adding operational vehicle costs gives the cost relationship that is the average variable cost (AVC). The extra cost of adding a vehicle to the flow is the short-run marginal cost (SRMC). The cost curves as well as a demand curve, that represents the willingness to pay for various quantities of trips, are shown in figure 2-1. In this figure the backward bending portion of the AVC curve is not illustrated

because the optimal flow will never occur in this region due to the fact that the same flow can be achieved at a lower cost.

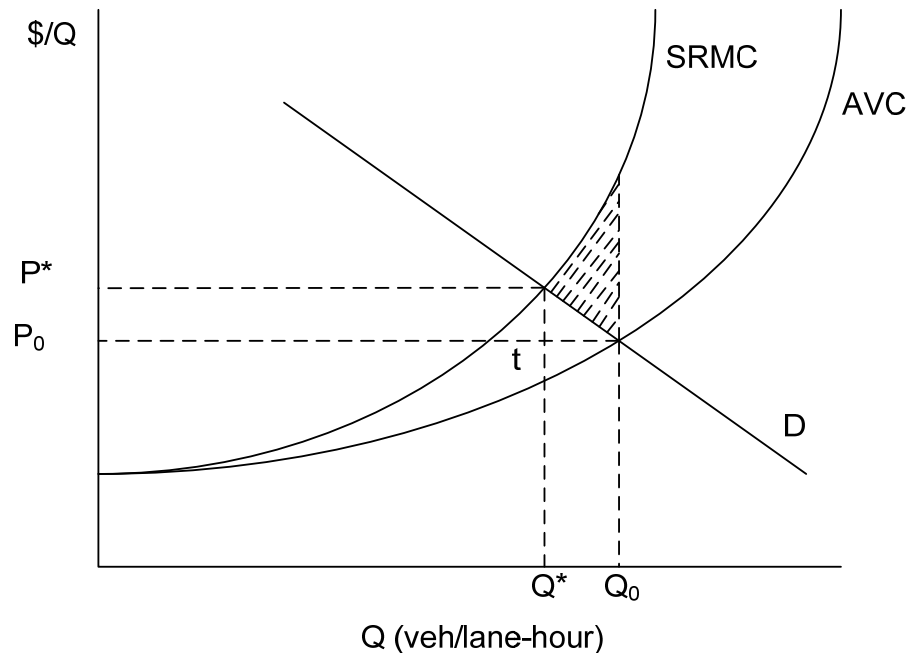


Figure 2-1. Optimal congestion toll (Morrison,1986).

If there is not any toll (t), equilibrium will occur at Q_0 which is the intersection point of demand and AVC curves. In this case, users who value the trip more than they can afford, travel. At this point, the additional cost to society from, for instance, considering other users exceeds the benefit derived by the last traveler. This is happening for all trips beyond point Q^* . The amount by which the extra cost of these $Q_0 - Q^*$ exceeds the additional benefits represents the welfare loss from non-optimal pricing. In order to have equilibrium a toll, t , must be added to optimal quantity Q^* . This toll is equal to congestion externality, that is, the difference between the cost a traveler affords (AVC) and the cost s/ he imposes on society (SRMC). In other words, externality is the congestion cost that each additional user of a congested road or other facility imposes on other users, by slowing down, increasing inconvenience, risk of accidents, etc. (Carey and Srinivasan, 1993).

Kraus et al (1976) and Keeler and Small (1977) further analyzed in two independent papers the congestion problem in a long-run perspective. They argue that, in a long-run analysis, optimal tolls depend on highway capacity cost due to the fact that optimal highway capacity and congestion depend on the cost of the additional capacity. Kraus et al estimated long-run peak tolls using “pseudo-empirical” analysis for a generic U.S. urban expressway. They found that tolls vary according to the capital cost and to location of the expressway. Keeler and Small (1977) conducted their research taking into consideration speed-flow relationships for uninterrupted flow conditions and highway construction costs for the San Francisco Bay Area and their results were that peak tolls vary very much among the types of roads.

Congestion pricing or value pricing—as it is sometimes called—is a tool for mitigating traffic congestion because it has been observed that people tend to make socially efficient choices when they are faced with the costs of their actions and the social benefits (Lindsey and Verhoef, 2000). Congestion pricing leads the rush-hour travel to shift to off-peak periods or to other transportation modes. By removing even a small percentage such as 5% of the vehicles that travel during the peak periods from a congested facility by implementing pricing, the system performs much better (FHWA, 2006). Congestion pricing entails setting tolls depending on how severe the congestion is. This implies that tolls must vary according to time, location, current circumstances such as bad weather, accidents and special events, type of vehicle and occupancy of the vehicle. Congestion pricing except for traffic alleviation is used also in other sections of the economy such as to telephone rates, air fares, electricity rates, hotels and other public utilities. FHWA (2006) refers four types of variable pricing strategies. The first involves priced lanes with variable tolls on separated lanes within a highway, such as express toll lanes and HOT lanes while the second includes variable tolls on entire roadways, that is, on toll roads, bridges,

and existing toll-free roads during rush hours. The third pricing strategy is cordon charges to drive into a congested area within a city and the fourth is the area-wide charges which are charges per mile on all roads within an area that vary by level of congestion. Our literature review focuses on HOT lanes operations and associated pricing strategies.

2.3 HOT Lanes

In this section the general principles of HOT lanes including definition, purpose, objectives, motivation and the current practice are summarized.

2.3.1 HOT Lane Concept

HOT lanes are facilities that combine pricing and vehicle eligibility to maintain free-flow conditions while still providing a travel time-saving incentive for high-occupancy vehicles (Obenberger, 2004). This allows additional HOV lane capacity to be used while acting as a stimulant for mode shifting.

HOT lanes were first advocated by those who believed that congestion pricing can reduce the congestion levels on freeways because drivers have to pay an amount of money in order to use the congested facility. Advocates support HOT lanes as first step toward more widespread pricing of congested roads (Dahlgren, 1999).

Dahlgren (2002) investigated when to implement HOT, HOV and General Purpose (GP) lanes, and suggested that HOT lane seemed to perform as well as or even better than an HOV lane in any circumstance. More specifically, HOT lanes may offer a solution to the issue of under-utilization of HOV lanes.

The concept of HOT lanes combines two very effective highway management tools: lane management and value pricing. More precisely, lane management includes limited-access to designated highway lanes that depend on the vehicle's occupancy and type, and other objectives. The desirable level of traffic service is maintained by limiting the number of vehicles on the

designated lanes. Lane management can promote a range of policies such as car pools and transit vehicles to encourage higher occupancy or low emission vehicles to improve air quality or vehicles equipped for electronic toll collection to improve operational efficiency (FHWA, 2006). Dynamic tolling includes the introduction of road user charges that vary over the time of day and according to congestion level. During the peak periods where the volumes are high, even the shift of a small number of vehicles can reduce significantly the overall congestion levels and to lead to more reliable travel times (FHWA, 2006).

2.3.2 Benefits of HOT Lanes

HOT lanes are an important management tool with a variety of benefits to transportation networks and to society.

2.3.2.1 Benefits of HOT lanes on transportation networks

The implementation of HOT lanes is appealing because it has many advantages to individuals in a variety of traffic conditions on road facilities. First of all, the HOT lanes can ensure free flow conditions due to management of the demand during the periods of peak congestion (FHWA, 2006) and mitigate congestion by giving drivers a financial incentive to seek alternative modes of transportation, such as carpooling and transit or to drive during off-peak hours (Replogle, 2006). This keeps the practical traffic capacity of these lanes from shrinking as usually occurs in periods of high congestion. That is, managing traffic with tolls adds new road capacity without building more lanes (Replogle, 2006). Secondly, the congestion pricing can lead to savings in the total travel time compared to the situation without tolls and can minimize the total travel time in the network (Jaksimovic et al., 2005). Thirdly, tolling lanes offer travel options for saving time and enhance the reliability of travel times (Obenberger, 2004). In addition, dynamic tolling improves transit speeds and the reliability of transit service and prevents the loss of vehicle throughput that comes from a breakdown of traffic flow (FHWA,

2006). Toll lanes often reduce traffic in the free lanes. This happens especially in cases where dynamic toll lanes are created by converting HOV lanes that are underused (Replogle, 2006). HOT lanes can, also, reduce turbulence among the vehicles because they separate traffic streams (Sisiopiku and Sullivan, 2007). Moreover, lanes with tolls improve safety on the road and the performance of freeways (Obenberger, 2004). Another important benefit of congestion tolling is that it can take into consideration not only congestion at traffic rush hour but also congestion that is caused by special events such as accidents, sport events, parades, construction, maintenance and severe weather conditions (hurricanes, floods, etc) (Halem et al, 2007).

2.3.2.2 Benefits of HOT lanes to society

HOT lane operations benefit drivers by reducing delays and stress and by increasing the predictability of trip times. Pricing roads guarantees vehicles a reliable speed and reduced travel time in comparison with roads without toll. Also, the HOT lanes can help sellers to deliver more products per hour to the market (FHWA, 2006) and leads to reductions in money lost when sitting in traffic (Replogle, 2006). Additionally, HOT lanes benefit the society as a whole due to reduction of fuel consumption and vehicle emissions and to allowance of more efficient land use decisions (FHWA, 2006). New toll lanes created from existing capacity do not have negative impact on the environment and if the tolls are appropriate they lead to less air pollution. This has been proved from studies of the toll roads in U.S.A. and the cordon pricing program in London. Another significant benefit of the implementation of tolling to congested roads is the improvement of the quality of transportation services without increase of taxes or large capital expenditures by providing additional revenues for funding transportation (FHWA, 2006).

2.3.2.3 Implementation difficulties of HOT lanes

The operation of HOT lanes has many advantages to both motorists and society. However it also has some disadvantages with respect to their implementation. Many transportation agencies recognize that there is a risk associated with inaccurate traffic forecasts on toll roads. If inaccurate traffic models are used in predicting the tolls, the expectations of toll roads will not be met (Replogle, 2006). Traffic forecasts that were not accurate were used, for example, at Dulles Greenway in Virginia (U.S. DOT, 2006). In addition, the application of tolling to highway lanes requires funding from the government or companies for the toll collection infrastructure. Moreover, although the payment of a toll is not actually a cost but an economic transfer from travelers to the toll authority, consumers feel as having an extra financial cost for paying the toll (Smith, 2007). How much consumers are affected by this transfer depends on the value of the time savings and the redistributed revenues.

2.3.3 Implementation of HOT Lanes in the U.S.

Currently, six HOT lanes (more precisely, managed lanes) are in operation in the United States:

- 1) The State Route 91 in Orange County California opened in December 1995. It is a four lane toll facility in the median of one of the most congested highways in the U.S. The tolls vary over the day to ensure that the toll lanes operate with free flow conditions. Priced lanes are separated with plastic pylons from free lanes on SR-91. Sullivan (2000) conducted a study to evaluate the impacts of the value pricing on SR-91. Some interesting results from his study are that the toll lanes attracted a substantial share of the traffic using the corridor SR-91, the congestion in the free lanes and the new trips induced by better travel conditions were for non-work purposes. Furthermore, the number of accidents in the express lanes decreased significantly after the implementation of the tolling.
- 2) I-15 lanes at San Diego opened in December 1996. I-15 lanes were initially HOV lanes and then a pricing program was implemented for solo drivers. Single-occupancy vehicles were allowed to use these lanes after the payment of a toll. Tolling lanes are separated with a barrier from general purpose lanes. Tolls in this case vary dynamically according to traffic conditions in the HOV lanes (FHWA, 2006). The express lanes are better utilized and the number of HOVs increased and users' travel times are reduced (Smith, 2007).

- 3) In Houston, tolls were introduced for the two-person carpoolers on existing HOV lanes at corridor I-10 in January 1998. In November 2000, tolling was implemented and in US 290 HOV lanes in Houston. The average number of trips on the HOT lanes increased and the main source of the travelers on toll lanes were those who used to travel in SOVs in regular lanes (FHWA, 2006).
- 4) In May 2005, the MnPass program converted the HOV lanes on I-394 at Minneapolis into HOT lanes. The tolls vary dynamically and change value even every three minutes depending on the traffic density.
- 5) In June 2006 opened the I-25/US-36 managed lanes in Denver. Solo drivers have to pay a toll to use these lanes while carpools, buses, and motorcycles not. The usage of this facility increased after the first month of its operation more than 46% (Colorado Department of Transportation, 2006).
- 6) Recently, in August 2008, the 95 Express Program converted the HOV lanes on I-95 at Miami into HOT lanes. There are two HOT lanes separated from the local traffic lanes with pylons. The tolls depend on the traffic conditions and they are expected to be between \$0.25 and \$2.65 during the average traffic conditions. During the peak hours, tolls can be up to \$6.20 in order to provide operating speeds between 45 and 50 mph at the HOT lanes.

More managed lanes are planned in Miami, Washington D.C., Northern Virginia, Seattle, Maryland, Austin, Dallas, Atlanta, the San Francisco Bay Area, Raleigh-Durham and Portland.

2.4 Determination of Dynamic Pricing Strategies

In this section, we summarize the models that have been developed over the years for the determination of dynamic pricing strategies and the current practice for the HOT lanes.

2.4.1 Modeling Approaches

This section discusses the methodologies that have been developed for determining dynamic pricing strategies. More specifically, it focuses on bottleneck and network models.

2.4.1.1 Bottleneck models

The bottleneck model was introduced by Vickrey (1969) and further developed by Arnott, de Palma and Lindsey (1993). It focuses on the time at which motorists want to travel and the principal innovation of it was to homogenize travelers' departure times. Motorists in this model travel along a single road with a bottleneck or bottlenecks downstream of certain flow capacity.

Bottleneck model doesn't take into consideration route choice because there is one single route every time.

Vickrey (1969) was the first to consider trip costs with respect on desired arrival times (i.e. travelers arriving later or earlier than the desired time experience not only their travel time costs but also schedule costs). The simple bottleneck model is dynamic, deriving the time pattern of congestion over the peak hour (Arnott et al, 1995; Arnott, 1998). It assumes that every morning a fixed number, N , of individuals living in suburbs would like to travel from home, which is the origin, O , to work, which is the destination, D , at the same time t . Each person travels by his/her own car along a single road connecting the origin and destination which has a bottleneck downstream. Traffic conditions are uncongested except at the bottleneck that has a deterministic capacity of s cars per unit time. Due to bottleneck's capacity, it is not possible for all the motorists to be at the same time at their destination. As a result, travelers arrive at work different times, some early and others late that entails a cost of delay. Furthermore, travelers incur expenses including a fixed component which can be set equal to zero for computational simplifications and a variable component which depends on the time spent waiting at the bottleneck. The bottleneck model addressed individuals' decisions with regard to the time depart from their homes. The basic insight is that the total trip's cost including the travel cost, the delay cost and the toll must be constant over the departure interval under equilibrium. For analysis simplification, in the model total trip cost is assumed linear in its components:

$$C_t = \alpha(\text{queuing time}) + \beta(\text{time early}) + \gamma(\text{time late}) + (\text{toll})$$

where,

C_t = the trip's cost when an individual arrives at time t

α = is the shadow value of queuing time

β = is the shadow value of time early

γ = is the shadow value of time late

It is assumed that $\alpha > \beta$ because all the individuals have the some official starting time, t^* (Small, 1982). Let t_f denote the time at which the first traveler arrives to work, t_l the time at which the last traveler arrives to work and $T(t)$ the variable travel time. During the peak hour the capacity throughput of the bottleneck must be used because in any other case, a traveler could depart in the interior of the peak hour having zero queuing cost and less delay cost than either the first or the last person to arrive for the equilibrium consistency. This implies that $t_l - t_f = N / s$, the first individual to arrive, he/ she doesn't face queue and experiences only delay cost equal to $\beta(t^* - t_f)$ and the last individual to arrive who, also, doesn't face queue experiences only delay cost equal to $\gamma(t_l - t^*)$. Under equilibrium, delay costs of the first and last person to arrive must be equal so the equilibrium price, \bar{p} , is equal to $\frac{\beta\gamma}{(\beta + \gamma)} N / s$.

Therefore, the travel cost function without toll is $c(N) = \frac{\beta\gamma}{s(\beta + \gamma)} N$.

If a dynamic toll is introduced that equals the queuing cost component in the equilibrium without toll, the queue will be eliminated without changing the rush-hour interval. In this case, every traveler has the same trip cost as before (equilibrium without toll) and as the trip costs have become equal, no one wants to change his/ her behavior. Imposing this toll the social optimum is decentralized, the delay cost is minimized because all travelers face the same delay cost, the queuing costs are eliminated and the bottleneck is used to capacity during the peak hour.

The amount of travel doesn't change while the total social cost is reduced by reallocating the traffic over the peak hour.

The bottleneck model presented above is simple and theoretical so many extensions are needed in order to make it more realistic. Over the years, many studies, considering elastic demand, heterogeneous individuals, stochastic capacity and demand, simple networks and alternative congestion treatments, have been conducted to improve the simple bottleneck model.

Arnott et al (1993) used, also, the 'bottleneck' approach to determine the time-varying tolls. They assumed a high number of individuals who travel on a network and want to be at work at the same time. However, Arnott et al (1993) also assumed that there is a bottleneck somewhere in the network with limited capacity so not all the travelers can reach destination on the desired time. In this model, the social optimum and the distribution of travel delays, the scheduling costs at the peak period, and the duration of the peak are determined endogenously. The optimal toll depends on time and has its maximum value when drivers arriving at the desired arrival time. The most important thing in this approach is that private costs of the use of the road which include the toll, the travel time cost and the delay cost, should be constant over the peak period.

Iryo and Kuwahara (2000) consider travelers that can choose their departure times for the minimization of their travel cost that includes the queuing delay on a bottleneck and schedule delay at the destination. They used a mathematical model to analyze the travelers' decision assuming one bottleneck between a residential area and a working area that must be used from all the commuters with constant capacity and FIFO service. At first, they derived the model without considering congestion toll and then they applied it to evaluate a dynamic road pricing. The purpose of Iryo and Kuwahara (2000) was to create a tool that can evaluate policies that

have been proposed to mitigate congestion such as Traffic Demand Management (TDM) policies. Specifically, Iryo and Kuwahara (2000) considered a policy that disperses travel demand over time because individual variation in time is very important when a road pricing scheme is analyzed. Their conclusions after the application of the method to road pricing were that dynamic congestion tolls that reduce the waiting time are not proportional to waiting time without the existence of the tolls. Moreover, commuters are likely to change their departure times and that can cause different travel costs for them. This case will not be true if all travelers have the same willingness-to-pay. Finally, Iryo and Kuwahara (2000) concluded that although the individual variation of time values, there is a dynamic congestion toll which can reduce queuing delay to zero. In this case, toll changes the travelers' behavior and their costs.

Although the single bottleneck model that Vickrey (1969) used gives a good insight into the travel times, the optimal toll and its benefits, it has an important deficiency. It doesn't consider the spatial extent of queues which is a significant aspect in the analysis of extended networks because it gives more realistic patterns to avoid congestion. Yperman et al (2005) followed the same procedure of Vickrey (1969) but replaced the simple bottleneck model by a traffic flow model in order to take into consideration the road space that is occupied by the queues. Specifically, they used the LWR traffic flow model which considers the spatial extent of queues and in the same time is not a very complicated model. The LWR model assumes that traffic is behaving as in a kinematic wave model. Yperman et al (2005) used a simple multi-destination network in their study and tried to determine the advantages of congestion pricing and understanding the congestion's mechanisms. They used the traffic demand from Vickrey's bottleneck model and determined user equilibrium and system optimum network conditions using the LWR traffic model. After the analysis, they concluded that congestion can be mitigated

if an optimal toll is imposed and that the benefits of introducing this time-varying optimal toll are higher than those expected by conventional point-queuing bottleneck models. The toll must be equal to the delay costs that commuters would experience if there is not any toll. After the imposing of the toll, commuters that want to travel through the bottleneck have the same travel costs as in the case without toll but their travel time is reduced while the commuters that don't want to use the bottleneck but experience queue that spill over from the bottleneck have reduced trip costs. Thus, traveling demand will increase without increase in congestion. Therefore, optimal tolling can lead to reduced trip costs for the travelers, to more efficient use of the transportation network and even to extra revenues for the government.

Verhoef (1997) considered a dynamic model of road congestion for the determination of time-varying tolls. The elastic demand for the morning peak road usage and a congestion technology used being 'flow congestion' and not 'bottleneck congestion' but the model is based on the bottleneck approach. Such elasticity of demand could come from the presence of different transport modes. In this case, the optimal time-varying toll should include a time-invariant component when drivers share the same desired arrival time. This is very important in reality because it means that the regulator should have information of the distribution of travelers' desired arrival times in order to set the optimal tolls, because the underlying reason of the time-invariant component is the assumption that desired travel times are equal among the users. This time-invariant component is relevant only in studies of road traffic congestion with elastic demand. In this approach, the optimal toll is not zero even in the case that congestion, in terms of travel time delays, has gone to zero.

Although bottleneck models taking into consideration real-world complications give a good insight about the tolls to introduced for mitigating congestion levels they don't incorporate

route choice and they always consider a bottleneck at the road. Therefore, bottleneck models cannot be applied to large networks. That led researchers to develop models called network models that can include more parameters with regard to individual choices and can be used to determine pricing strategies to networks.

2.4.1.2 Network models

In the recent literature, in traffic engineering and transport economics, network models, also, have been examined to find policies to alleviate the problem of congestion that is present in the most of the transportation facilities. Network models in contrast with bottleneck models encompass the mode, departure time and route choice and longer-run choices. The traffic models must be as realistic as possible in order to derive logical and effective policies. In the literature, there are many types of network models, some of them consider fixed departure times, others variable demand, others many alternative modes, others destination choice or route choice and others combination of the factors mentioned. Some of them are presented in the next paragraphs.

Palma and Lindsey (1997) focused on the efficiency of use of private toll roads assuming a simple network with two parallel routes that can have different free-flow travel times and capacity, connecting one origin and one destination. In addition, they assumed that congestion takes the form of queuing and that every traveler has three options: whether to travel by car, and if choose to drive to decide the route and the time of traveling. For the analysis, Palma and Lindsey (1997) considered three cases, firstly, a free access on the one of the routes and a private road with tolls on the other, secondly, private roads on both routes and thirdly a public road with tolls competing with a private road. In each case, they measured the efficiency gain by determining the social surplus relatively to efficiency gain if both routes had first-best optimal tolls. The conclusion of the study was that the efficiency gain is much higher if tolls are imposed dynamically than using a fixed toll in each assumed case.

Time-dependent tolls on a general network are determined by Joksinovic et al (2005a) using a dynamic traffic model to describe the network performance. Joksinovic et al determined the time-varying road prices that minimize the total travel time in the network, taking into account the time changes of the route and departure as a response to the prices with the formulation of a network design problem. For the analysis, they considered dynamic traffic flows, dynamic road pricing strategies; they formatted the problem as a Mathematical Programming with Equilibrium Constraints and analyzed a small and simple network. In their research, they demonstrated that dynamic pricing can lead to savings in the total travel time in the network. Finally, they concluded that it is very difficult to find any simple solution to the dynamic toll design problem in real size dynamic traffic networks because the objective function is non-linear and non convex so it is hard to find a global optimal solution, that is, the optimal toll values and they suggested that in order to find a global solution to large networks, a global search algorithm should be developed. That was the reason that they considered a small hypothetical network in order to analyze the uniform and varying pricing.

Joksinovic et al (2005b) considered elastic demand and applied second-best tolling scenarios only to a subset of links on the network. They used the same methodology with Joksinovic et al (2005a) in order to determine the optimal toll.

Friesz et al (2006) introduced a dynamic optimal toll problem with user equilibrium constraints (DOTPEC). They further presented and tested two algorithms for solving the optimal control representations of the DOTPEC. Firstly, Friesz et al (2006) studied a dynamic efficient toll problem by employing a form of dynamic user equilibrium model to compare the efficient tolls with DOTPEC which is not equivalent to dynamic generalization of the static efficient toll problem. Then, they formulated DOTPEC with two different ways and solved it using both a

descent in Hilbert space without time discretization and a finite dimensional approximation solved as a nonlinear program algorithm. The mathematical representation, in their approach, is detailed enough to capture the behavioral and technological considerations about dynamic tolling and it is referred as a computable theory.

The studies about network models for pricing strategies mentioned are some among a large number. Nowadays, in each tolling facility different models are used for determining the tolls.

2.4.2 Current Practice of Toll Determination in HOT Lanes Operations

Currently, the approaches are used to determine dynamic tolls depend on the availability of data, modeling software, model structure and objective of the study. In the following, the current operation practices on Interstate I-394 and Interstate I-95 will be described as an example of how tolls are determined on HOT lanes.

2.4.2.1 I-394 HOT lanes in Minnesota

The I-394 is an east-west oriented facility that links I-94 on the east end to I-494 and downtown Minneapolis and the Twin Cities western suburbs. It runs for 11 miles and has three lanes in each direction. It has 3 miles reversible lanes separated with barriers and 8 miles of diamond lanes. The diamond lanes are designed as in Figure 2-2. The facility opened with HOV and general-purpose lanes in 1992 but significant congestion of GP lanes and the underutilization of the HOV lanes led to requests that the MnDOT open HOV lanes to solo drivers. The transportation service in Minnesota called MnPASS conducted a major study, which was completed in 2001, that concluded that opening the HOV lanes to solo drivers would lead to a congested facility so the conversion of these lanes to HOT lanes will be the most effective solution (Halvorson and Buckeye, 2006).

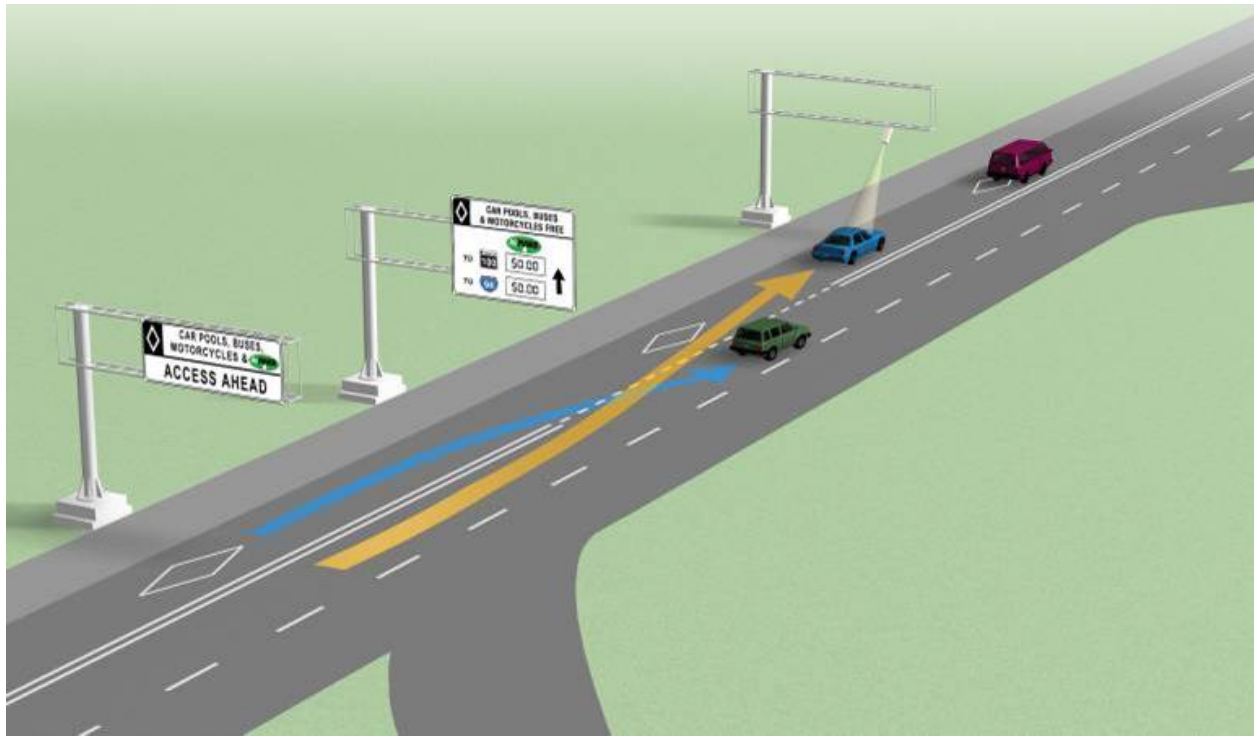


Figure 2-2. Diamond lane design (Scott, 2007).

The MnPass program which was implemented in May 2005, which converted the HOV lanes on I-394 into HOT lanes. The SOVs are charged to use the lanes while HOVs and motorcyclists can use for free. The toll rates vary from \$0.25 to \$4.00 per trip, with the highest value at \$8.00, adjusted as often as every three minutes based on the detected traffic density. The objective of the HOT lane operation is to maintain free flow conditions at all times (speeds greater than 50 mph). The MnPASS project introduced dynamic tolling on multiple consecutive roadway segments. Separate pricing to segments is necessary for managing the demand to this corridor. To maintain free-flow travel speed for all vehicles, the algorithm that adjusts the toll rates dynamically is based on the total number of vehicles per mile in the lane, the rate of change of traffic conditions and on the speed of the vehicles. When a change in the density is detected by the road-way sensors, the toll is adjusted upward or downward, determined from a “look-up” table (Halvorson et al., 2006). The toll rates are calculated at each entry point according to the

maximum traffic density downstream of each entry point. The rate calculation interval is adjusted so that traffic conditions that change rapidly to be measured. High differences between the toll rates in a single calculation interval are avoided. The toll rates for one hour of one day on I-394 are presented in Figure 2-3.

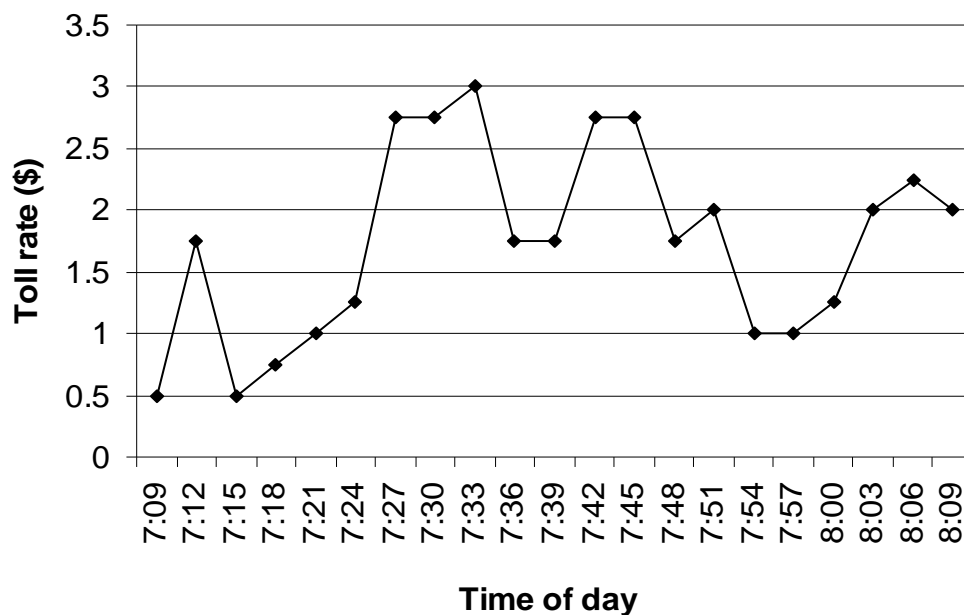


Figure 2-3. I-394 toll rates on May 18, 2005 (source: www.MnPASS.net)

A comprehensive evaluation plan is being implemented to thoroughly assess conditions and public attitudes before and during the project operations (FHWA, 2006a). Preliminary performance data of this new facility were obtained for the first six months: the average number of toll trips per week is around 15,918 and the average revenue per week is around \$12,484 (FHWA, 2006a).

2.4.2.2 I-95 HOT lane project in South Florida

I-95 is one of the main highways on the East Coast of the United States serving some of the largest urban areas in the United States. In summer 2008, Express Toll lanes opened to

reduce congestion on I-95 from I-395 in downtown Miami to Broward Boulevard in Fort Lauderdale. The modeling approach to determine the tolls depend on the existing demand which is obtained by data collection (traffic counts and on the amount of money that people are willing to pay to use the tolling facility. More precisely, the procedure for the dynamic pricing on I-95 is as follows. Firstly, real time traffic data are collected and the level of HOT lane operation is examined. Next, if the HOT lanes operate over the capacity threshold the toll rate will be increased by a preset margin of 25 cents. After the toll change, if the coming volumes decrease the toll remains the same until the new traffic data obtained. However, if the coming volumes increase or remain the same, the toll rate will be increased by the specific preset margin. In case of special event or an incident, the incident management manual override will be considered. If the HOT lanes operate with very low volumes and the minimum toll is in effect, no changes are required but if the toll is higher than the minimum the toll rate will be reduced by the preset margin and then according to the change in volumes after the toll reduction, the toll remains the same or changes again. The range of toll rates on I-95 is going to be from \$0.25 to \$10.00 according to operational conditions (normal or single lane operation).

2.4.3 Self-Learning Control Approaches for Dynamic Tolling

Yin and Lou (2006) proposed two practical approaches for the determination of pricing strategies for operating managed lanes. They considered dynamic tolls that change according to traffic conditions in order to maximize the throughput of a freeway and keep superior free flow conditions to travelers. The first approach is a feedback-control approach where one loop detector station located downstream is required to detect the traffic condition along the facility's segment. The second approach is a reactive self-learning approach that calibrates the willingness to pay using revealed preference information and then determines the optimal toll rate using the approaching flow rates, the estimated travel time and the calibrated willingness to pay. For the

implementation of this approach, two loop detectors are required, one before the toll entrance to detect approaching flows and one after the entrance to detect the flows on both the managed and the regular lanes. For the two approaches' validation, simulation experiments were conducted. The conclusions of this research are that although both approaches are simple and easy to implement, the results of the two approaches may lead to wildly varying toll pattern than can cause unstable traffic conditions. Moreover, the toll price is determined for each entry point without considering other entries, which may create inequality among motorists that enter to the managed lane from different points.

Lou et al (2007) further expanded the approach proposed by representing traffic dynamics more realistically and with an explicit formulation to optimize tolls. The impacts or the lane-changing before the entrance of the toll lanes on the freeway throughput and the travel time are considered using the multi-lane hybrid traffic flow model that was proposed by Laval and Daganzo (2006). The optimal tolls are determined for specific time intervals solving a nonlinear optimization model in order to maximize the freeway's throughput and to ensure that the critical density of HOT lanes will not be exceeded from the density of the corridor. Lou et al (2007), also, examined further the conversion of HOV lanes to HOT lanes and presented some extensions to the approach that they proposed considering more realistic cell representations for differences to HOT lane slip ramp configurations. For the validation and demonstration of the approach, simulation experiments were conducted using data from loop detectors.

2.5 Conclusions That Can Be Drawn from the Literature Review

The studies presented in this literature review are some of a large number of studies that have already been conducted for applying dynamic pricing to congested roads. The high interest for dynamic tolling all over the world as a congestion mitigation policy and its implementation

on many networks show that dynamic pricing is a viable mean of combating traffic congestion and not just a topic for academic discussion.

Although a large number of studies considering different aspects of dynamic tolling have been conducted, there are issues that have not been solved yet and deficiencies to some of the existing studies. The literature does not offer a practical and sensible approach for determining dynamic toll rates for HOT lanes. Many of the studies (see, e.g., Morrison, 1986; Palma and Lindsey, 1997; Arnott et al., 1998; Liu and McDonald, 1999) consider hypothetical and idealized situations in which analytical solutions can be derived. For example, the travel demand function or travel demand is usually assumed to be known. Therefore, there is need for the use of real data for the determination of more accurate toll rates. In contrast, Yin and Lou (2006) and Lou et al (2007) proposed two readily-implementable approaches for determining time-varying tolls in response to the traffic arrival. The first approach adjusts the toll rate based on the concept of the feedback control, while the second approach is a reactive self-learning approach and it provides first the willingness to pay using previous preference information and then it determines the optimal toll rate using the approaching flow rates. However, a more proactive approach will give better amount of tolls according to the incoming flow. In addition, there is not any robust model that takes into consideration different scenarios of demand for the determination of the toll rates. Moreover, none of the studies that have been conducted for dynamic tolling takes into consideration that the capacity of a facility is not a constant value but fluctuates over the time.

CHAPTER 3 RESEARCH APPROACH

3.1 Introduction

This chapter describes the proposed methodology for the research about dynamic tolling with real-time demand and willingness-to-pay learning for HOT lanes. The research approach includes the determination of the willingness to pay of travelers using a recursive Kalman filter, the real-time demand learning based on Bayesian inference and the empirical investigation of demand using data for SR-91 which is located at California. Furthermore, a scenario-based robust toll optimization with linear representation of cell-transmission model, stochastic demand and stochastic capacity will be addressed.

3.2 Methodology Overview

The present methodology provides a robust and proactive approach to determine toll rates for HOT lanes. The goal is to adjust the toll rates according to traffic demand and freeway's prevailing traffic conditions in a smooth manner to maximize the throughput of the freeway and to maintain free-flow conditions. For the toll determination, historical data (arrival flows and toll rates) are used in order to predict the short-term inflows and the motorists' willingness-to-pay. The flows and densities of toll and GP will be described by propagating the inflow according to the cell-transmission model. According to flows, willingness-to-pay and capacity, toll rates will be optimized (see Figure 3-1).

In the proactive and robust toll optimization approach, stochastic demand and stochastic capacity are considered. More precisely, demand scenarios are generated using the negative binomial distribution and capacity scenarios are generated using the Weibull distribution. Toll optimization will be done with respect to each scenario. However, before taking into account the demand and capacity scenarios, we consider deterministic demand and capacity to simply

illustrate the methodology. In both cases, traffic dynamics will be modeled using the cell-transmission model. Finally, simulation experiments will be conducted to validate and clearly illustrate the proposed methodology.

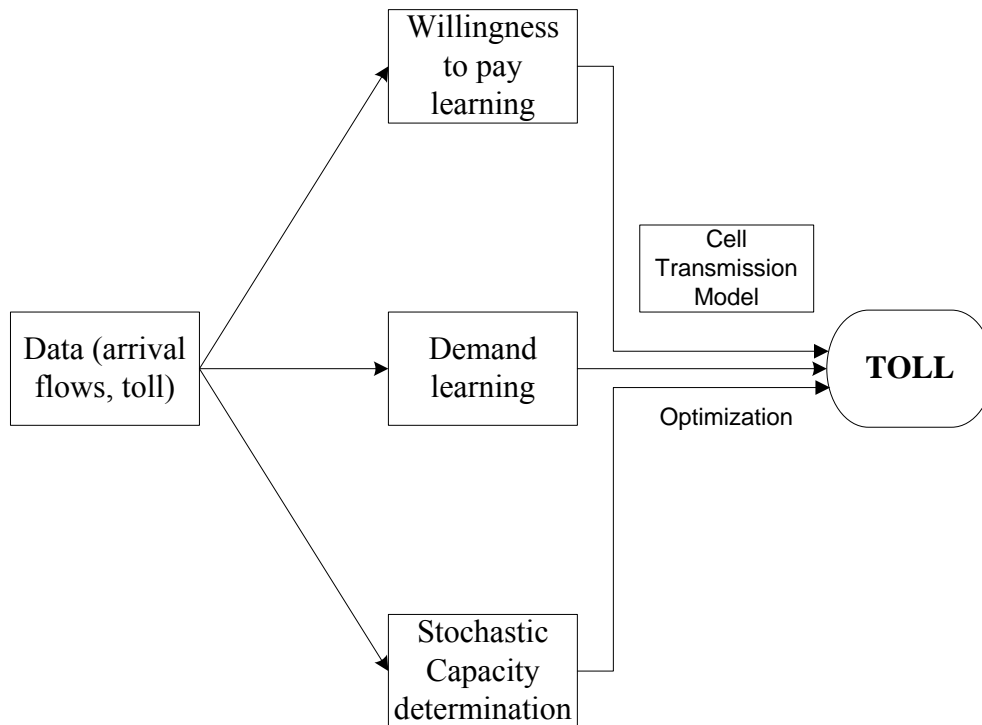


Figure 3-1. Toll determination procedure

The methodologies for the willingness-to pay learning, the demand learning, the stochastic capacity and the scenario-based robust toll optimization that will be used in this research are described below.

3.3 Modeling Preparation

This section addresses three major components of toll optimization including willingness-to-pay learning, the demand learning and the stochastic capacity determination.

3.3.1 Willingness-to-Pay Learning

3.3.1.1 Basic concept

Willingness-to-pay is the amount of money that individuals are willing to pay to obtain a good. It refers to the maximum monetary amount that a person would pay for a good. In the present research the good is the use of the HOT lanes, more precisely, travel time savings.

Motorist's willingness to pay can be gradually learned by mining data from loop detectors. The gained knowledge can be applied afterwards to determine optimal tolls to maximize the freeway's throughput rate while maintaining free-flow conditions. More precisely, it is assumed that the number of the travelers who are willing to pay in order to use the HOT lane can be formulated as an aggregate Logit model with unknown parameters if a specific toll is given. However, the Logit model's parameters can be estimated recursively using revealed-preference information that includes the flow rates before and after the lane choice which can be measured directly from the loop detectors and the travel times on the regular purpose and the toll lanes which can be estimated or directly measured by installing additional toll-tag readers along the freeway. The optimal tolls can be determined explicitly to achieve the operating objectives using the estimated Logit model.

The present approach decomposes the pricing strategies into two consecutive steps. First, the motorists' willingness to pay learning using the previous revealed-preference information and second, the determination of the optimal toll rate based on the approaching flows (here different demand scenarios will be developed), the calibrated willingness to pay and the estimated travel times.

For the implementation of this approach two set of detectors are required as it shown at figure 3-2. The first set of detectors is installed before the toll-tag reader for detecting the approaching traffic flows and the second set is installed after the reader for detecting the flows

on the toll and the general purpose lanes. Without loss of generality, we assume that there are one HOT lane, one general purpose lane and a recurrent bottleneck downstream (Figure 3-2). Furthermore, there are only one entrance and one exit to the toll lane and no on/off ramps between. In figure 3-2, $\mu(t)$ is the total approaching flow during time interval, $\lambda_T(t)$ is the flow on regular lane during time interval t and $\lambda_R(t)$ is the flow on regular lane during time interval t.

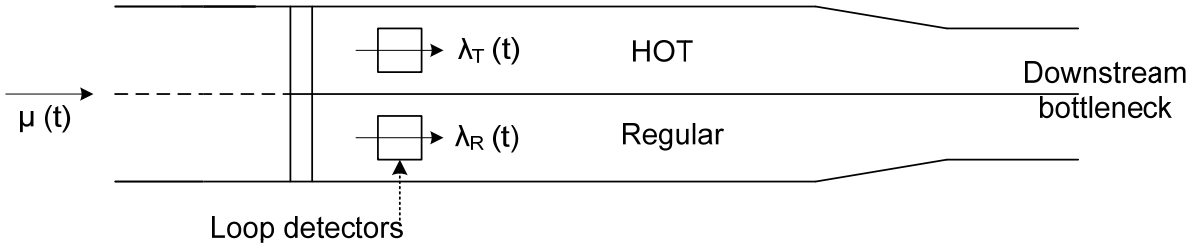


Figure 3-2. System configuration

3.3.1.2 Calibration of willingness-to-pay

To formulate the motorist's decision on whether to choose the HOT lane or not, an aggregate Logit model whose parameters are unknown is adopted as in Lou et al. (2007) and Yin and Lou (2006), given a specific value of toll $\beta(t)$ that changes over the time. Making the assumptions that motorists are homogeneous and they have the same willingness to pay, the relationships between the flow rates on HOT and regular lanes and the approaching flow rates are presented below:

$$\mu(t) = \lambda_T(t) + \lambda_R(t) \quad (3-1)$$

$$\lambda_T(t) = \mu(t) \times \frac{1}{1 + \exp(\alpha_1(c_T(t) - c_R(t)) + \alpha_2\beta(t) + \gamma)} \quad (3-2)$$

Where,

$\mu(t)$ = total approaching flow at time interval t

$\lambda_T(t)$ = flow on HOT lane at time interval t

$\lambda_R(t)$ = flow on regular lane at time interval t

$c_T(t)$ = (average) travel time on the HOT lane at time interval t

$c_R(t)$ = (average) travel time on the regular lane at time interval t

$\beta(t)$ = toll at time interval t

α_1 = marginal effect of travel time on motorists' utility

α_2 = marginal effect of toll on motorists' utility

γ = parameter that encapsulates other factors that affect motorists' willingness to pay

The ratio $\frac{\alpha_1}{\alpha_2}$ represents the motorists' value of time, in other words, their trade-off

between time savings and tolls. The variables $\mu(t)$, $\lambda_T(t)$ and $\lambda_R(t)$ can be obtained directly from loop detectors. The variables $c_T(t)$ and $c_R(t)$ can be estimated using traffic flow models or directly measured and the variable $\beta(t)$ is set by the operator. It must be mention that for the approximation of the lane choices' probabilities, the volume splits $\frac{\lambda_T(t)}{\mu(t)}$ and $\frac{\lambda_R(t)}{\mu(t)}$ are used.

A discrete Kalman filtering or a recursive least-squares technique can be used to estimate the constant parameters α_1 , α_2 and γ , in real-time operation. The equation (2) which is actually the demand function of the HOT lane can be written as follows:

$$\ln\left(\frac{\mu(t)}{\lambda_T(t)} - 1\right) = \alpha_1(c_T(t) - c_R(t)) + \alpha_2\beta(t) + \gamma \quad (3-3)$$

$$\text{Let } y(t) = \ln\left(\frac{\mu(t)}{\lambda_T(t)} - 1\right), \quad x(t) = \begin{pmatrix} \alpha_1(t) \\ \alpha_2(t) \\ \gamma(t) \end{pmatrix} \text{ and } H(t) = (c_T(t) - c_R(t) \quad \beta(t) \quad 1), \text{ the}$$

system/ observation equation (3-3) can be written as below:

$$y(t) = H(t) \cdot x(t) + e_y \quad (3-4)$$

Where,

$y(t)$: the observed system output

e_y : a random measurement error with zero mean and a known variance of σ_y

The value of σ_y can vary according to the type of sensor.

If we apply the Kalman filtering technique to estimate the parameters, we obtain:

$$\begin{cases} \begin{pmatrix} \hat{\alpha}_1(t+1) \\ \hat{\alpha}_2(t+1) \\ \hat{\gamma}(t+1) \end{pmatrix} = \begin{pmatrix} \hat{\alpha}_1(t) \\ \hat{\alpha}_2(t) \\ \hat{\gamma}(t) \end{pmatrix} + G(t) \left[\ln \left(\frac{\mu(t)}{\lambda_T(t)} - 1 \right) - (c_T(t) - c_R(t) \quad \beta(t) \quad 1) \begin{pmatrix} \hat{\alpha}_1(t) \\ \hat{\alpha}_2(t) \\ \hat{\gamma}(t) \end{pmatrix} \right] \\ G(t) = P(t) \begin{pmatrix} c_T(t) - c_R(t) \\ \beta(t) \\ 1 \end{pmatrix} \left[(c_T(t) - c_R(t) \quad \beta(t) \quad 1) P(t) \begin{pmatrix} c_T(t) - c_R(t) \\ \beta(t) \\ 1 \end{pmatrix} + \sigma_y \right]^{-1} \\ P(t) = [I - G(t)(c_T(t) - c_R(t) \quad \beta(t) \quad 1)] P(t-1) \end{cases} \quad (3-5)$$

Where,

“^”: indicate the variables which are estimates

P : the expected covariance matrix of the estimation errors

G : the Kalman gain

Equation (3-5) can update the estimates of α_1 , α_2 and γ in real time each time that a new information is obtained with an initialization of $\alpha_1(0)$, $\alpha_2(0)$, $\gamma(0)$ and $P(0)$. As time involves, accurate estimates of the impact of α_1 , α_2 and γ are expected and the initialization will not have much impact on them. In this research, homogeneous motorists are assumed with same willingness to pay. However, this procedure can be applied and even if the motorists are heterogeneous with different willingness to pay. In this case, there will be distributions associated α_1 , α_2 and γ across the population.

3.3.2 Demand Learning

This paragraph addresses the methodology of real-time demand learning based on Bayesian Inference.

First of all, traffic arrival is assumed that follows a Poisson process whose average arrival rate is unknown, denoted as Λ . Therefore, during $(0, t]$, the number of vehicles that arrived, N , follows the distribution below:

$$P(N = n | \Lambda = \lambda, t) = \frac{e^{-\lambda t} \cdot (\lambda t)^n}{n!} \quad (3-6)$$

Since Λ is unknown, from historical data, we may estimate its prior distribution as gamma distribution with parameters (α, k) and the following density function:

$$f_{\Lambda}(\lambda) = \frac{\alpha e^{-\alpha \lambda} \cdot (\alpha \lambda)^{k-1}}{\Gamma(k)} \quad (3-7)$$

Where,

$\Gamma(k) = \int_0^{\infty} e^{-x} x^{k-1} dx$, a gamma function

When k is an integer, $\Gamma(k) = (k-1)!$. Note that $k = \mu_{\Lambda}^2 / \sigma_{\Lambda}^2$ and $a = \mu_{\Lambda} / \sigma_{\Lambda}^2$. An integer k can be selected to simplify the computation. The arrival rate Λ is assumed to follow a gamma distribution. A normal distribution has at least two disadvantages in this case. Firstly, the normal distribution can give negative values while the arrival rate can never be negative and, secondly, as the pricing moves forward, Λ doesn't have a closed distribution form and this is mathematically intractable. Moreover, with the appropriate choice of the gamma distribution's parameters its shape can be like the bell shape of normal distribution.

The prior distribution of Λ may be updated in real time during operations, as demonstrated in Lin (2006). Assuming that during $(0, t]$, i number of vehicles have been observed from loop

detectors. After the observation that i vehicles have passed, the researcher has a better sense about the true arrival rate. If i is relatively large, Λ is more likely to be large too. As a consequence, the posterior probability density function of Λ can be updated by using the Bayes' theorem:

$$f_{\Lambda|N(t)=i}(\lambda) = \frac{f(\lambda) \cdot P(N(t)=i|\Lambda=\lambda)}{\int_0^\infty f(\lambda) \cdot P(N(t)=i|\Lambda=\lambda) d\lambda} = \frac{\frac{\alpha e^{-\alpha\lambda} \cdot (\alpha\lambda)^{k-1}}{\Gamma(k)} \cdot \frac{e^{-\lambda t} \cdot (\lambda t)^i}{i!}}{\int_0^\infty \frac{\alpha e^{-\alpha\lambda} \cdot (\alpha\lambda)^{k-1}}{\Gamma(k)} \cdot \frac{e^{-\lambda t} \cdot (\lambda t)^i}{i!} d\lambda} \quad (3-8)$$

$$= \frac{(\alpha+t) \cdot e^{-(\alpha+t)\lambda} \cdot ((\alpha+t) \cdot \lambda)^{k+i-1}}{\Gamma(k+i)}$$

From above, it can be detected that the posterior distribution is another gamma distribution with parameters $(\alpha+t, k+i)$ this time. With it, the number of vehicles that may arrive during time interval $(t, t+\Delta t]$ can be estimated. It turns out that vehicles follow a negative binomial distribution shown below:

$$P(N(t+\Delta t) - N(t) = n | N(t) = i) = \int_0^\infty \frac{e^{-\lambda \cdot \Delta t} \cdot (\lambda \cdot \Delta t)^n}{n!} \cdot \frac{(\alpha+t) e^{-(\alpha+t)\lambda} \cdot ((\alpha+t)\lambda)^{k+i-1}}{\Gamma(k+i)} d\lambda$$

$$= \frac{\Gamma(k+i+n)}{n! \Gamma(k+i)} \cdot \left(\frac{\alpha+t}{\alpha+t+\Delta t} \right)^{k+i} \cdot \left(\frac{\Delta t}{\alpha+t+\Delta t} \right)^n \quad (3-9)$$

$$= \binom{k+i+n-1}{n} \cdot \left(\frac{\alpha+t}{\alpha+t+\Delta t} \right)^{k+i} \cdot \left(\frac{\Delta t}{\alpha+t+\Delta t} \right)^n \quad (k \text{ is integer})$$

The mean and the variance for the above binomial distribution are:

$$\mu_b = (k+i) \frac{\frac{\Delta t}{\alpha+t+\Delta t}}{\frac{\alpha+t+\Delta t}{\alpha+t+\Delta t}} = (k+i) \frac{\Delta t}{(\alpha+t)} \quad (3-10)$$

$$\sigma_b^2 = (k+i) \cdot \frac{\frac{\Delta t}{\alpha+t+\Delta t}}{\left(\frac{\alpha+t}{\alpha+t+\Delta t}\right)^2} = (k+i) \frac{\Delta t (\alpha+t+\Delta t)}{(\alpha+t)^2} \quad (3-11)$$

The result that the number of vehicles follows the negative binomial distribution seems logical because the future vehicles come one immediately after another and not in a continuous-time setting. The negative binomial distribution is a discrete probability distribution.

3.3.3 Stochastic Capacity Determination

The capacity of a freeway is not a constant value but varies according to traffic, roadway, control, and weather conditions. The concept of stochastic capacity is more realistic and gives better insight for the maximum flow that can pass from a point. In the present paragraph, the approach for stochastic capacity determination based on the methodology proposed by Brilon et al. (2005) is addressed.

The capacity distribution function is difficult to be estimated because the capacity itself cannot be measured directly. The capacity distribution function is defined as follows:

$$F_c(q) = p(c \leq q) \quad (3-12)$$

Where,

$F_c(q)$ = capacity distribution function

c = capacity

q = traffic volume

To find an estimate that will be appropriate for the whole capacity distribution function, more information about the mathematical type of the distribution function $F_c(q)$ must be known.

Brilon and Zurlinden (2003) examined various distribution functions like Normal, Gamma and Weibull. A maximum likelihood technique was used for the estimation of the distribution functions' parameters. The likelihood function used is:

$$L = \prod_{i=1}^n f_c(q_i)^{\delta_i} \cdot [1 - F_c(q_i)]^{1-\delta_i} \quad (3-13)$$

Where,

$f_c(q_i)$ = statistical density function of capacity c

$F_c(q_i)$ = cumulative distribution function of capacity c

n = number of intervals

$\delta_i = 1$, if uncensored (the observed volume causes a breakdown)

$\delta_i = 0$, elsewhere

The parameters of the distribution function are calibrated by maximizing the likelihood function or its natural logarithm. Based on the value of the likelihood function, it showed that the capacity of a freeway section follows the Weibull distribution with a nearly constant shape parameter.

The function of the Weibull distribution is shown below:

$$F(x) = 1 - e^{-\left(\frac{x}{\beta}\right)^\alpha} \quad \text{for } x \geq 0 \quad (3-14)$$

Where,

α = shape parameter

β = scale parameter

The random capacity is considered because it leads to more robust toll optimization and more realistic roll rates.

3.4 Robust Toll optimization

Robust Toll Optimization is an approach to optimize the amount of toll that motorists have to pay to use the HOT lanes. This approach is called robust because it has to deal with uncertainty. Some of the parameters are random variables and the problem data are described with possible realizations called as scenarios.

3.4.1 Deterministic Case

To build a basis for robust toll optimization, we consider deterministic case first. In the deterministic case, the traffic demand of the planning horizon, the willingness-to-pay and the capacity are considered deterministic and given. In this case the toll optimization is done by solving the maximization problem presented in paragraph 3.4.1.2.

3.4.1.1 Modeling traffic dynamics

Modeling traffic dynamics is the key for the toll optimization. In the present research, traffic dynamics will be modeled using the linear representation of cell-transmission model (CTM) (Daganzo, 1994, 1995). The CTM is chosen because it covers all the fundamental flow-density, speed-density, speed-flow relationships and has been validated by field data (Lin and Daganzo, 1994; Lin and Ahanotu, 1995). In addition, because it is a dynamic formulation, time variant demands can be modeled easily.

The evolution of traffic is examined over a one-way road with one access and one egress. Actually, the traffic is modeled for both the toll and GP lanes. The traffic conditions are updated every few seconds. To model traffic, the road is divided to cells, i.e. homogeneous sections (Figure 3-3).

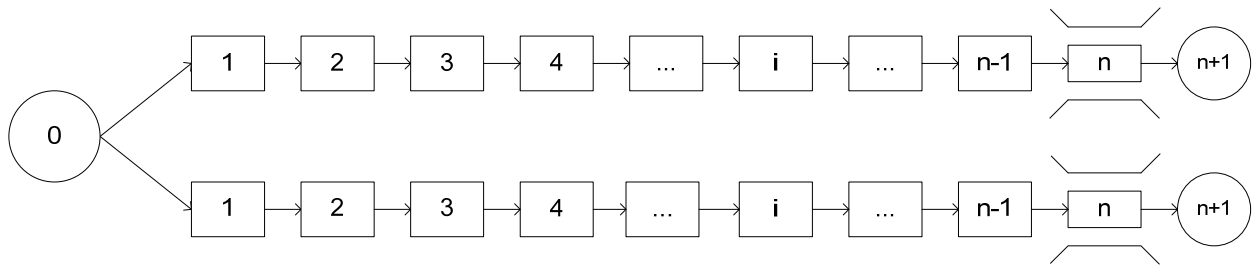


Figure 3-3. Representation of the road with cells

In Figure 3-4, the flow-density diagram is shown. More precisely, there are two triangles in Figure 3-4, the big one is for the lanes upstream of the bottleneck and the small is for the lanes in

the downstream bottleneck. In the diagram, q_{\max} and $q_{b\max}$ are the maximum flows at the lanes upstream and in the bottleneck, respectively, while k_o and k_{bo} are the optimum densities at the lanes upstream and in the bottleneck, respectively. The jam density (k_j) is the same for all the lanes.

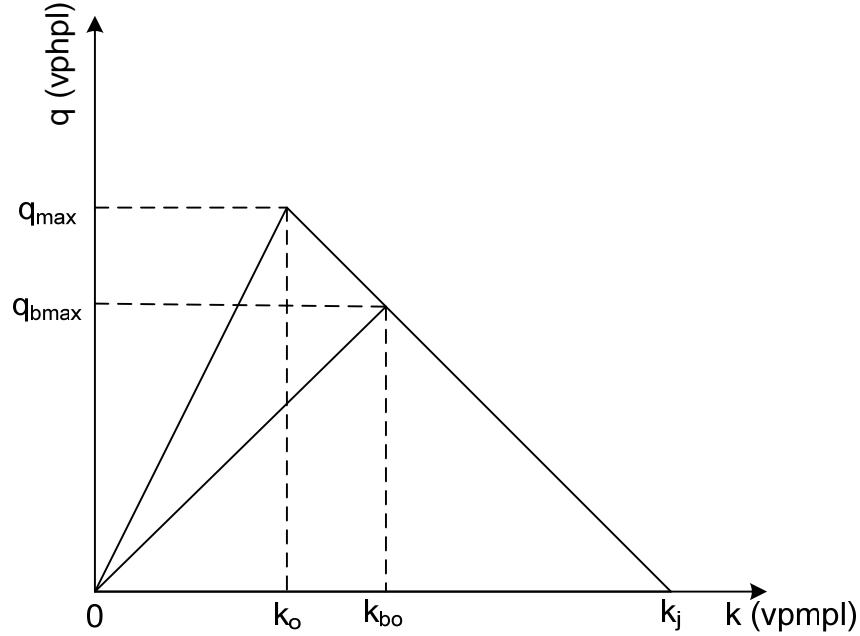


Figure 3-4. Flow-density relationship

In cell-transmission model, the road is divided into cells (Figure 3-3). The length of the cells is set equal to the distance traveled by a typical vehicle in free-flow conditions in one time interval. Therefore, under uncongested conditions, all the vehicles in a cell will travel to the next cell in one time interval so the system evolution obeys (Daganzo, 1994):

$$n_{i+1}(t+1) = n_i(t) \quad (3-15)$$

Where,

$n_i(t)$ = number of vehicles in cell i at time interval t

The equation (3-15) holds for all flows, unless queues are developed. To incorporate queuing in the model, two constants are introduced: $N_i(t)$ which is the maximum number of vehicles that can be presented in cell i at time interval t and $Q_i(t)$ which is the maximum number of vehicles that can flow into cell i at time interval t . Thus, the number of vehicles that can flow from cell i into cell $i+1$ at time interval t is the smallest of the three values: $n_i(t)$, $Q_{i+1}(t)$ and $\frac{W}{V}(N_{i+1}(t) - n_{i+1}(t))$ which is the empty space in cell i at time interval t , W is the backward wave speed and V is the free-flow speed. The cell-transmission model is based on equation (3-16) (flow conservation):

$$n_i(t+1) = n_i(t) + y_i(t) - y_{i+1}(t) \quad (3-16)$$

Where $y_{i+1}(t) = \min \left\{ n_i(t), Q_{i+1}(t), \frac{W}{V}(N_{i+1}(t) - n_{i+1}(t)) \right\}$ indicating that flows depend on the current conditions at time t . The cell occupancies will be updated with each time interval.

The boundary conditions are specified by means of the first (input) and the last (output) cell. The first cell (cell 0) has infinite size $N_0(t) = \infty$ and the inflow $Q_0(t)$ is equal to the desired link input flow. The last cell has, also, infinite size $N_{i+1}(t) = \infty$ and a suitable, maybe time-varying, capacity.

3.4.1.2 Model formulation

The formulation of the toll optimization problem is shown below:

$$\max_{(y,n)} \sum_{t=0}^T \left\{ \sum_{j=0}^t (y_{n+1}^T(j) + y_{n+1}^G(j)) + \theta \cdot \sum_{i=1}^n \left[\min(k_{i,opt}^T(t) - n_i^T(t), 0) \right] \right\} \quad (3-17)$$

s..t.

$$y_1^T(t) = q(t) \cdot \frac{1}{1 + \exp(\alpha_1(c^T(t) - c^G(t)) + \alpha_2\beta(t) + \gamma)} \quad \forall t = 0, \dots, T \quad (3-18)$$

$$c^T(t) = \sum_{i=1}^n \left[l / \left(W^T \left(N_i^T(t) / n_i^T(t) - 1 \right) \right) \right] \quad \forall t = 0, \dots, T \quad (3-19)$$

$$c^G(t) = \sum_{i=1}^n \left[l / \left(W^G \left(N_i^G(t) / n_i^G(t) - 1 \right) \right) \right] \quad \forall t = 0, \dots, T \quad (3-20)$$

$$y_1^T(t) + y_1^G(t) = q(t) \quad \forall t = 0, \dots, T \quad (3-21)$$

$$n_i^T(t+1) = n_i^T(t) + y_i^T(t) - y_{i+1}^T(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-22)$$

$$y_i^T(t) \leq n_{i-1}^T(t) \quad \forall t = 0, \dots, T \text{ and } i = 2, \dots, n+1 \quad (3-23)$$

$$y_i^T(t) \leq Q_i^T(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-24)$$

$$y_i^T(t) \leq \frac{W^T}{V^T} \left(N_i^T(t) - n_i^T(t) \right) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-25)$$

$$n_i^G(t+1) = n_i^G(t) + y_i^G(t) - y_{i+1}^G(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-26)$$

$$y_i^G(t) \leq n_{i-1}^G(t) \quad \forall t = 0, \dots, T \text{ and } i = 2, \dots, n+1 \quad (3-27)$$

$$y_i^G(t) \leq Q_i^G(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-28)$$

$$y_i^G(t) \leq \frac{W^G}{V^G} \left(N_i^G(t) - n_i^G(t) \right) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-29)$$

$$0 \leq \beta(t) \leq \beta_{\max} \quad \forall t = 0, \dots, T \quad (3-30)$$

$$y_i^T(t) \geq 0 \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-31)$$

$$y_i^G(t) \geq 0 \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-32)$$

$$n_i^T(t) \geq 0 \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-33)$$

$$n_i^G(t) \geq 0 \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-34)$$

Where,

i = refers to a cell i

$y_i^T(t)$ = actual inflow to cell i at time interval $(t, t+1)$ at HOT lane

$y_i^G(t)$ = actual inflow to cell i at time interval $(t, t+1)$ at GP lane

θ = penalty parameter

$k_{i,opt}^T(t)$ = optimal (or critical) density of cell i at time interval t of the HOT lane

$q(t)$ = demand at time interval t

$c^T(t)$ = (average) travel time on the HOT lane at time interval t

$c^R(t)$ = (average) travel time on the GP lane at time interval t

l = length of each cell

$N_i^T(t)$ = maximum number of vehicles that can be presented in cell i at time interval t at HOT lane

$N_i^G(t)$ = maximum number of vehicles that can be presented in cell i at time interval t at GP lane

$\beta(t)$ = toll at time interval t

β_{\max} = maximum toll that can be set

α_1 = marginal effect of travel time on motorists' utility

α_2 = marginal effect of toll on motorists' utility

γ = parameter that encapsulates other factors that affect motorists' willingness to pay

$n_i^T(t)$ = number of vehicles in cell i at time interval t at HOT lane

$n_i^G(t)$ = number of vehicles in cell i at time interval t at GP lane

$Q_i^T(t)$ = maximum number of vehicles that can flow into cell i at time interval t at HOT lane

$Q_i^G(t)$ = maximum number of vehicles that can flow into cell i at time interval t at GP lane

W^T = backward wave speed at HOT lane

V^T = free-flow speed at GP lane

W^G = backward wave speed at HOT lane

V^G = free-flow speed at GP lane

The objective function (3-17) is to maximize the sum of the freeway's throughput at the downstream bottleneck while ensuring that the density of cell i of the HOT lane does not exceed the critical density. Equations (3-18)–(3-20) indicate how many vehicles are going to choose the HOT lane prescribed by a Logit model. Constraint (3-21) assures that all the incoming flow goes either to HOT or GP lane. Constraint (3-22) and (3-26) are the flow conservation constraints for HOT and GP lane, respectively. Constraints (3-23)–(3-25) and (3-27)–(3-29) anticipate that the actual inflow at HOT and GP lane, respectively, will take up the maximum allowed. Actually, constraints (3-22)–(3-29) model the traffic dynamics prescribed by CTM. Constraint (3-30) doesn't allow toll rate to take values lower than zero and greater than a specific maximum value. Constraints (3-31)–(3-34) ensure no negativity.

The above formulation has some drawbacks. First of all, the problem is not linear so it may require more computational time. Although, in the deterministic case, this is not very important, after the development of demand and capacity scenarios is very critical because this problem will be solved many times. Moreover, the planning horizon and the time interval that toll changes must be the same. In this case traffic propagates every some seconds while toll changes every some minutes. Tolls cannot be set every some seconds because this is not practical and the traffic modeling cannot be every some minutes because the results won't be accurate. This issue can be solved by adding more constraints to the formulation ensuring that the toll will be constant every some minutes but this will increase the computational complexity.

The reasons mentioned above lead to the need of a two step formulation. In the first step, the flows to each lane (general purpose or toll) will be determined given a demand and in the second step, the toll for each time interval will be computed. The two step formulation is as follows:

Stage 1:

$$\max_{(y,n)} \sum_{t=0}^T \left\{ \sum_{j=0}^t (y_{n+1}^T(j) + y_{n+1}^G(j)) + \theta \cdot \sum_{i=1}^n \left[\min(k_{i,opt}^T(t) - n_i^T(t), 0) \right] \right\} \quad (3-17)$$

s.t.

$$y_1^T(t) + y_1^G(t) = q(t) \quad \forall t = 0, \dots, T \quad (3-21)$$

$$n_i^T(t+1) = n_i^T(t) + y_i^T(t) - y_{i+1}^T(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-22)$$

$$y_i^T(t) \leq n_{i-1}^T(t) \quad \forall t = 0, \dots, T \text{ and } i = 2, \dots, n+1 \quad (3-23)$$

$$y_i^T(t) \leq Q_i^T(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-24)$$

$$y_i^T(t) \leq \frac{W^T}{V^T} (N_i^T(t) - n_i^T(t)) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-25)$$

$$n_i^G(t+1) = n_i^G(t) + y_i^G(t) - y_{i+1}^G(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-26)$$

$$y_i^G(t) \leq n_{i-1}^G(t) \quad \forall t = 0, \dots, T \text{ and } i = 2, \dots, n+1 \quad (3-27)$$

$$y_i^G(t) \leq Q_i^G(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-28)$$

$$y_i^G(t) \leq \frac{W^G}{V^G} (N_i^G(t) - n_i^G(t)) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-29)$$

$$q(t) \cdot p_{\min} \leq y_1^T(t) \leq q(t) \cdot p_{\max} \quad \forall t = 0, \dots, T \quad (3-30)$$

$$y_i^T(t) \geq 0 \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-31)$$

$$y_i^G(t) \geq 0 \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-32)$$

$$n_i^T(t) \geq 0 \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-33)$$

$$n_i^G(t) \geq 0 \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-34)$$

Stage 2:

$$\min \sum_{t=1}^{\hat{T}} \left[q(t) \cdot \frac{1}{1 + \exp(\alpha_1(c^T(t) - c^G(t)) + \alpha_2\beta + \gamma)} - y_1^T(t) \right]^2 \quad (3-36)$$

$$\text{s.t.} \quad 0 \leq \beta \leq \beta_{\max} \quad \forall t = 0, \dots, \hat{T} \quad (3-30)$$

Where,

$$c^T(t) = \sum_{i=1}^n \left[l / \left(W^T(N_i^T(t)/n_i^T(t) - 1) \right) \right] \quad \forall t = 0, \dots, \hat{T} \quad (3-19)$$

$$c^G(t) = \sum_{i=1}^n \left[l / \left(W^G(N_i^G(t)/n_i^G(t) - 1) \right) \right] \quad \forall t = 0, \dots, \hat{T} \quad (3-20)$$

In the above formulation, p_{\min} and p_{\max} are the minimum and maximum probability, respectively, of the incoming vehicles to choose the HOT lane.

The objective function of the stage 1, (3-17), as mentioned before, is to maximize the sum of the freeway's throughput at the downstream bottleneck while ensuring that the density of cell i of the HOT lane does not exceed the critical density. The objective function of stage 2, (3-36), is to minimize the difference between the actual inflow at HOT lane and the predicted inflow at the same lane. Constraint (3-35) ensures that the vehicles choose to travel to HOT lane is a percentage of the entire incoming flow.

3.4.2 Scenario-based optimization

Scenario-based optimization, as mentioned above, is an approach to optimize the tolls considering different scenarios for some parameters. The purpose is to determine a toll that

performs reasonably well across all the scenarios. In other words, the approach to determine the toll rates is proactive and robust.

In this part, different scenarios for the approaching flows (demand) and the available capacity will be generated, using the Weibull distribution for capacity and the updated negative binomial distribution for demand. Each scenario will be associated with positive probability of occurrence and it will specify how many vehicles may come during the planning horizon and what the available capacity of the downstream bottleneck is. Due to the fact that designing for the worst case may be very conservative, a toll against “high-consequence” scenarios is determined. The performance measure for examining the toll is the conditional value-at-risk, mean shortfall, mean excess loss or tail value-at-risk. This performance measure allows handling large number of scenarios and offers computational efficiency (Rockafellar and Uryasev, 2002). Figure 3-4 illustrates the probability density function and the mass function of a loss.

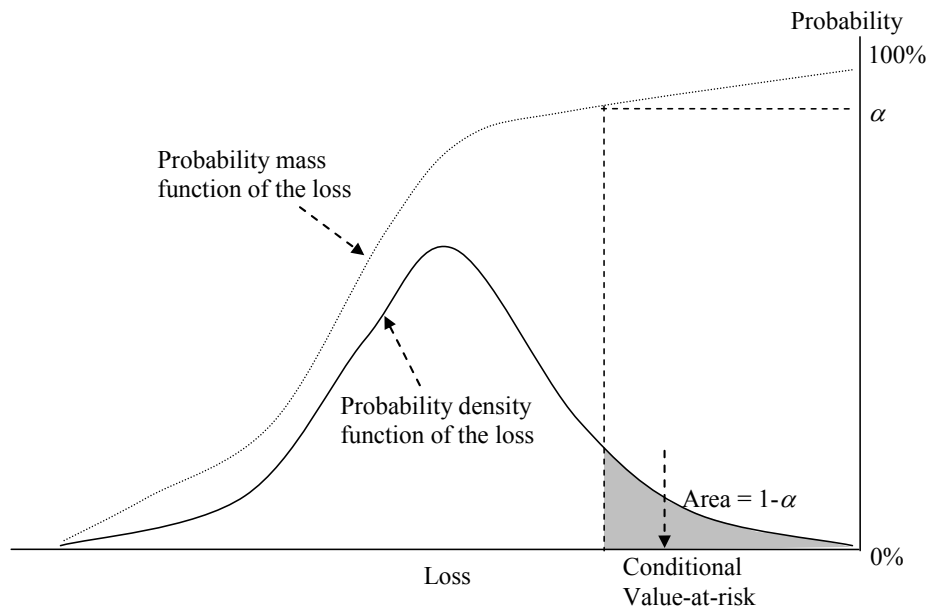


Figure 3-5. A loss function (Yin, 2007)

Each scenario, as mentioned above, indicates the number of vehicles that may arrive on each time interval and the available capacity at downstream bottleneck at each time interval. We generate a set of scenarios, S , and each scenario is indicated by the superscript s . Therefore, the robust toll optimization model is formulated as follows:

Stage 1:

$$\max_{(y,n)} \sum_{t=0}^T \left\{ \sum_{j=0}^t (y_{n+1}^{T s}(j) + y_{n+1}^{G s}(j)) + \theta \cdot \sum_{i=1}^n [\min(k_{i,opt}^T(t) - n_i^T(t), 0)] \right\} \quad (3-37)$$

s.t.

$$y_1^{T s}(t) + y_1^{G s}(t) = q^s(t) \quad \forall t = 0, \dots, T \quad (3-38)$$

$$y_i^T(t) \leq Q_i^{T s}(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-39)$$

$$y_i^{G s}(t) \leq Q_i^{G s}(t) \quad \forall t = 0, \dots, T \text{ and } i = 1, \dots, n \quad (3-40)$$

$$q^s(t) \cdot p_{\min} \leq y_1^T(t) \leq q^s(t) \cdot p_{\max} \quad \forall t = 0, \dots, T \quad (3-41)$$

(3-21) – (3-29), (3-31) – (3-35)

Stage 2:

$$\min_{\beta, \xi} Z = \xi + \frac{1}{1-a} \sum_{s \in S} p_s \cdot \max(L^s(\beta) - \xi, 0) \quad (3-42)$$

s.t.

$$0 \leq \beta \leq \beta_{\max} \quad \forall t = 0, \dots, \hat{T} \quad (3-30)$$

Where,

$$L^s(\beta) = \sum_{t=1}^{\hat{T}} \left[q^s(t) \cdot \frac{1}{1 + \exp(\alpha_1(c^{T s}(t) - c^{G s}(t)) + \alpha_2\beta + \gamma)} - y_1^{T s}(t) \right]^2 \quad (3-43)$$

s = each scenario

α = specified confidence level

p_s = probability of occurrence for scenario s

$q^s(t)$ = demand for scenario s

$y_1^{Ts}(t)$ = demand at toll lane at time interval t with respect to each scenario s

$Q_i^{Ts}(t)$ = capacity at cell i at time interval t at HOT lane

$Q_i^{Gs}(t)$ = capacity at cell i at time interval t at GP lane

The scenario-based robust toll optimization is another two step model like in the deterministic case. At stage 1, the objective function, (3-37), is to maximize the freeway's throughput considering the scenarios by deciding the demand split between toll and regular lanes. The constraints are the same as in the stage 1 of the deterministic case [constraints (3-21)–(3-29), (3-31)–(3-35)] but will be respect to each scenario. The problem of stage 2 is to minimize the conditional value-at-risk by deciding a robust toll.

CHAPTER 4 DEMAND LEARNING RESULTS

4.1 Introduction

This chapter implements the real-time demand learning based on empirical data from California SR-91. The methodology proposed in the previous chapter will be applied and the results will be presented. Last, the tasks which are to be achieved to accomplish the objective are discussed.

4.2 Empirical Analysis

The approach proposed in chapter 3 for the demand learning has been exercised using flow data from the SR-91 (Figure 4-1). State Route 91 is a major east-west freeway located entirely within Southern California and serving several regions of the Los Angeles metropolitan area. Specifically, it runs from Vermont Avenue in Gardena, just west of the junction with the Harbor Freeway (Interstate 110), east to Riverside at the junction with the Pomona (State Route 60 west of 91), Moreno Valley (State Route 60 east of the 91), and Escondido (Interstate 215).



Figure 4-1. SR-91

The data of total flow at SR-91 have been collected every five minutes during the morning period from 5:00a.m. to 10:00a.m. in September 2001. For the analysis, the data of five days are used. The flows of the first four days are used in order to calculate the parameters of the gamma distribution k and a . The flows of the fifth day are predicted using the approach presented above. Then the results of the presented approach are compared with the actual data.

The analysis conducted in two ways. First, the data of the entire period between 5:00 a.m. to 10:00 a.m. are used to predict the flows of the fifth day. Then, the analysis is conducted only with the data of the peak hour.

4.2.1 Analysis Using the Data of the Entire Period

The data collected are presented in table 4-1 and the analysis in table 4-2.

Table 4-1. Data from SR-91

	Total arrival flow (veh/5min)	Total arrival flow (veh/5min)	Total arrival flow (veh/5min)	Total arrival flow (veh/5min)
Time (a.m.)	9/12/2001	9/18/2001	9/19/2001	9/25/2001
5:00	8664	7968	8400	8613
5:05	8916	8616	8736	7908
5:10	8076	9384	8496	7740
5:15	8052	8352	8304	8013
5:20	7608	7920	8052	8076
5:25	7416	8352	8316	8316
5:30	7680	7968	8340	8184
5:35	7932	7956	7932	7812
5:40	8040	8064	8387	8172
5:45	7800	7680	7335	7980
5:50	7416	7512	6216	7680
5:55	7668	7968	5244	7572
6:00	7596	7245	5040	7752
6:05	7356	7932	6696	8016
6:10	7488	7356	7128	7740
6:15	7536	7668	6888	8664
6:20	7176	8088	6264	8196
6:25	7404	8064	6456	7884
6:30	7476	7968	6960	8244
6:35	7212	7824	6156	8136

Table 4-1. Continued

Time (a.m.)	9/12/2001	9/18/2001	9/19/2001	9/25/2001
6:40	8076	7920	6720	8892
6:45	7920	7896	6612	8208
6:50	7752	7548	6960	8364
6:55	7620	7668	7755	8400
7:00	7536	8080	7613	7740
7:05	7848	7980	8508	7848
7:10	7620	8136	8244	8175
7:15	7704	8160	8400	7980
7:20	7248	7680	7776	8076
7:25	7908	8292	6888	7596
7:30	7848	8100	7212	7653
7:35	7896	7944	7080	7812
7:40	7440	8496	7164	8088
7:45	7428	7920	6900	8124
7:50	7464	8388	7480	7488
7:55	7380	7860	8016	7548
8:00	7347	7973	8328	7560
8:05	7680	7644	7812	7272
8:10	7416	7440	7860	7464
8:15	7320	7872	8220	8052
8:20	7824	7404	7800	7236
8:25	7656	7608	8205	7404
8:30	7512	7824	7980	7956
8:35	7740	7704	7656	7227
8:40	7524	7704	7827	7644
8:45	7200	7464	7320	7932
8:50	7164	7884	7935	6876
8:55	8100	7200	8124	7344
9:00	7596	7980	8100	7613
9:05	7152	7368	7836	6192
9:10	7044	7733	7452	8100
9:15	7470	7752	7656	7788
9:20	7476	7476	8028	7284
9:25	7764	7680	7836	8052
9:30	7740	7560	7596	7212
9:35	7596	7764	7704	7140
9:40	7584	7716	7872	7524
9:45	7632	7260	7392	7368
9:50	7920	7416	7668	7476
9:55	7452	7320	7128	7092
10:00	7575	7188	7404	6732
Average	7634	7834	7531	7774

From table 4-1, the average μ_{Λ} and the standard deviation σ_{Λ} of the average flows are calculated.

$$\mu_{\Lambda} = 7693.5$$

$$\sigma_{\Lambda} = 136.80$$

The above parameters refer to 5 minutes period so the average μ_{Λ} and the standard deviation σ_{Λ} for 1 minute period are:

$$\mu_{\Lambda} = 7693.5 / 5 = 1538.70$$

$$\sigma_{\Lambda} = 136.80 / 5 = 27.36$$

Therefore, $k = \mu_{\Lambda}^2 / \sigma_{\Lambda}^2$

$$k = \mu_{\Lambda}^2 / \sigma_{\Lambda}^2 \approx 3163$$

$$a = \mu_{\Lambda} / \sigma_{\Lambda}^2 = 2.0555$$

The number of vehicles that may arrive during time interval $(t, t + \Delta t]$ can be estimated using the negative binomial distribution shown below ($\Delta t = 5$ min):

$$P(N(t+5) - N(t) = n | N(t) = i) = \binom{3163 + i + n - 1}{n} \cdot \left(\frac{2.0555 + t}{2.0555 + t + 5} \right)^{3163 + i} \cdot \left(\frac{5}{2.0555 + t + 5} \right)^n$$

The mean and the variance for the above binomial distribution are:

$$\mu_b = (3163 + i) \frac{5}{(2.0555 + t)}$$

$$\sigma_b = (3163 + i) \frac{5 \cdot (2.0555 + t + 5)}{(2.0555 + t)^2}$$

The analysis and the estimated flows per 5 minutes are presented in table 4-2.

Table 4-2. Estimated flows

Cumulative arrivals	Actual Time	Time t	Δt	Probability Mean (Estimated flows)	Variance	Standard deviation
0						
8440	5:00 AM	0	5	7693.65	26407.62	162.50
16885	5:05 AM	5	5	8222.56	14049.54	118.53
25225	5:10 AM	10	5	8314.81	11763.35	108.46
32905	5:15 AM	15	5	8322.20	10761.93	103.74
40801	5:20 AM	20	5	8176.61	10030.25	100.15
49345	5:25 AM	25	5	8124.75	9626.25	98.11
57373	5:30 AM	30	5	8190.15	9467.64	97.30
64693	5:35 AM	35	5	8168.27	9270.43	96.28
72637	5:40 AM	40	5	8067.42	9026.55	95.01
80029	5:45 AM	45	5	8054.30	8910.13	94.39
87565	5:50 AM	50	5	7990.69	8758.20	93.59
94921	5:55 AM	55	5	7950.84	8647.60	92.99
103201	6:00 AM	60	5	7902.91	8539.67	92.41
110773	6:05 AM	65	5	7931.03	8522.41	92.32
118789	6:10 AM	70	5	7906.12	8454.73	91.95
126529	6:15 AM	75	5	7913.25	8426.72	91.80
134437	6:20 AM	80	5	7902.69	8384.24	91.57
142225	6:25 AM	85	5	7903.00	8356.90	91.42
150601	6:30 AM	90	5	7896.75	8325.66	91.25
158461	6:35 AM	95	5	7921.44	8329.53	91.27
166609	6:40 AM	100	5	7918.43	8306.38	91.14
174517	6:45 AM	105	5	7929.15	8299.48	91.10
182521	6:50 AM	110	5	7928.21	8281.97	91.01
190513	6:55 AM	115	5	7931.45	8270.23	90.94
198126	7:00 AM	120	5	7933.93	8258.94	90.88
206226	7:05 AM	125	5	7921.31	8233.04	90.74
213918	7:10 AM	130	5	7928.08	8228.25	90.71
221310	7:15 AM	135	5	7919.46	8208.38	90.60
229182	7:20 AM	140	5	7900.90	8178.99	90.44
236814	7:25 AM	145	5	7899.91	8168.52	90.38
244134	7:30 AM	150	5	7891.11	8150.59	90.28
252270	7:35 AM	155	5	7872.92	8123.56	90.13
259842	7:40 AM	160	5	7881.04	8124.20	90.13
267462	7:45 AM	165	5	7871.79	8107.39	90.04
275154	7:50 AM	170	5	7864.47	8093.02	89.96
283290	7:55 AM	175	5	7859.60	8081.56	89.90
291426	8:00 AM	180	5	7867.19	8083.26	89.91
299226	8:05 AM	185	5	7874.38	8084.86	89.92
307122	8:10 AM	190	5	7872.44	8077.39	89.87
314742	8:15 AM	195	5	7873.04	8072.81	89.85
322350	8:20 AM	200	5	7866.78	8061.45	89.79

Table 4-2. Continued

Cumulative arrivals	Actual Time	Time t	Δt	Probability Mean (Estimated flows)	Variance	Standard deviation
330030	8:25 AM	205	5	7860.53	8050.35	89.72
337458	8:30 AM	210	5	7856.27	8041.51	89.67
344766	8:35 AM	215	5	7846.41	8027.15	89.59
351942	8:40 AM	220	5	7834.28	8010.69	89.50
359574	8:45 AM	225	5	7819.79	7991.99	89.40
367164	8:50 AM	230	5	7815.74	7984.14	89.35
374856	8:55 AM	235	5	7810.98	7975.73	89.31
381924	9:00 AM	240	5	7808.52	7969.82	89.27
389364	9:05 AM	245	5	7793.54	7951.26	89.17
396408	9:10 AM	250	5	7786.52	7940.98	89.11
404400	9:15 AM	255	5	7772.08	7923.26	89.01
412164	9:20 AM	260	5	7776.28	7924.65	89.02
419919	9:25 AM	265	5	7776.05	7921.63	89.00
427407	9:30 AM	270	5	7775.66	7918.57	88.99
435171	9:35 AM	275	5	7770.47	7910.70	88.94
443511	9:40 AM	280	5	7770.35	7908.10	88.93
450651	9:45 AM	285	5	7780.28	7915.79	88.97
458631	9:50 AM	290	5	7769.31	7902.33	88.90
465903	9:55 AM	295	5	7772.86	7903.69	88.90
473439	10:00 AM	300	5	7764.57	7893.10	88.84

The actual and the predicted flows are illustrated in table 4-3.

Table 4-3. Actual and predicted flows

Actual Flow	Probability Mean	Standard deviation	mean-3* σ	mean+3* σ
9/26/2001				
8440	7694	162.50	7206	8181
8445	8223	118.53	7867	8578
8340	8315	108.46	7989	8640
7680	8322	103.74	8011	8633
7896	8177	100.15	7876	8477
8544	8125	98.11	7830	8419
8028	8190	97.30	7898	8482
7320	8168	96.28	7879	8457
7944	8067	95.01	7782	8352
7392	8054	94.39	7771	8337
7536	7991	93.59	7710	8271
7356	7951	92.99	7672	8230
8280	7903	92.41	7626	8180
7572	7931	92.32	7654	8208
8016	7906	91.95	7630	8182
7740	7913	91.80	7638	8189

Table 4-3. Continued

Actual Flow	Probability Mean	Standard deviation	mean-3* σ	mean+3* σ
7908	7903	91.57	7628	8177
7788	7903	91.42	7629	8177
8376	7897	91.25	7623	8170
7860	7921	91.27	7648	8195
8148	7918	91.14	7645	8192
7908	7929	91.10	7656	8202
8004	7928	91.01	7655	8201
7992	7931	90.94	7659	8204
7613	7934	90.88	7661	8207
8100	7921	90.74	7649	8194
7692	7928	90.71	7656	8200
7392	7919	90.60	7648	8191
7872	7901	90.44	7630	8172
7632	7900	90.38	7629	8171
7320	7891	90.28	7620	8162
8136	7873	90.13	7603	8143
7572	7881	90.13	7611	8151
7620	7872	90.04	7602	8142
7692	7864	89.96	7595	8134
8136	7860	89.90	7590	8129
8136	7867	89.91	7597	8137
7800	7874	89.92	7605	8144
7896	7872	89.87	7603	8142
7620	7873	89.85	7603	8143
7608	7867	89.79	7597	8136
7680	7861	89.72	7591	8130
7428	7856	89.67	7587	8125
7308	7846	89.59	7578	8115
7176	7834	89.50	7566	8103
7632	7820	89.40	7552	8088
7590	7816	89.35	7548	8084
7692	7811	89.31	7543	8079
7068	7809	89.27	7541	8076
7440	7794	89.17	7526	8061
7044	7787	89.11	7519	8054
7992	7772	89.01	7505	8039
7764	7776	89.02	7509	8043
7755	7776	89.00	7509	8043
7488	7776	88.99	7509	8043
7764	7770	88.94	7504	8037
8340	7770	88.93	7504	8037
7140	7780	88.97	7513	8047
7980	7769	88.90	7503	8036
7272	7773	88.90	7506	8040
7536	7765	88.84	7498	8031

The following figure (figure 4-2) shows the plot of the actual flows measured every five minutes and the plot of the predicted (probability) mean after subtracting three standard deviations and after adding three times the standard deviation of the probability. According to the Chebyshev's inequality, within 6 standard deviations from the mean, there are at least 97% of the values predicting from the distribution. From the plot, one can conclude that the majority of the flows that have been observed are within this area, so the use of this approach in order to determine the demand learning is reasonable.

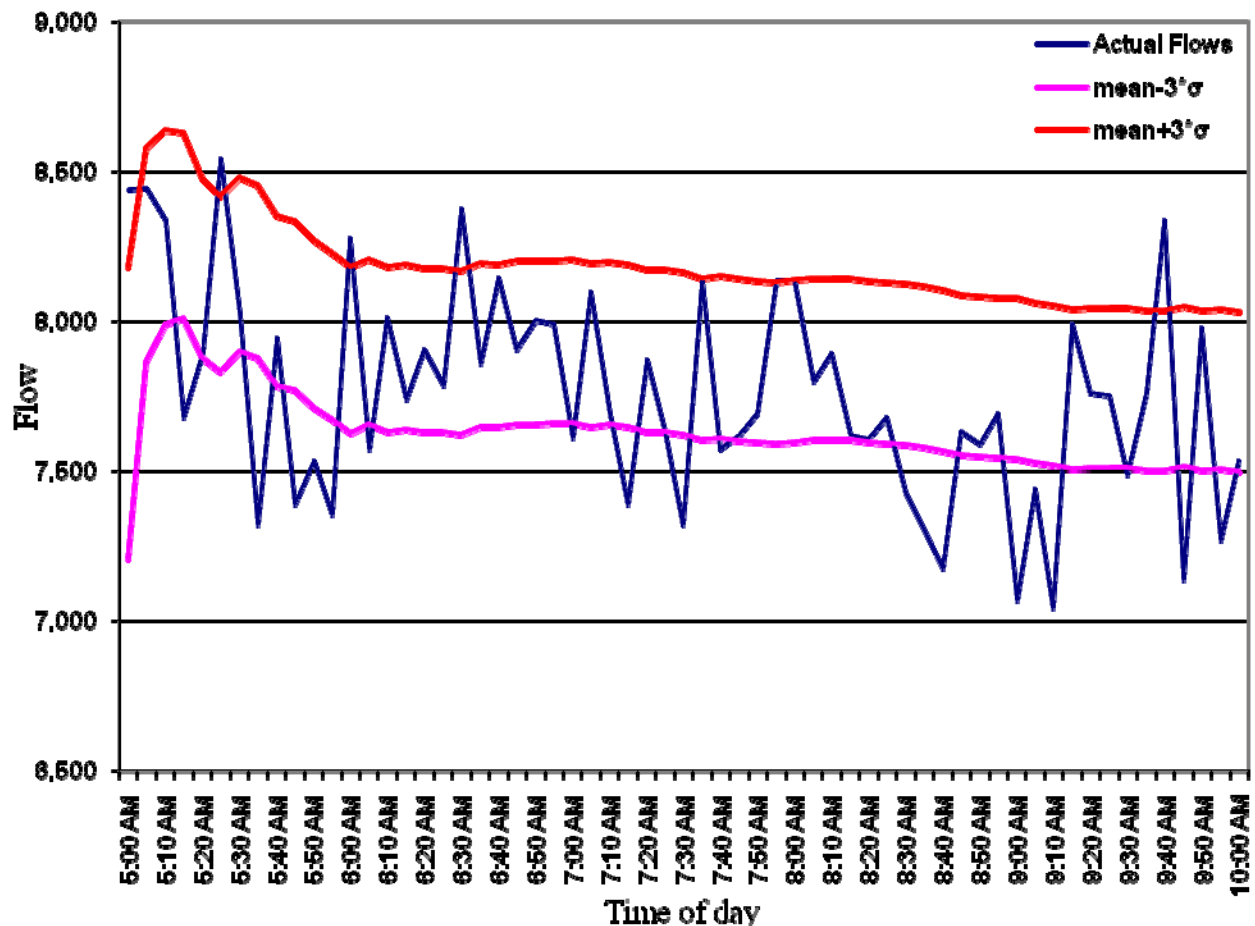


Figure 4.2. Actual flows and weaker bounds.

4.2.2 Analysis Using the Data of the Peak Hour.

The data collected are presented in table 4-3 and the analysis in table 4-4.

Table 4-4. Data from SR-91

	Total arrival flow (veh/5min)	Total arrival flow (veh/5min)	Total arrival flow (veh/5min)	Total arrival flow (veh/5min)
Time (a.m.)	9/12/2001	9/18/2001	9/19/2001	9/25/2001
5:00	8664	7968	8400	8613
5:05	8916	8616	8736	7908
5:10	8076	9384	8496	7740
5:15	8052	8352	8304	8013
5:20	7608	7920	8052	8076
5:25	7416	8352	8316	8316
5:30	7680	7968	8340	8184
5:35	7932	7956	7932	7812
5:40	8040	8064	8387	8172
5:45	7800	7680	7335	7980
5:50	7416	7512	6216	7680
5:55	7668	7968	5244	7572
6:00	7596	7245	5040	7752
Average	7913	8076	7600	7986

From table 4-4, the average μ_{Λ} and the standard deviation σ_{Λ} of the average flows are calculated.

$$\mu_{\Lambda} = 7894$$

$$\sigma_{\Lambda} = 206.88$$

The above parameters refer to 5 minutes period so the average μ_{Λ} and the standard deviation σ_{Λ} for 1 minute period are:

$$\mu_{\Lambda} = 7894 / 5 = 1578.80$$

$$\sigma_{\Lambda} = 206.88 / 5 = 41.37$$

Therefore,

$$k = \mu_{\Lambda}^2 / \sigma_{\Lambda}^2 \approx 1456$$

$$a = \mu_{\Lambda} / \sigma_{\Lambda}^2 = 0.9221$$

The number of vehicles that may arrive during time interval $(t, t + \Delta t]$ can be estimated using the negative binomial distribution shown below ($\Delta t = 5$ min):

$$P(N(t+5) - N(t) = n | N(t) = i) = \binom{1456 + i + n - 1}{n} \cdot \left(\frac{0.9221 + t}{0.9221 + t + 5} \right)^{1456 + i} \cdot \left(\frac{5}{0.9221 + t + 5} \right)^n$$

The mean and the variance for the above binomial distribution are:

$$\mu_b = (1456 + i) \frac{5}{(0.9221 + t)}$$

$$\sigma_b = (1456 + i) \frac{5 \cdot (0.9221 + t + 5)}{(0.9221 + t)^2}$$

The analysis and the estimated flows per 5 minutes are presented in table 4-4.

Table 4-5. Estimated flows

Cumulative arrivals	Actual time	Time t	Δt	Probability Mean (Estimated flows)	Variance	Standard deviation
0						
8440	5:00 AM	0.00	5.00	7894.67	50700.88	225.17
16885	5:05 AM	5.00	5.00	8355.09	15409.20	124.13
25225	5:10 AM	10.00	5.00	8396.25	12239.93	110.63
32905	5:15 AM	15.00	5.00	8378.58	11009.70	104.93
40801	5:20 AM	20.00	5.00	8211.64	10174.06	100.87
49345	5:25 AM	25.00	5.00	8150.75	9722.91	98.60
57373	5:30 AM	30.00	5.00	8214.34	9542.57	97.69
64693	5:35 AM	35.00	5.00	8188.40	9328.15	96.58
72637	5:40 AM	40.00	5.00	8082.30	9069.82	95.24
80029	5:45 AM	45.00	5.00	8067.24	8945.60	94.58
87565	5:50 AM	50.00	5.00	8000.94	8786.55	93.74
94921	5:55 AM	55.00	5.00	7959.37	8671.02	93.12
103201	6:00 AM	60.00	5.00	7909.85	8559.03	92.52

The actual and the predicted flows are illustrated in table 4-5.

Table 4-6. Actual and predicted flows

Actual Flows	Probability Mean	Standard deviation	mean-3* σ	mean+3* σ
9/26/2001				
8,440.00	7894.671	225.17	7219.17	8570.18
8,445.00	8355.086	124.13	7982.69	8727.49
8,340.00	8396.248	110.63	8064.34	8728.15
7,680.00	8378.584	104.93	8063.80	8693.37
7,896.00	8211.636	100.87	7909.04	8514.24
8,544.00	8150.754	98.60	7854.94	8446.57
8,028.00	8214.341	97.69	7921.28	8507.40
7,320.00	8188.404	96.58	7898.66	8478.15
7,944.00	8082.3	95.24	7796.59	8368.01
7,392.00	8067.241	94.58	7783.50	8350.99
7,536.00	8000.94	93.74	7719.73	8282.15
7,356.00	7959.37	93.12	7680.01	8238.72
8,280.00	7909.85	92.52	7632.31	8187.40

The following figure (figure 4-3) shows the plot of the actual flows measured every five minutes during the peak hour and the plot of the predicted (probability) mean after subtracting three standard deviations and after adding three times the standard deviation of the probability. According to the Chebyshev's inequality, within 6 standard deviations from the mean, there are at least 89% of the values predicting from the distribution.

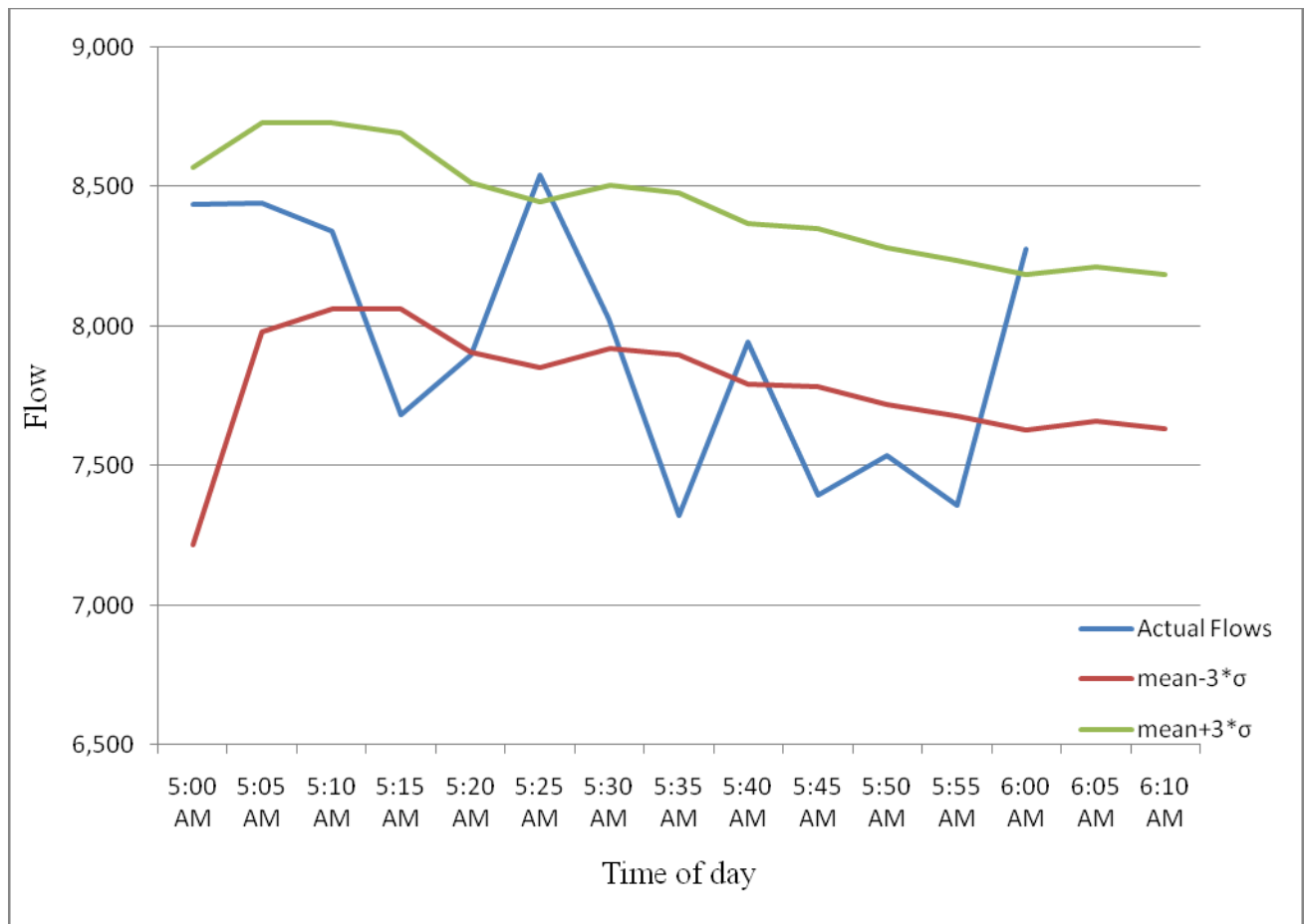


Figure 4-3. Actual flows and weaker bounds during the peak hour.

CHAPTER 5 PROACTIVE AND ROBUST PRICING STRATEGIES

5.1 Introduction

This chapter implements both the deterministic and the scenario-based toll optimization approach in order to validate and compare them. The methodology, which provides a robust and proactive approach to determine toll rates for the HOT lanes, proposed in the third chapter is applied and the results are presented.

5.2 Model Solution and Simulation Study

The toll optimization model is solved using GAMS 22.6 (GAMS Development Corporation, 2003) for pricing strategies for a freeway facility 3 miles long with a bottleneck downstream. The operation horizon is 90 minutes and other parameters assumed are as follows:

- Free-flow speed: 60 mph (or 88 ft/sec)
- Backward wave speed: 30 mph (or 44 ft/sec)
- Jam density: 120 vpm
- Saturation flow along the freeway: 2400 vphpl (or 3.33 vehicles/ 5sec)
- Saturation flow at the bottleneck: 1800 vphpl (or 2.50 vehicles/ 5sec)
- Toll values can vary from \$0.00 to \$8.00.

The propagation of traffic flow is based on the fundamental flow-density curves that are illustrated in Figure 5-1.

As previously discussed in the paragraph 3.4.1.2, toll optimization consists of two stages. In the first stage, given traffic demand, the optimal volume split between HOT and GP lanes is computed by maximizing freeway's throughput using the cell transmission model for the flow propagation. In this stage, the optimization horizon is 10 minutes while traffic propagation is done with a time step of 5 seconds. In the second stage, the optimal toll is determined for matching the optimal volume. In this stage, a value of toll is determined every 3 minutes which is a reasonable amount of time that the toll value can change in practice.

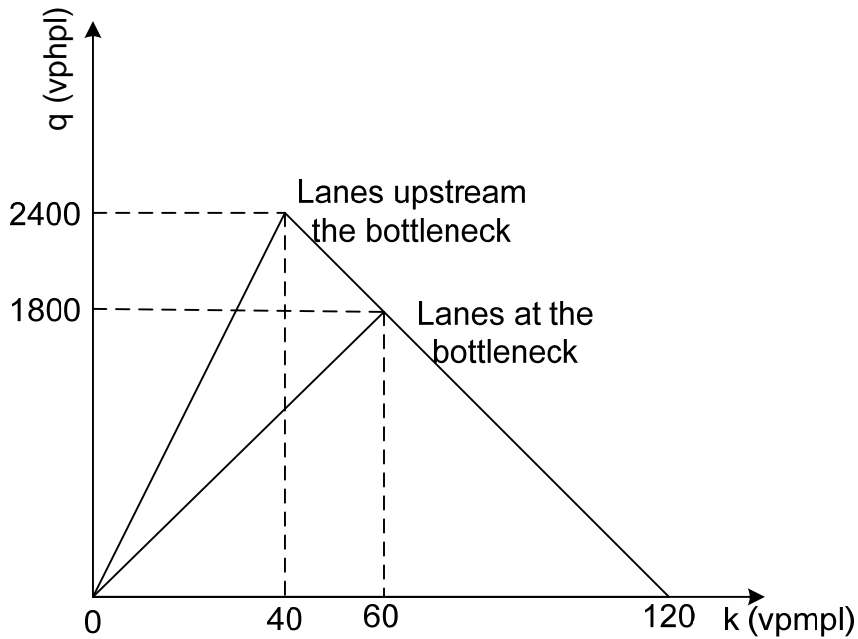


Figure 5-1. Flow-Density Curves

Both the deterministic and scenario-based toll optimization models are solved. In the former, there is only one demand scenario while 20 demand scenarios are generated for the latter. After the toll optimization, simulation experiments are conducted in order to demonstrate and validate the proposed robust toll optimization. A self-developed simulation platform is used, which consists of a monitor, a controller and a simulator.

The monitor is collecting information at each time interval including the total arrival flow, the flows at GP and HOT lanes, the densities, and the travel times. The controller solves toll optimization models using the information of the densities and flows from the monitor. The simulator attempts to replicate motorists' behavior with respect to which lane they will choose to travel. At each time interval, based on the coming flow, the instantaneous travel times provided from the monitor and the toll value obtained from the controller, the Logit model is applied by the simulator to compute how many motorists are going to choose the HOT lane versus the GP lane.

The simulation site is a 3 mile freeway segment as mentioned above. This segment has two lanes, one HOT lane and one GP lane that each of them obeys the big triangular flow-density curve shown in Figure 5-1. At the end of the segment, it is assumed that there is a bottleneck where the lanes obey the small triangular flow-density curve in Figure 5-1. All the relevant parameters are given in the Figure 5-1.

The simulation time step is 5 sec and both lanes are partitioned into small cells of 440 ft each. The values of α_1 , α_2 and γ used in the simulator are 0.5, 1, 0.2, respectively, as proposed by Lou et al (2007). Two experiments are conducted, one with low-high-low demand profile and the other low-medium-high-medium-low demand profile.

5.3 Numerical Results

The numerical results from the simulation experiments are presented in Figures 5-2 – 5-25.

5.3.1 Low-High-Low Demand Case

In the low-high-low demand case, traffic demand varies from 1200 vphpl to 2100 vphpl. More precisely, traffic demand is uniformly distributed from 2400 vph to 3200 vph for the first 15 min, from 3200 vph to 4200 vph for the next 30 min and from 2400 vph to 3200 vph for the rest of the simulation horizon (Figure 5-2).

5.3.1.1 Robust toll optimization

The Robust toll optimization is conducted by generating 20 different demand scenarios. Each scenario is associated with a probability of 5%. Figure 5-3 presents the optimal toll rates determined by the controller and Figure 5-4 reports the resulting freeway, HOT and GP throughputs.

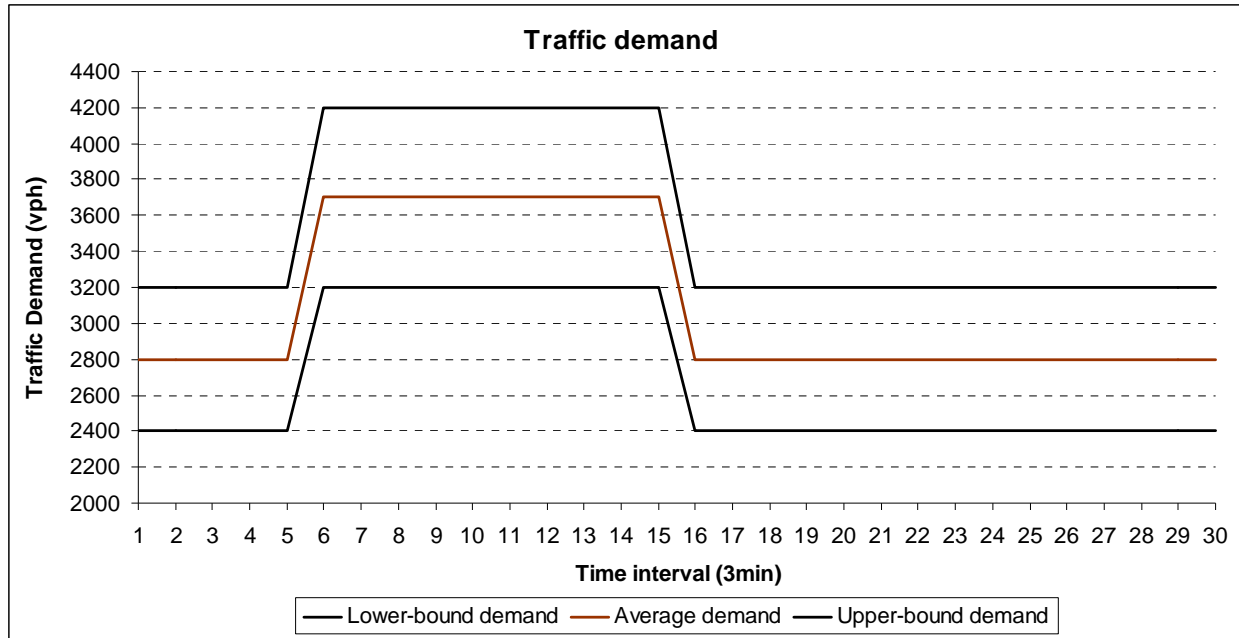


Figure 5-2. Traffic demand profile.

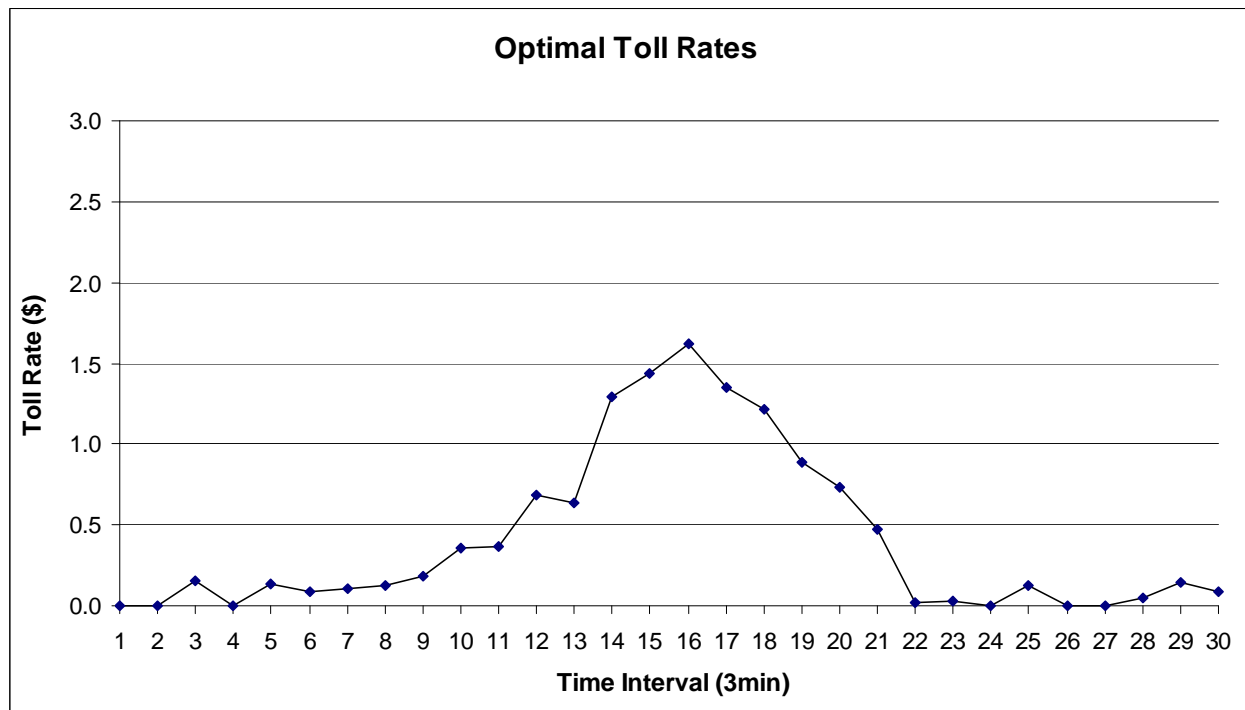


Figure 5-3. Optimal Toll Rates.

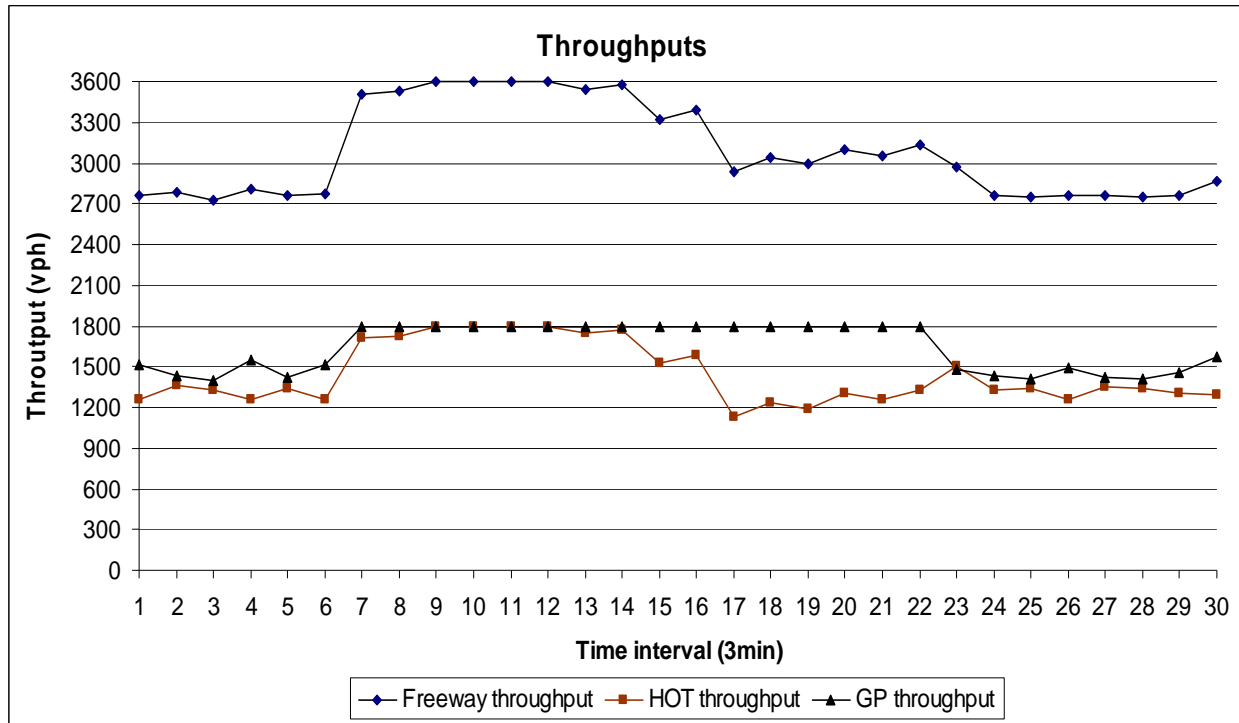


Figure 5-4. Freeway, HOT and GP throughputs.

It can be seen that the controller is able to apply a toll value to maintain a high throughput of the freeway while toll value is changing in a smooth manner. For an incoming demand from 1200 vph to 2100 vph for each lane, tolls are taking values from \$0.00 to \$1.62. When there is no difference in travel time between the HOT lane and the GP lane, toll implementation is not necessary so the toll takes the value of \$0.00. The average throughput is 3084, 1442 and 1643 vph for the freeway, the HOT and the GP lane, respectively. When the demand is high, between time intervals 6 and 15, the average of the freeway's throughput is 3464 vph, HOT lane's throughput 1693 vph and GP lane's throughput is 1771 vph. This indicates that the freeway can be managed well with the proposed approach.

Figure 5-5 shows the average densities along the HOT and GP lanes, and figure 5-6 presents the queue length upstream of the bottleneck at HOT and GP lane.

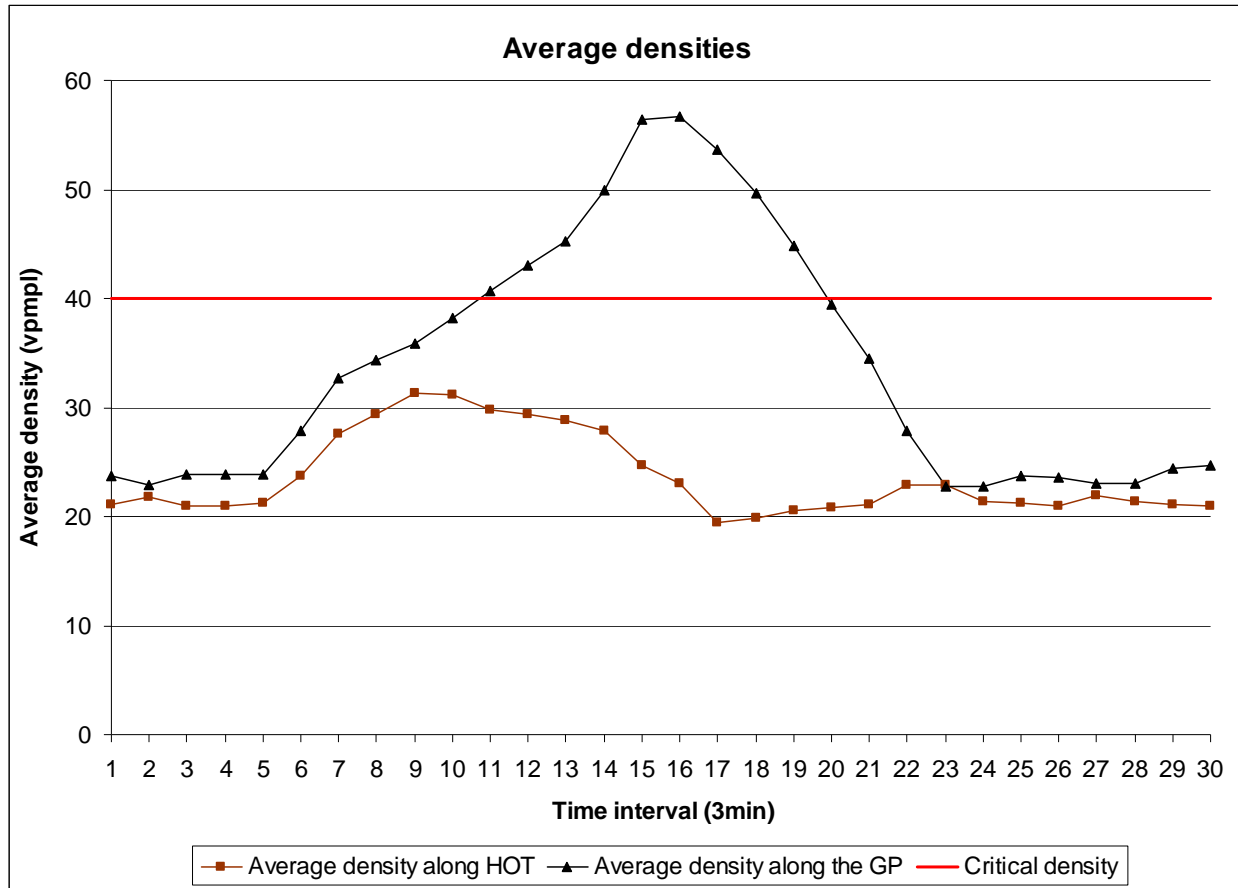


Figure 5-5. Average densities along HOT and GP lanes.

The average density for the HOT lane is 23.29 vpmp/l while for the GP lane is 33.47 vph which means that the HOT lane operates with higher speed for the vehicles. In addition, from figure 5-5, it can be observed that the average density of the GP lane is over the critical density for 9 time intervals or 27 min while the average density of the HOT lane never exceeds the critical density. This is important because if the density of one lane exceeds the critical value, it does not operate under free flow speed anymore but with lower speed depending on the flow-density curve (Figure 5-1).

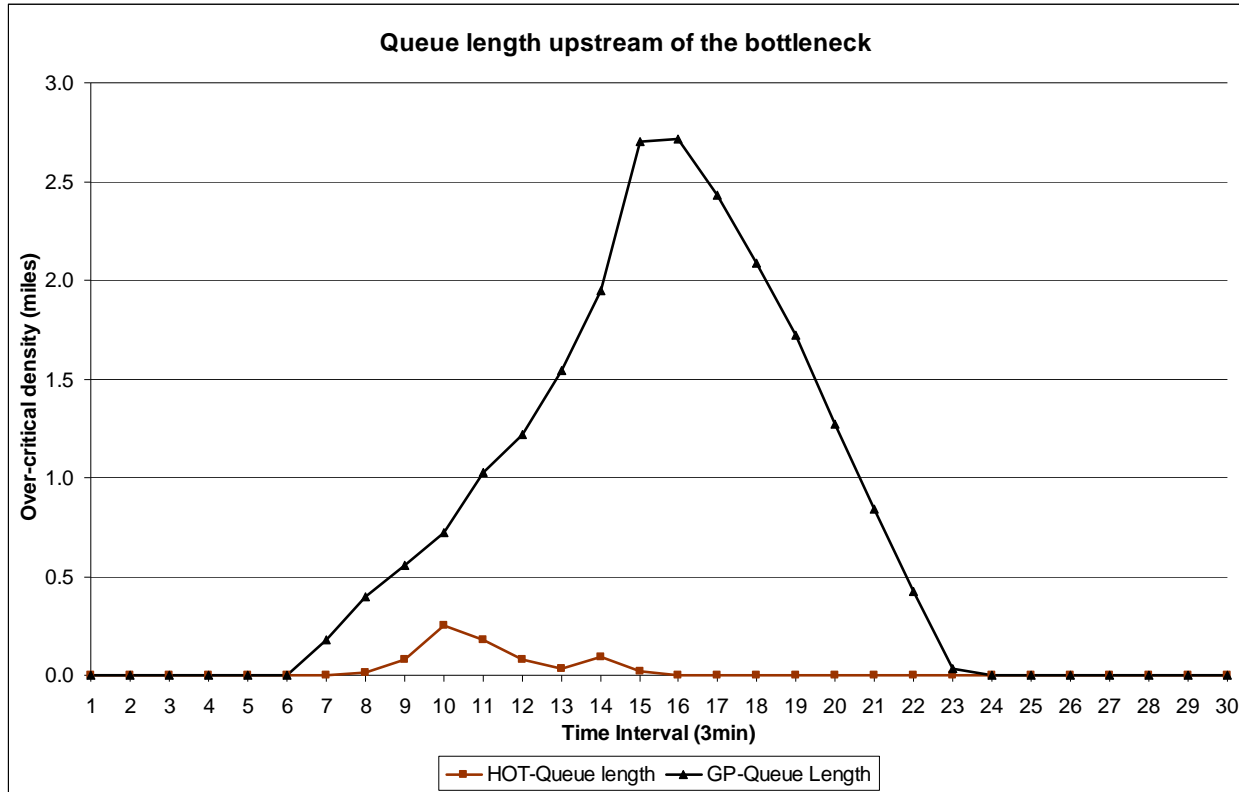


Figure 5-6. Queue length upstream of the bottleneck at HOT and GP lane.

As the above Figure illustrates, the proposed approach achieves very well the desired operating conditions at the HOT lane. Although the demand is very high for half hour, only a small queue is developed at HOT lane upstream the bottleneck, due to the random demand. The maximum queue length is 0.3 miles. However, the queue in the GP lane may extend up to 2.7 miles because the total arrival is beyond the freeway's capacity.

5.3.1.2 Deterministic toll optimization

The deterministic toll optimization is conducted by assuming that the incoming flow will be 1400 vphpl for the first 15 minutes, 1850 vphpl for the next 30 minutes and then 1400 vphpl for the rest of the experiment horizon, which are the averages of random traffic demands assumed (Figure 5-2).

Figure 5-7 presents the optimal toll rates determined by the controller and Figure 5-8 shows the freeway, HOT and GP throughputs.

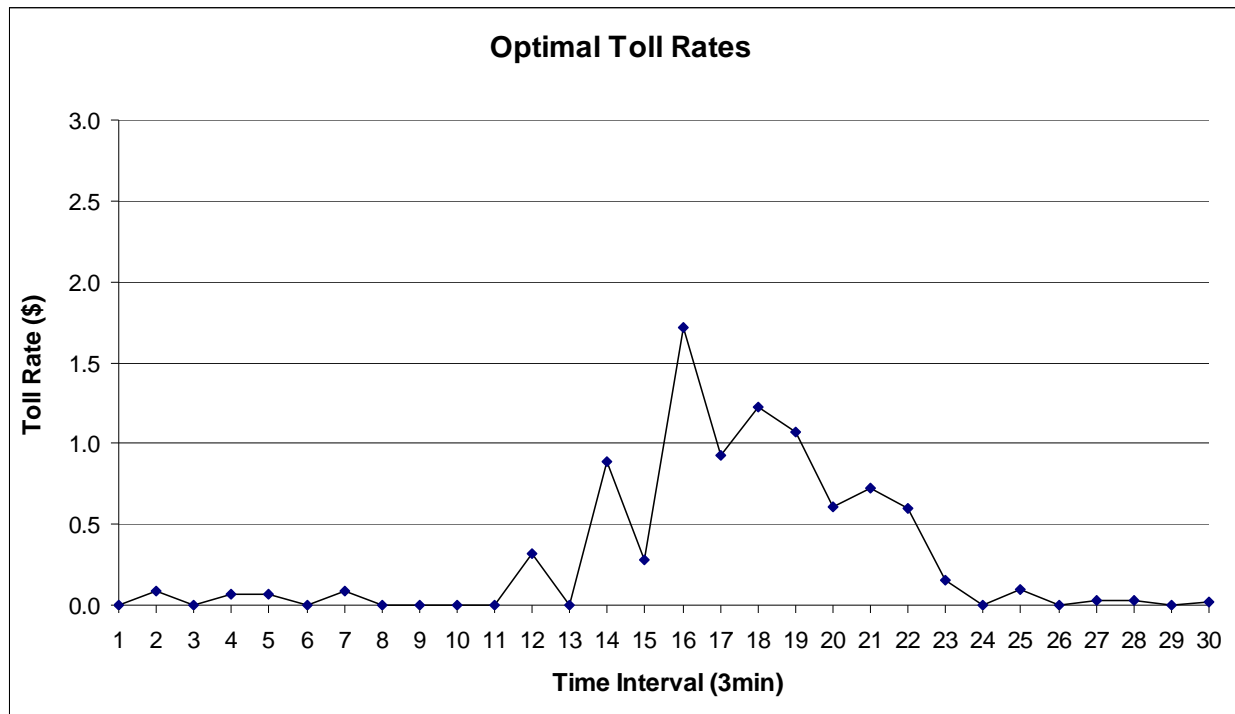


Figure 5-7. Optimal Toll Rates.

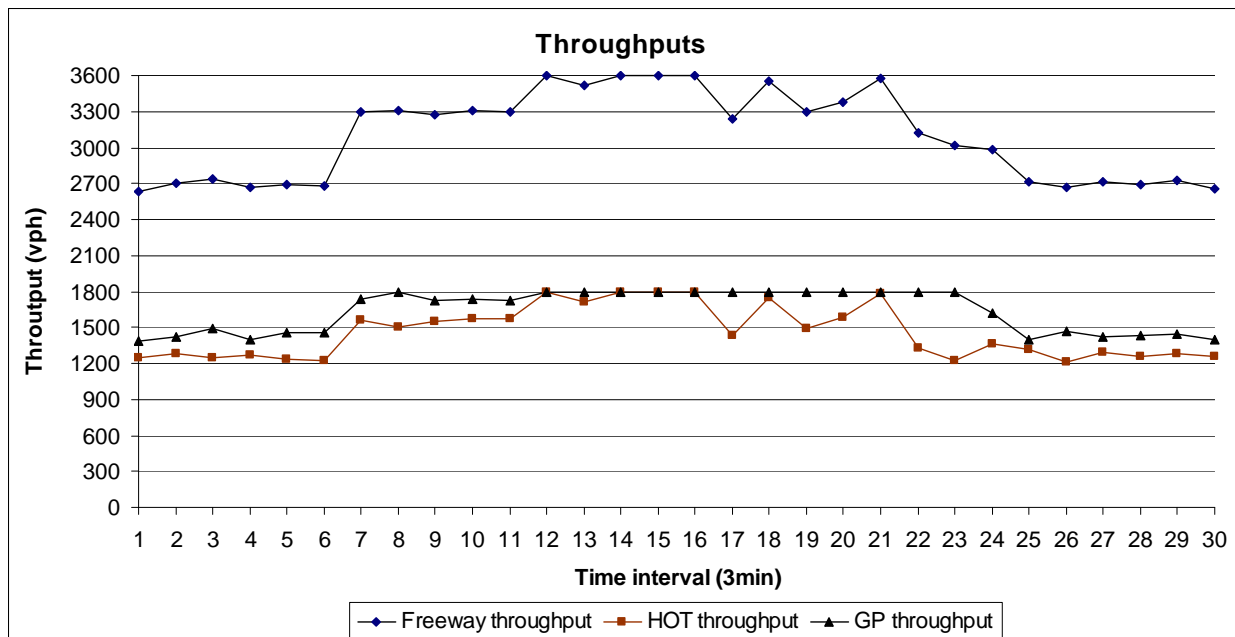


Figure 5-8. Freeway, HOT and GP throughputs.

As demonstrated in Figure 5-7, the controller is able to maintain a high throughput of the freeway with a toll varying from \$0.00 to \$1.71. The average throughput is 3095, 1458 and 1637 vph for the freeway, the HOT and the GP lane, respectively (Figure 5-8).

Figure 5-9 illustrates the average densities along the HOT and GP lanes. The average density for the GP lane is 32.84 vpmpl while for the HOT lane is 24.10 vph. However, GP has density over the critical for 24 min out of the 90 min experiment horizon, suggesting that it is congested and vehicles travel with less speed than the free flow speed.

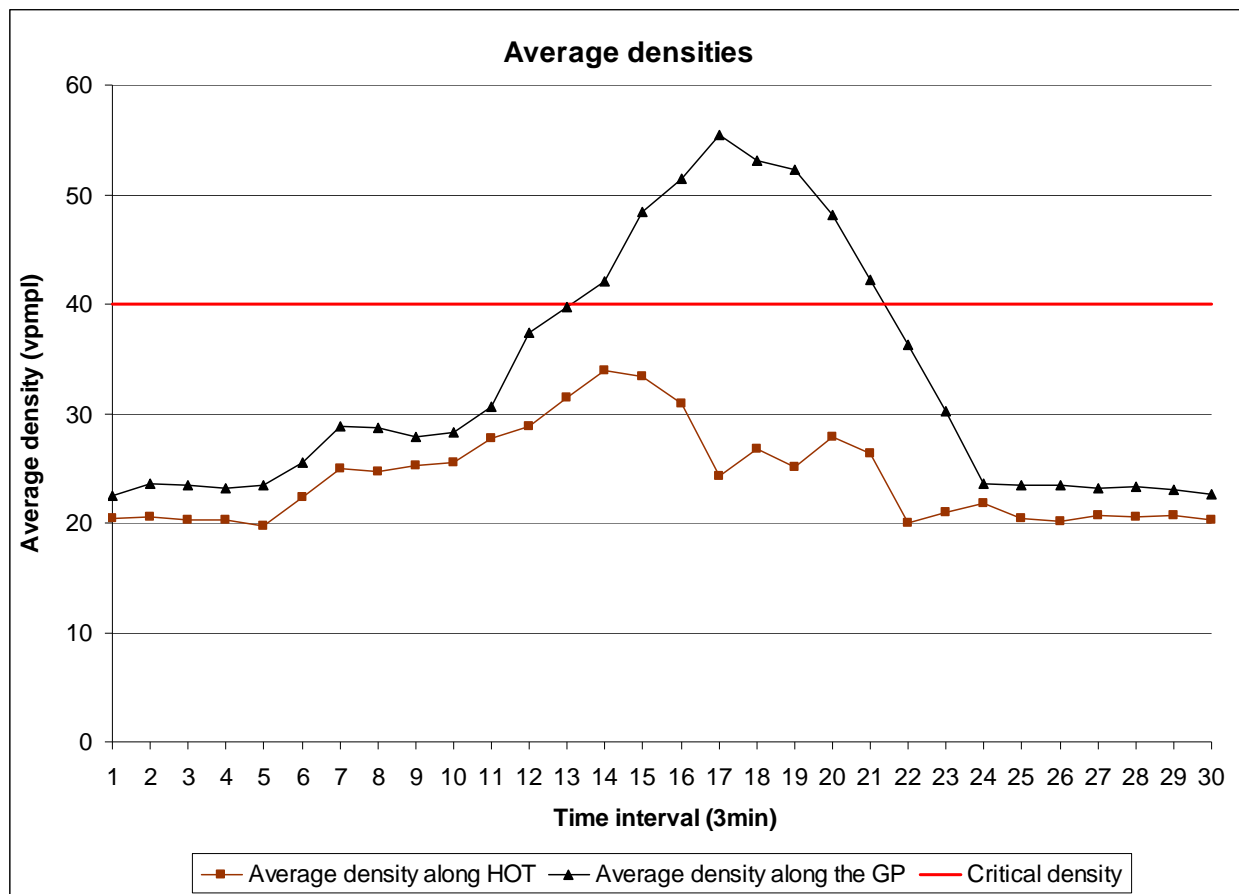


Figure 5-9. Average densities along HOT and GP lanes.

Figure 5-10 shows the queues developed at HOT and GP lanes when the demand is over the capacity. In this case, the HOT lane operates really well comparatively to GP lane that has long queues because no managing strategy is applied to it.

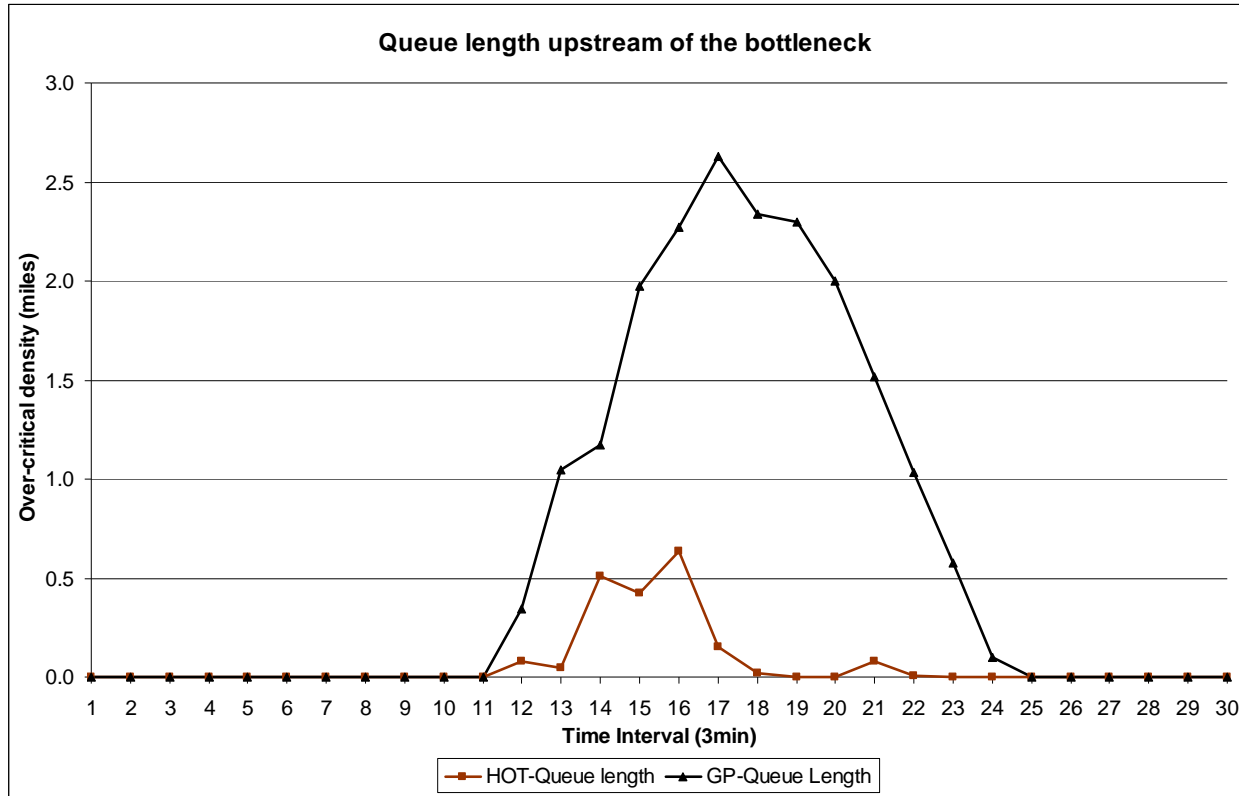


Figure 5-10. Queue length upstream of the bottleneck at HOT and GP lane.

5.3.1.3 Comparison of the robust versus the deterministic toll optimization

The results presented above show some performance differences between the two proposed approaches. To better compare the results, diagrams that merge the results from the two approaches for the toll rates and the HOT throughput, density and queue length are provided as follows.

Figure 5-11 compares the optimal toll rates.

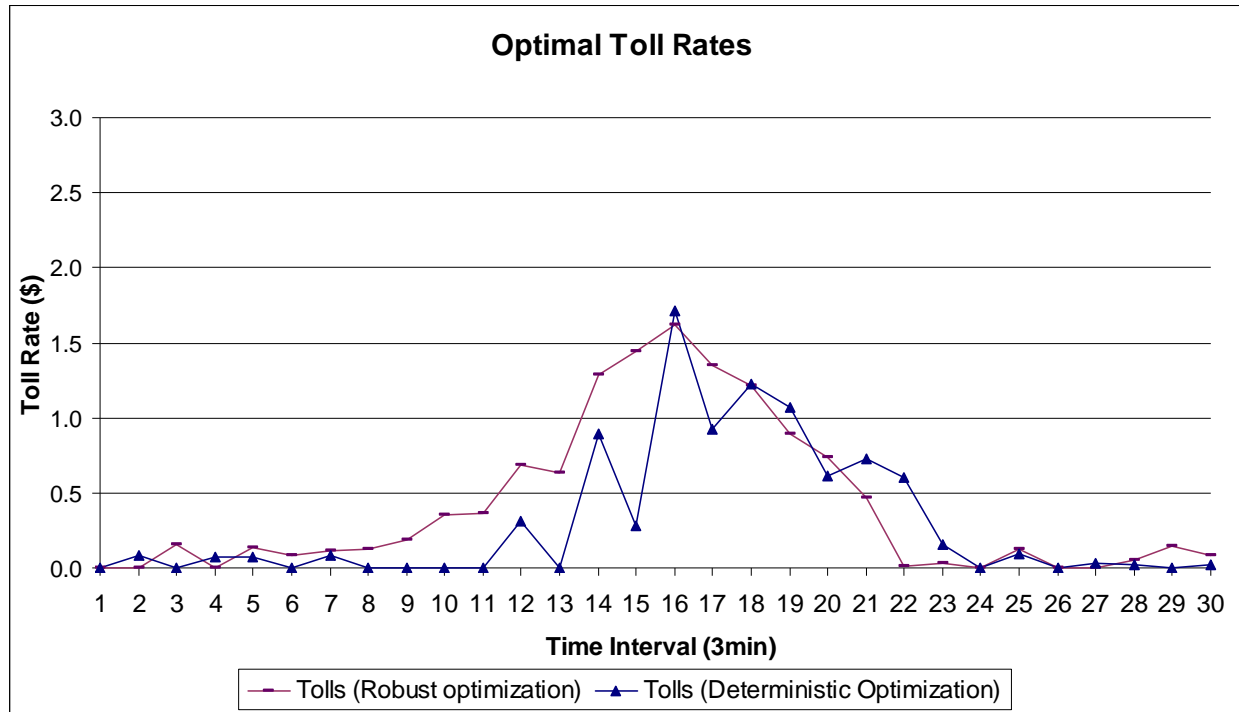


Figure 5-11. Optimal Toll Rates.

At robust optimization, the toll rates are more stable, change in smoother manner. They reach gradually a maximum value that is less than the maximum value under the deterministic case. In addition, the profit at the robust case is \$18,835 for the 90 min under analysis while it is \$14,844 for the deterministic case. This means that at robust optimization the profit is larger.

Figure 5-12 shows the HOT lane's throughput for the two approaches. During the 30 min that the coming demand is over the capacity, the robust approach maintains a slightly more stable and higher HOT throughput than the deterministic one. This can be seen by computing the standard deviation of HOT throughput for the time intervals 6 to 15. The standard deviation for the robust approach is 175.35 while for the deterministic case is 178.41 vph. The average throughput is 1458 vph and 1441 vph respectively. However, those differences are not statistically significant.

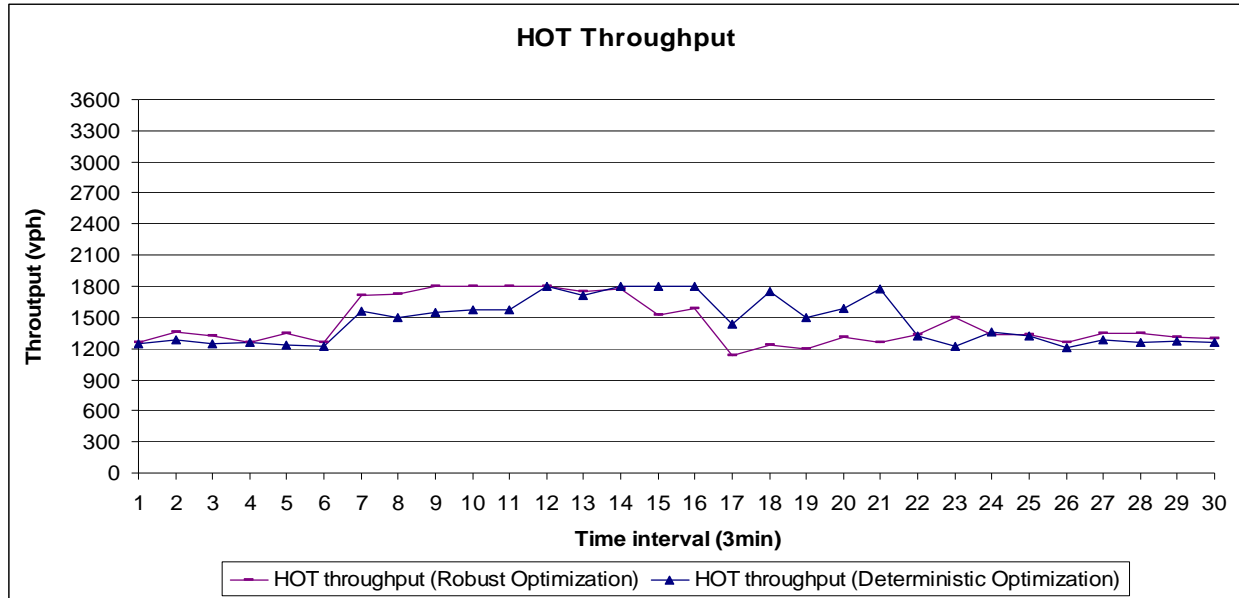


Figure 5-12. HOT throughput.

Figure 5-13 provides the density along the HOT lane. In this case the standard deviation for the robust approach is 2.51 vs. 3.95 for the deterministic approach. If the density does not fluctuate much, the performance of the HOT lane can be better.

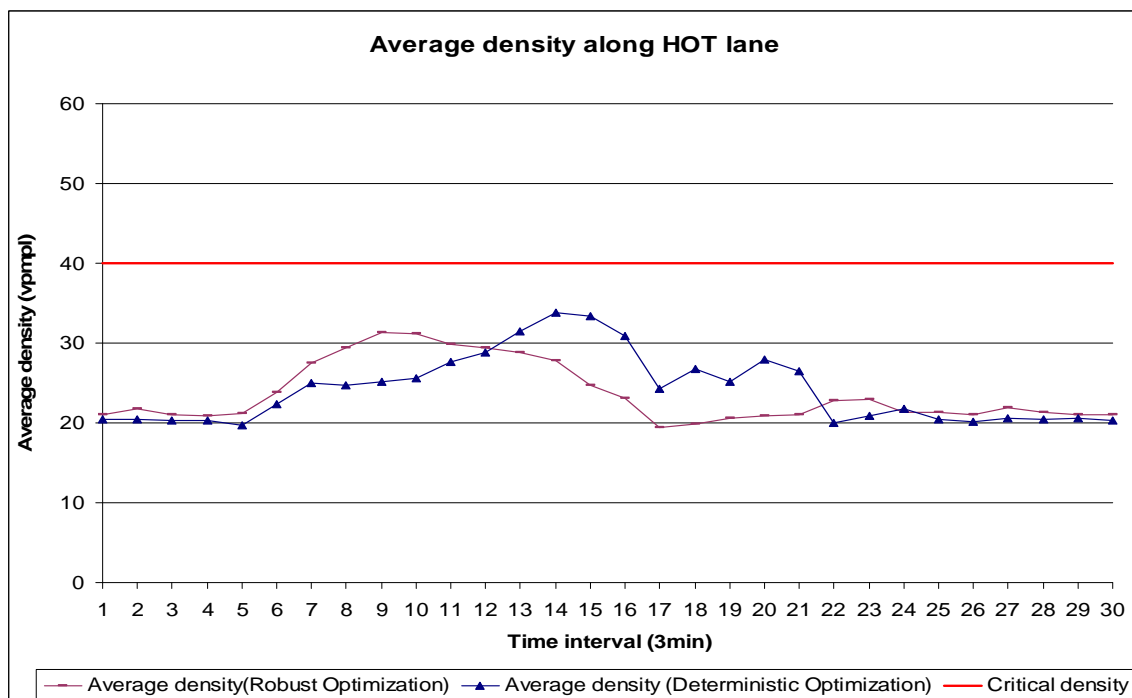


Figure 5-13. Average density along HOT lane.

Figure 5-14 illustrates the queues developed at HOT lane. At the deterministic case, the queue is longer which indicates that the lane does not perform so well as in the robust case.

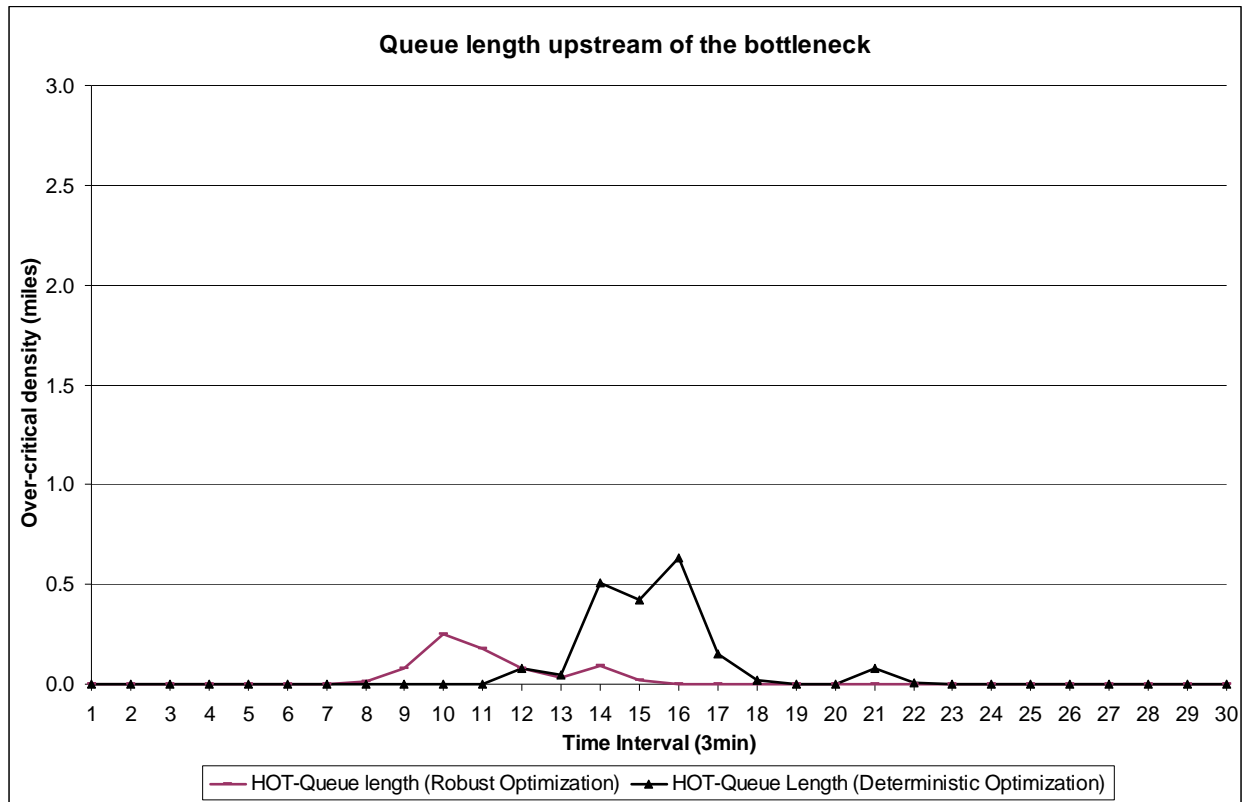


Figure 5-14. Queue length upstream of the bottleneck at HOT lane.

To summarize, the results of the two proposed approaches showed that:

- The robust approach leads to more smooth and stable toll pattern.
- The profit at the robust case is larger.
- There is not much difference at the throughput of the HOT lane.
- The average density is less than the critical density in both approaches but at the scenario-based case it varies less.
- The queue is longer at the deterministic case.

Overall, the robust approach gives slightly better results but both approaches perform really well in preventing the HOT lane for being congested while keeping the throughput in high levels.

5.3.2 Low-Medium-High-Medium-Low demand case

In the low-medium-high-medium-low demand case, traffic demand is uniformly distributed from 1200 vph to 1500 vph for the first 15min, from 1500 vph to 1800vph for the next 15 min, from 1800 vph to 2100 vph for the next 15 min, from 1500 vph to 1800vph for the following 15 min, and from 1200 vph to 1500 vph for the last 30 min of the simulation horizon (Figure 5-15).

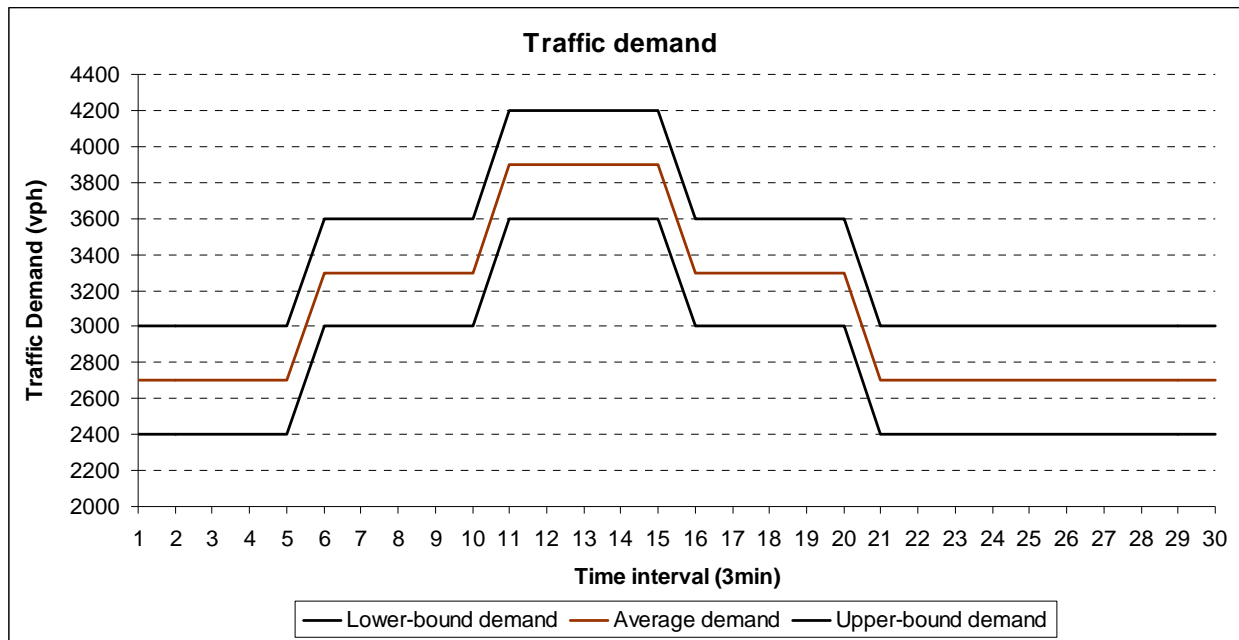


Figure 5-15. Traffic demand profile.

5.3.2.1 Robust toll optimization

Twenty different demand scenarios with a probability of 5% are generated for the robust approach. Figure 5-16 shows the optimal toll rates determined by the controller.

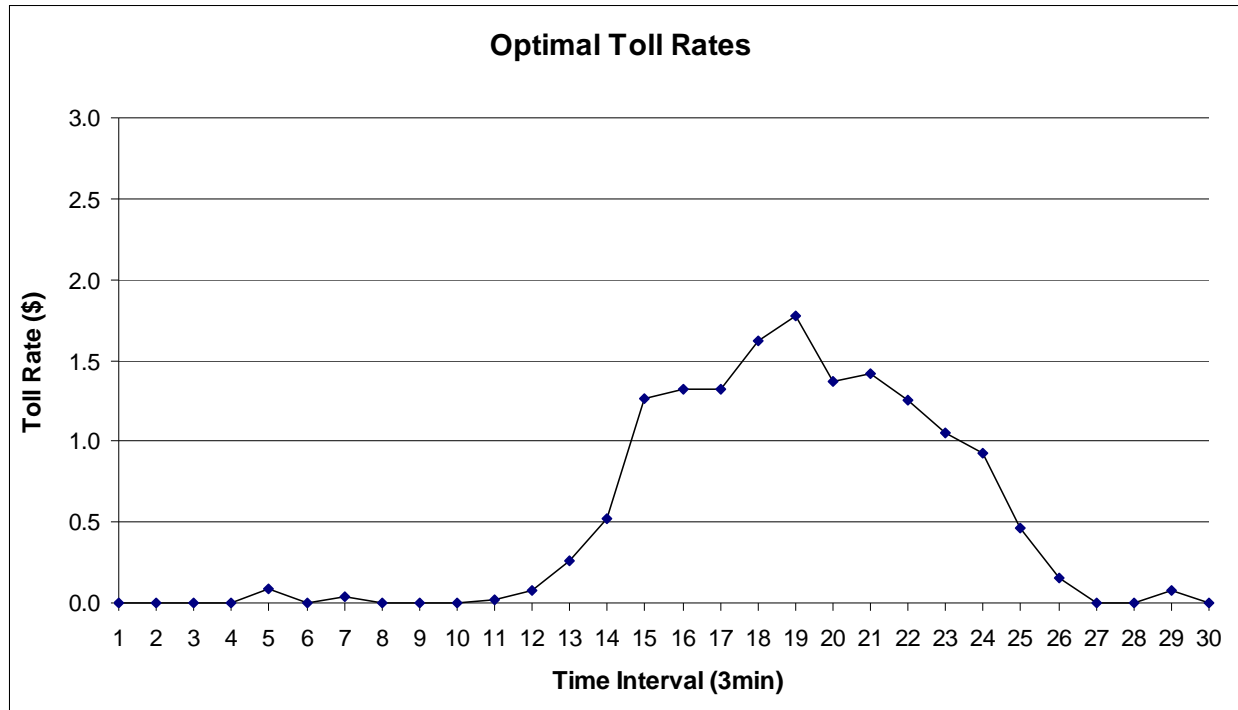


Figure 5-16. Optimal Toll Rates.

Figure 5-17 presents the freeway, HOT and GP throughputs.

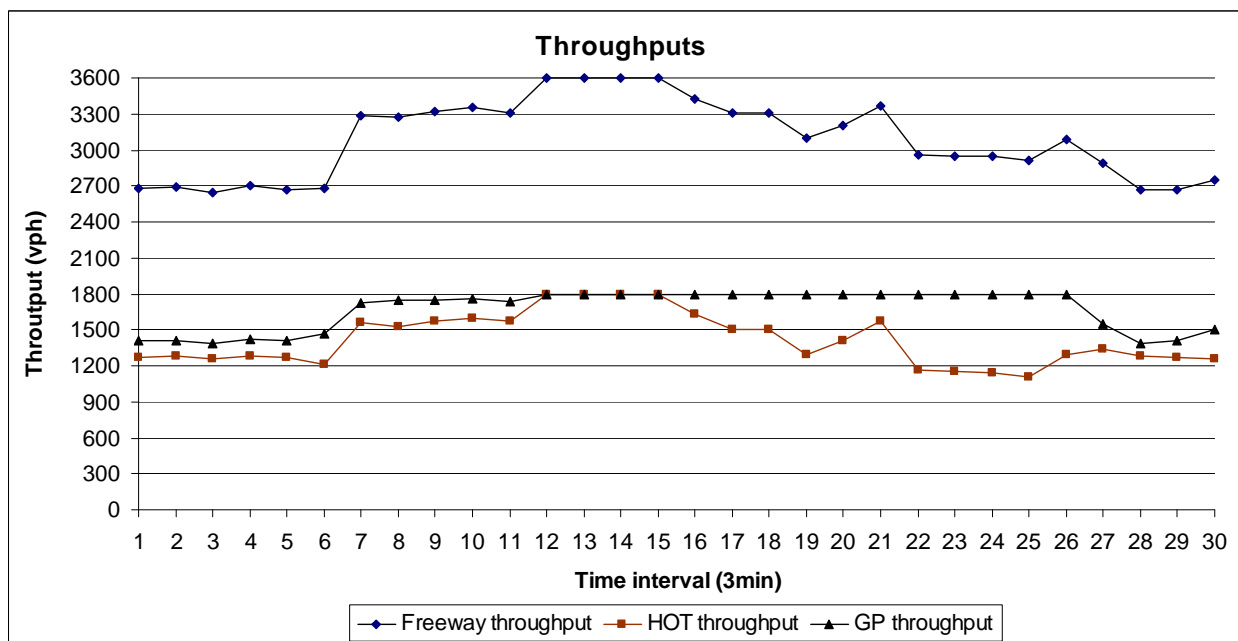


Figure 5-17. Freeway, HOT and GP throughputs.

From the Figures 5-16 and 5-17, it can be observed that the controller is able to apply a toll value to maintain a high throughput of the freeway while toll values are changing smoothly. For a coming flow from 1200 vph to 2100 vph for each lane, the lowest toll value is \$0.00 when the flow is less than the capacity for both lanes while the highest value is \$1.78 when the demand is very high. The average throughput after applying the toll rates is 3086, 1417 and 1668 vph for the freeway, the HOT and the GP lane, respectively.

Figure 5-18 presents the average densities along the HOT and GP.

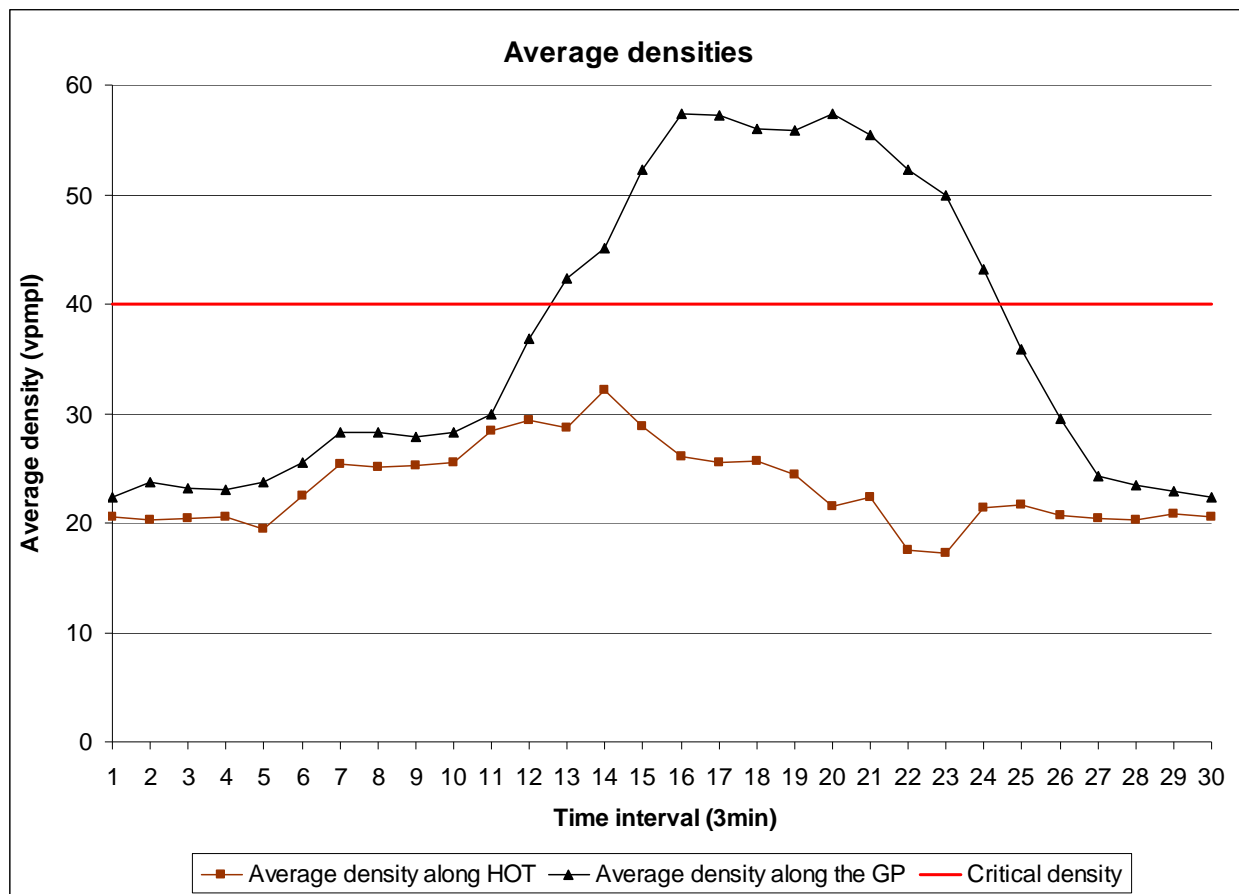


Figure 5-18. Average densities along HOT and GP lanes.

The average density for the HOT lane is 22.91 vpmp/l while for the GP lane is 35.70 vph. Furthermore, Figure 5-18 illustrates that the average density of the GP lane is over the critical density for 33 min. At the same time, the average density of the HOT lane never exceeds the

critical density. This is important because when the average density of one lane exceeds the critical value means that this lane does not operate under free flow speed in more than the half time.

Figure 5-19 shows the queue length upstream of the bottleneck at HOT and GP lane

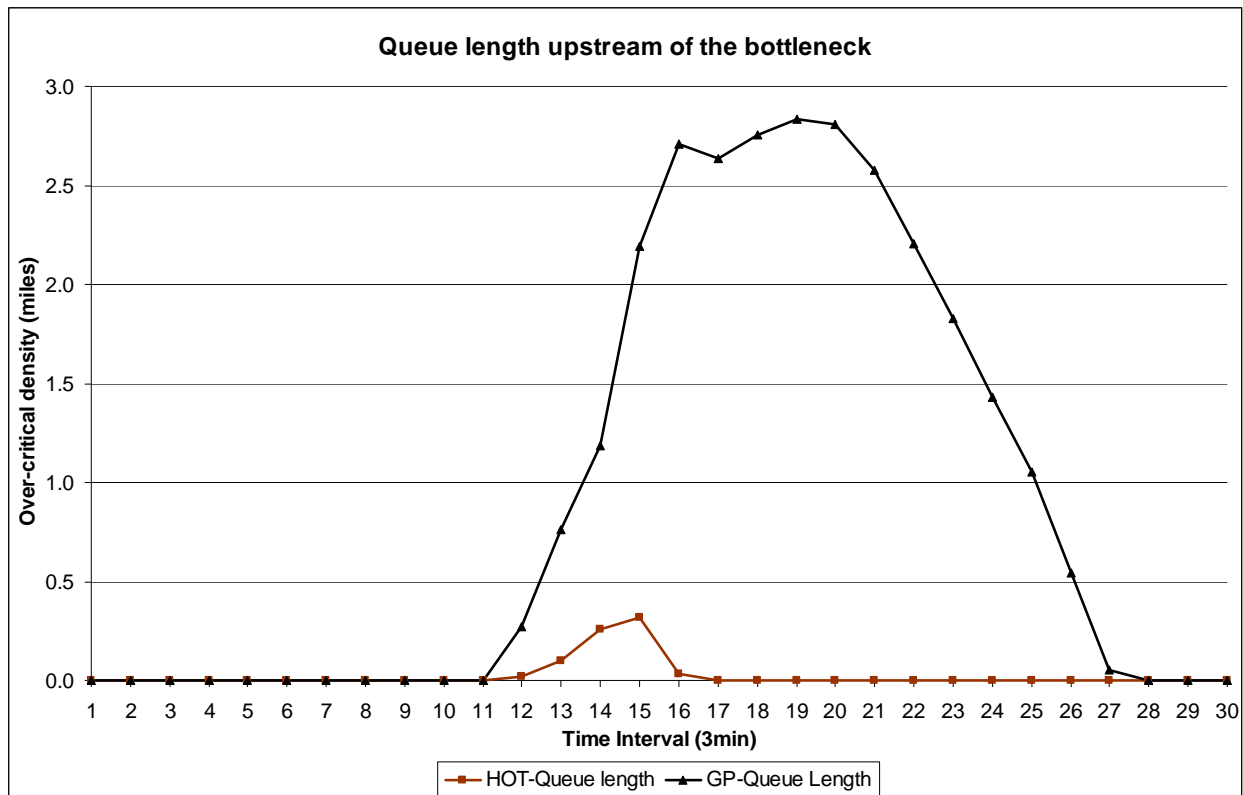


Figure 5-19. Queue length upstream of the bottleneck at HOT and GP lane.

The queue at HOT lane does not exceed the 0.32 miles upstream the bottleneck although the total demand is beyond the total capacity for 30 min. On the other hand, the queue in the GP lane reached 2.83 miles.

5.3.2.2 Deterministic toll optimization

The deterministic toll optimization in this case is conducted by assuming that the traffic demand will be 1350 vphpl for the first 15 min, 1650 vphpl for the next 15 min, 1850vphpl for the following 15 min, 1650 vphpl for the other 15 min and 1350 vphpl for the last 30 min for the

rest of the experiment horizon, with are the averages of the random traffic demands assumed (Figure 5-15).

Figure 5-20 reports the optimal toll values given by the controller and Figure 5-21 presents the freeway, HOT and GP throughputs.

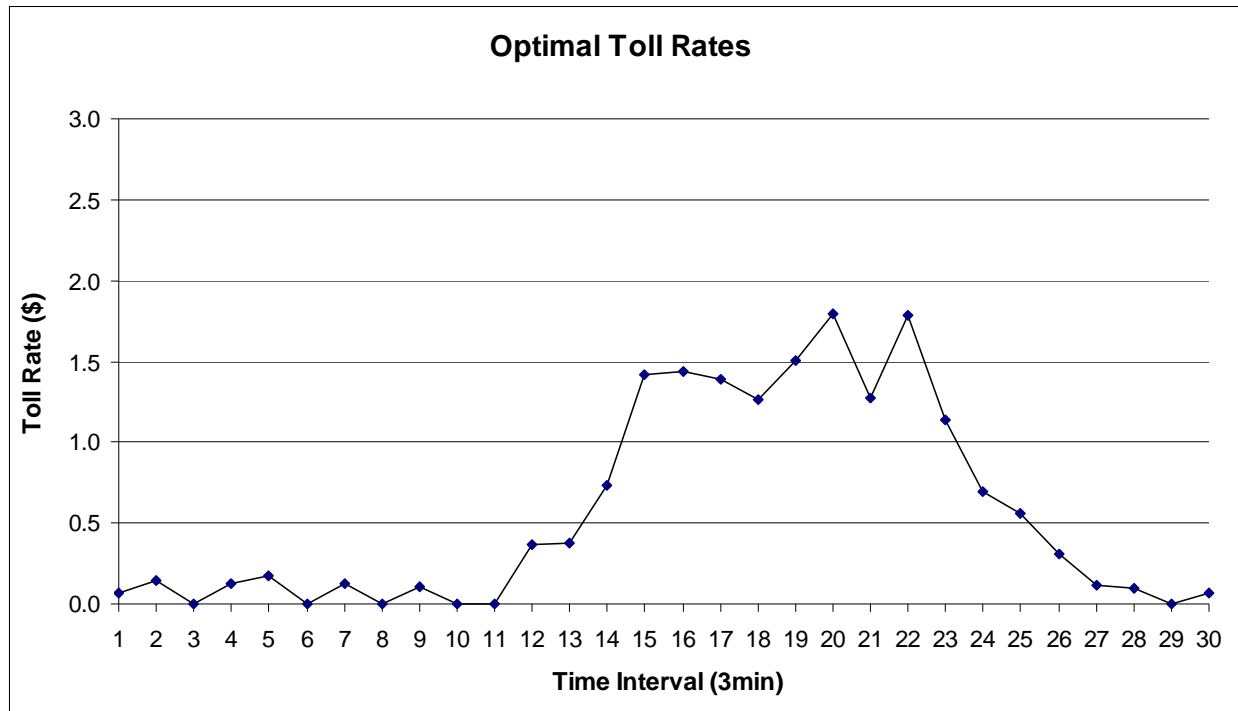


Figure 5-20. Optimal Toll Rates.

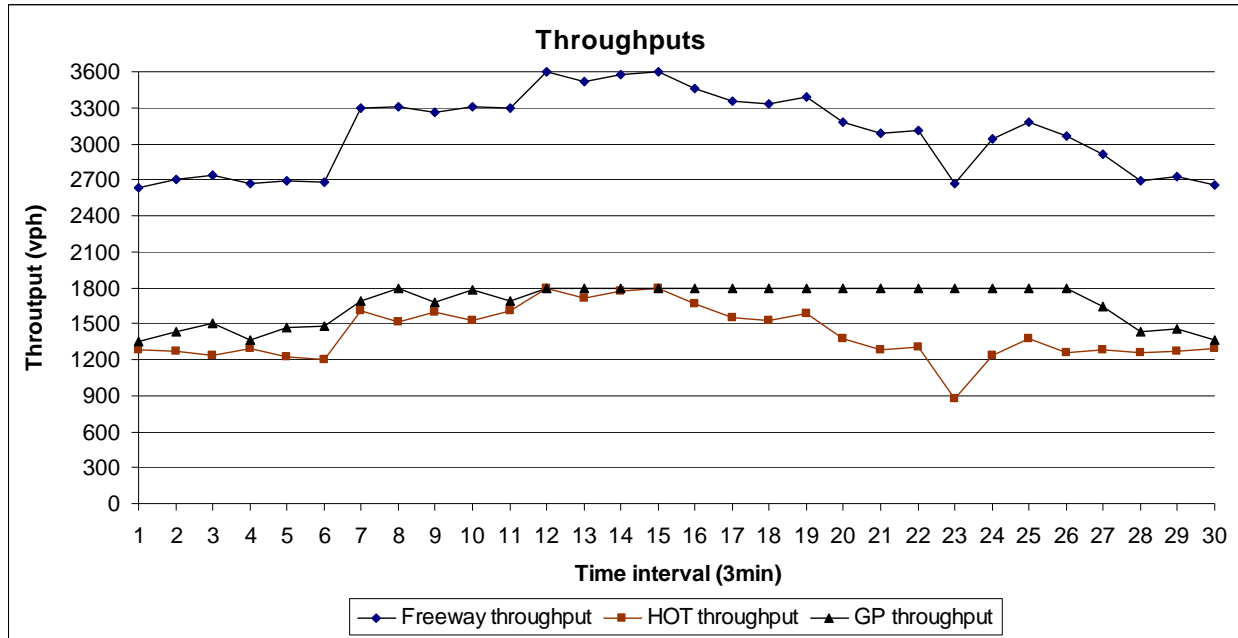


Figure 5-21. Freeway, HOT and GP throughputs.

The controller applies a toll value to maintain a high throughput of the freeway while keeping HOT lane uncongested. Toll values, in this case, are taking values from \$0.00 to \$1.79 and the average throughput is 3091, 1420 and 1671 vph for the freeway, the HOT and the GP lane, respectively.

Figures 5-22 provides the average densities along the HOT and GP lanes. The average density for the GP lane is 36.78 vpmpl while for the HOT lane is 23.28 vpmpl. However, GP has average density over the critical density for 12 time intervals or in other words for 36 min out of the 90 min experimental horizon.

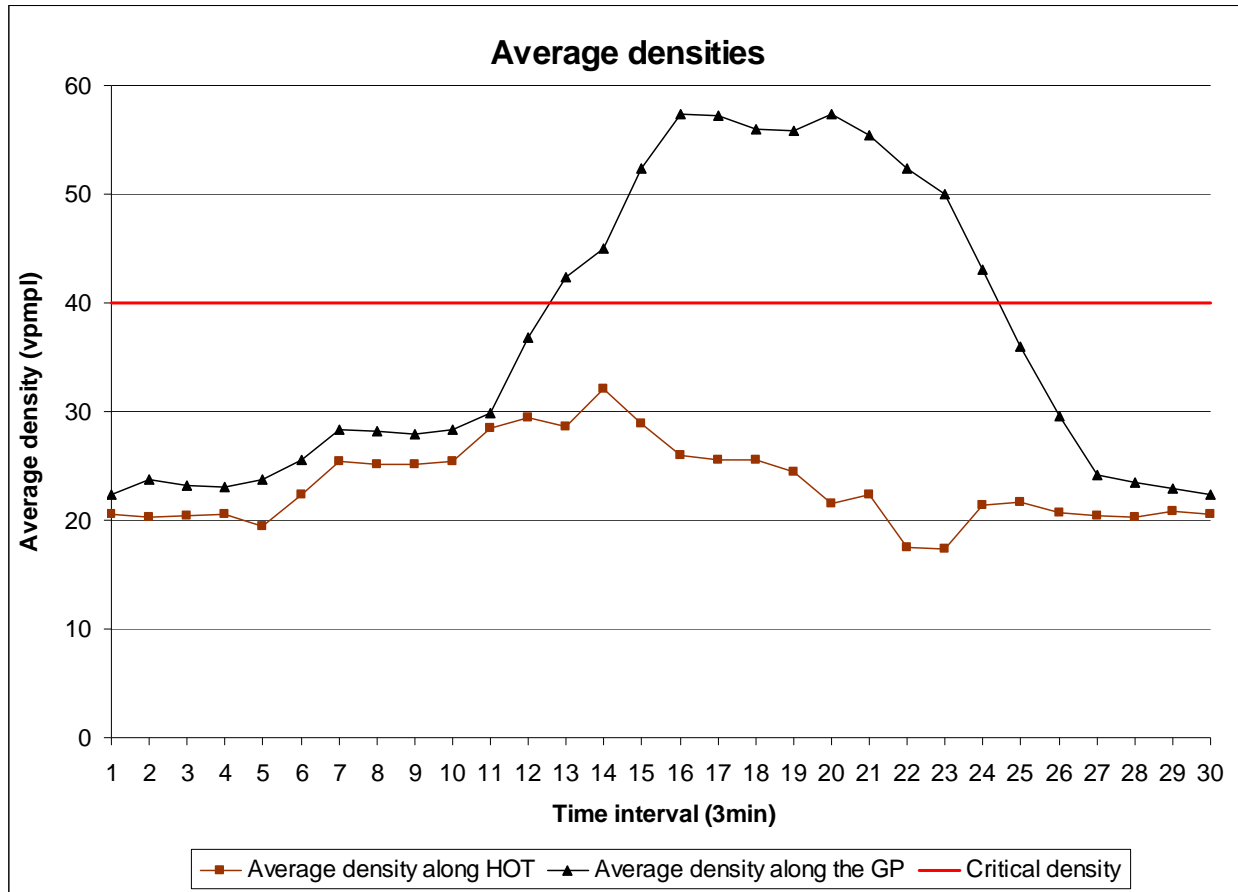


Figure 5-22. Average densities along HOT and GP lanes.

Figure 5-23 illustrates the extent of the HOT and GP lane queues upstream of the bottleneck. As in the robust optimization, HOT lane has a short queue and GP experiences long queues. This shows that the HOT lane operates under good traffic conditions.

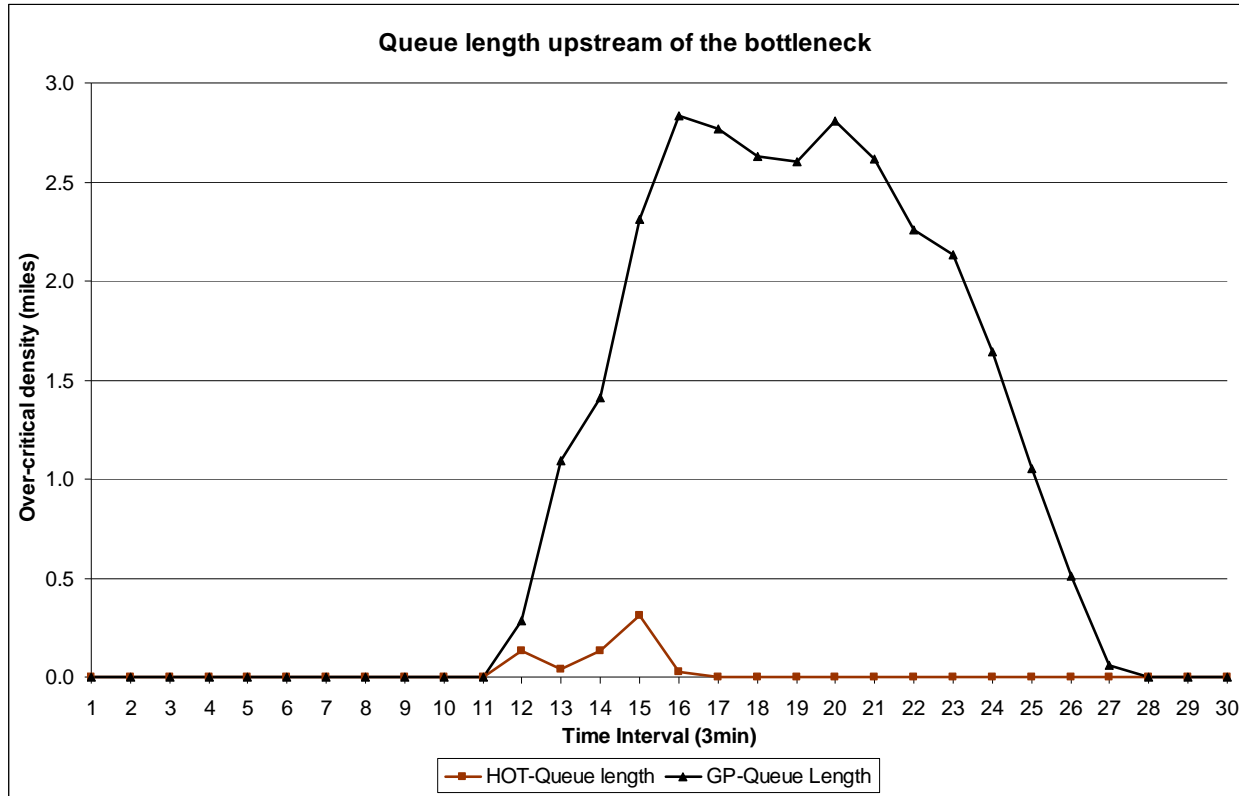


Figure 5-23. Queue length upstream of the bottleneck at HOT and GP lane.

5.3.2.3 Comparison of the robust versus the deterministic toll optimization

The results illustrated above show some performance differences between the two proposed approaches. To better compare the results, diagrams that merge the results from the approaches for the toll rates and the HOT throughput, density and queue length upstream of the bottleneck are given as follows.

Figure 5-24 presents the optimal toll rates for the two approaches.

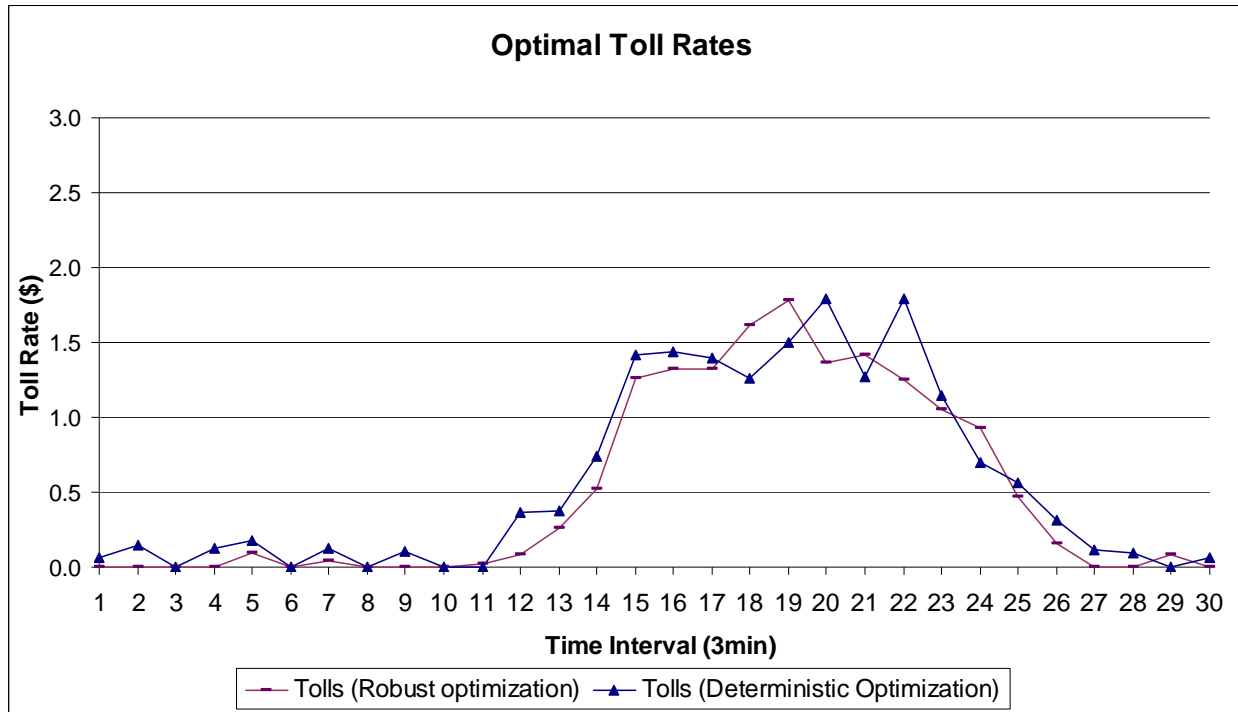


Figure 5-24. Optimal Toll Rates.

At robust optimization, the toll rates increase progressively and do not fluctuate much as at the deterministic case. However, the profit in this case is \$22,493 which is less than \$25,605 which is the profit at the deterministic case.

Figure 5-25 provides the HOT lane's throughput for the two approaches.

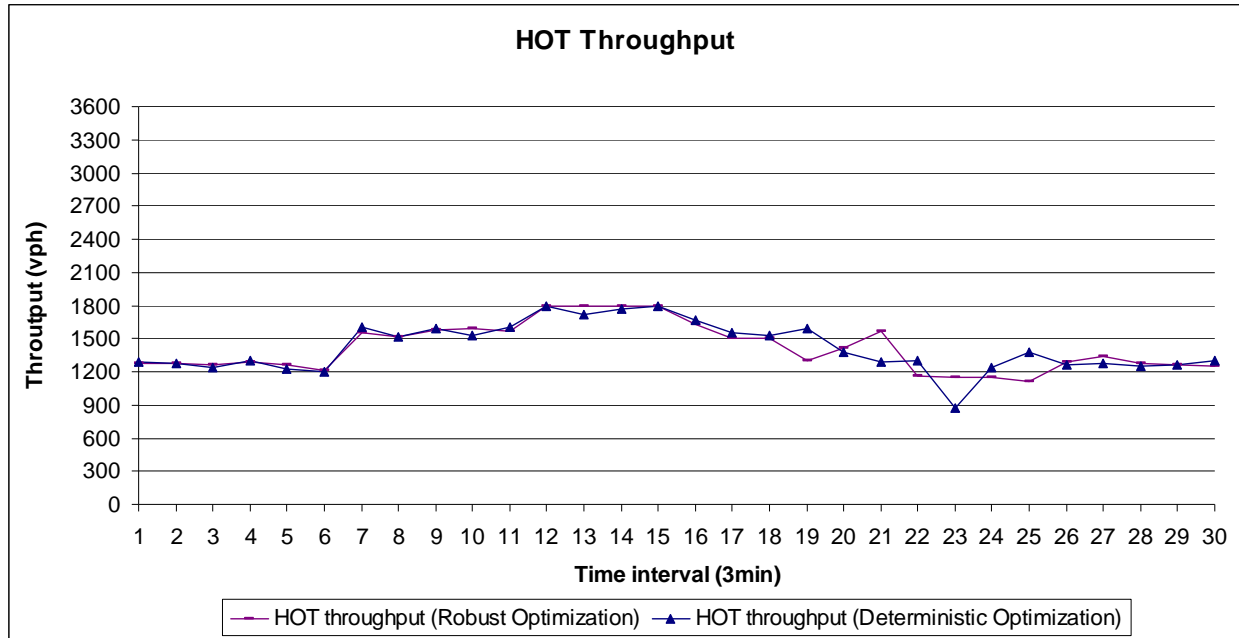


Figure 5-25. HOT throughput.

In this case, the average throughput as well as the standard deviation of the HOT lane is almost the same.

Figure 5-26 shows the density along the HOT lane. In both approaches the average density is less than the critical and there is not great difference on how the average density changes.

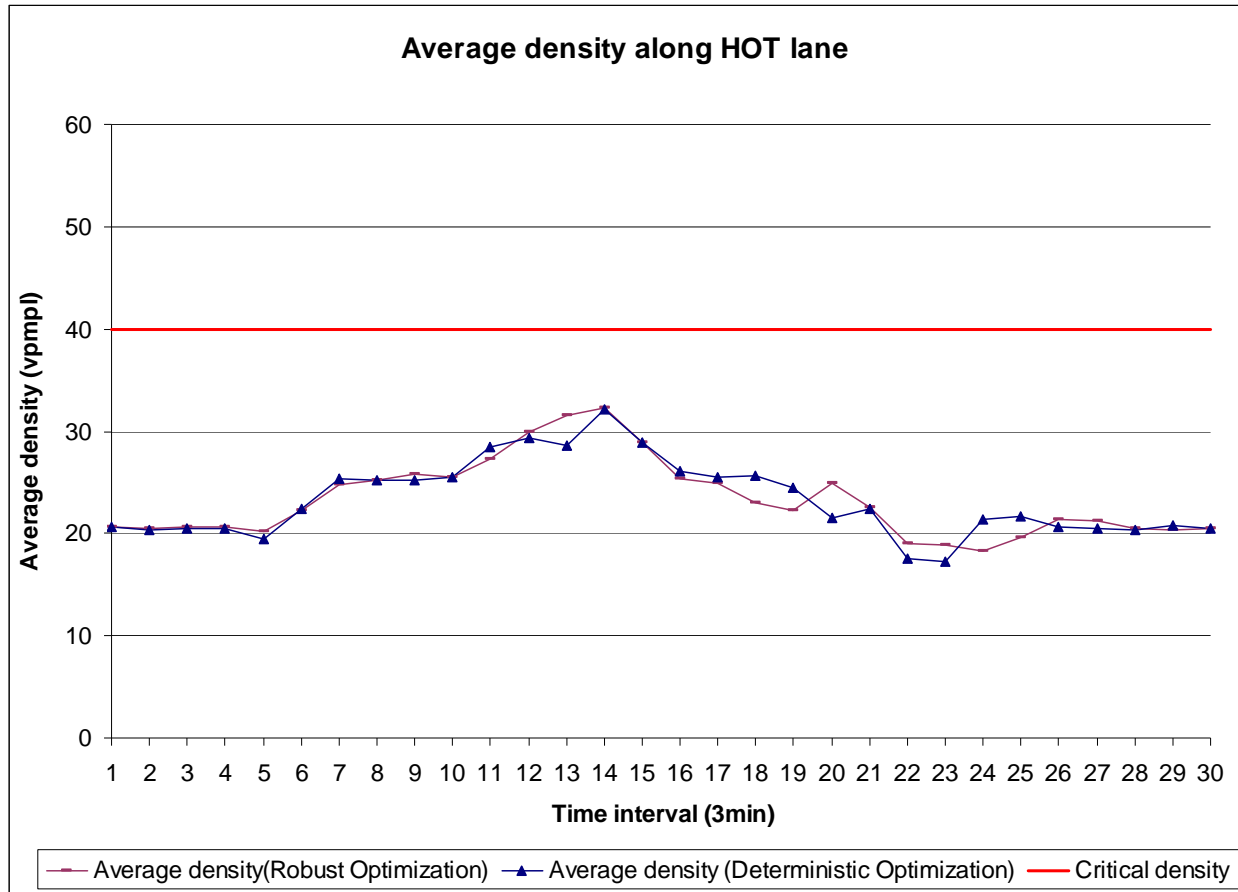


Figure 5-26. Average density along HOT lane.

Figure 5-27 illustrates the queues developed at HOT lane. At the deterministic case, the queue is present more time than in the robust case but the average is 0.31 miles while at the robust case is 0.26 miles which is slightly worse.

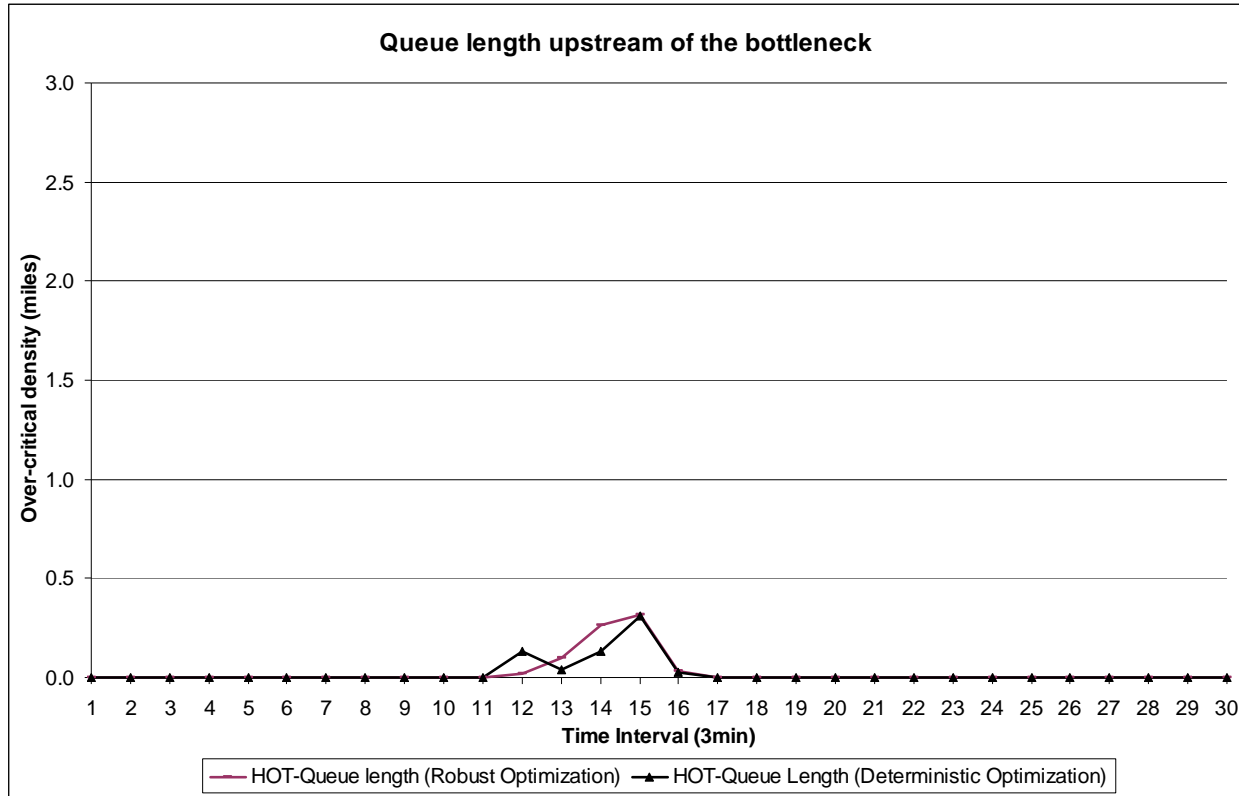


Figure 5-27. Queue length upstream of the bottleneck at HOT lane.

From the above diagrams, one can conclude that:

- The tolls fluctuate more in the deterministic case.
- The profit is more in the deterministic case.
- The throughput of the HOT lane is almost the same in both cases.
- The average density varies, also, almost in the same manner in both approaches.
- The queue lasts longer and it is slightly bigger at the deterministic case.

All in all, in the low-medium-high-medium-low case, the two approaches yield similar performances. Both the approaches perform pretty well in meeting the assumptions made for the operation of the HOT lane and the freeway.

5.4 Conclusions

In this Chapter the results of the two proposed approaches presented and compared in their effectiveness of providing free flow conditions at the HOT lane while utilizing the freeway's capacity.

It can be concluded that both approaches are effective and can be used to manage a freeway segment. However, the scenario-based or robust approach produces a smoother toll pattern and generally better performance of the HOT lane. Also, it is observed that the robust approach respond more adaptively to a sudden demand surge. On the other hand, the deterministic approach requires much less computational time and yields fairly good results. Therefore, the choice of one approach instead of the other must rely on how much accuracy is needed at the pricing strategies and how important is to have tolls that do not fluctuate much from one time interval to the next.

LIST OF REFERENCES

- Arnott R., de Palma A., and Lindsey R. 1998. Recent Developments in the Bottleneck Model. Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility (Kenneth J. Button and Erik T. Verhoef, eds), 79-110.
- Arnott, R., A. de Palma and Lindsey, R., 1993. A structural model of peak-period congestion: a traffic bottleneck with elastic demand. American Economic Review 83 (1), pp. 161-179.
- Arnott, R., de Palma, A. and Lindsey, R. 1995. Information and time-of-usage decisions in the bottleneck model with stochastic capacity and demand, manuscript.
- Arnott, R., de Palma, A. and Lindsey, R. 1997. Recent development in the bottleneck model. Road pricing, Traffic Congestion and Environment: Issues of Efficiency and Social Feasibility, Aldershot: Edward Elgar, 1998, 79-110.
- Brilon, W., Geistefeldt, J., Regler, M. 2005. Reliability of Freeway Traffic Flow: A stochastic Concept of Capacity, Proceedings of the 16th International Symposium on Transportation and Traffic Theory, pp. 125-144.
- Carey, M. and Srinivasan, A., 1993. Externalities, Average and Marginal Costs, and Tolls on Congested networks with time-varying Flows. Operation Research, Vol. 41, No 1, January – February 1993.
- Colorado Department of Transportation 2006. Available from: <http://www.dot.state.co.us/communications/news/DM20060828-1.htm>. Date accessed: November 2007.
- Daganzo, C. F. 1994. The cell-transmission model: A simple dynamic representation of highway traffic. Transportation Research, 28B(4), 269-287.
- Daganzo, C.F., 1995. The cell-transmission model, Part II: Network traffic. Transportation Research, 29B (2), 79-93.
- Dahlgren J. 1999 High-Occupancy Vehicle/Toll Lanes: How Do They Operate and Where Do They Make Sense? Intellimotion, Volume 8, No.2, pp1-3.
- Dahlgren J. 2002 High-occupancy/toll lanes: where should they be implemented? Transportation Research, Part A, Volume 36, Issue 3, pp 239-255.
- Federal Highway Administration, 2006a. Available from: http://ops.fhwa.dot.gov/tolling_pricing/value_pricing/projtyps/hovhotlanes.htm. Date accessed: November 2008.
- Federal Highway Administration, Department of Transportation, 2006b. Congestion Pricing, A Primer. Available from: <http://ops.fhwa.dot.gov/publications/congestionpricing/congestionpricing.pdf>. Date accessed: November 2007.

- Friesz, L.T., Kwon, C., Chow, A., and Heydecker, B., 2006. A Computable Theory of Dynamic Congestion Pricing. Proceeding of the 17th International Symposium on Transportation and Traffic Theory, 2006.
- GAMS 22.6, GAMS Development Corporation, 2003. Available from: <http://www.gams.com>. Date accessed: February 2009.
- Halvorson R., Nookala M. and Buckeye K.R. 2006. High Occupancy Toll Lane Innovations: I-394 MnPASS. The 85th Annual Meeting of the Transportation Research Board, Compendium of Papers CD-ROM, No. 06-1265, January 9 – 13, 2006.
- Iryo, T. and Kuwahara, M. 2000. A Theoretical Analysis on Departure Time Choice for Morning Commute Traffic Considering Individual Variation in Time Value and an Application to Road Pricing. Monthly Journal of Institute of Industrial Science, University of Tokyo. Available from: <http://www.transport.iis.u-tokyo.ac.jp/PDFs/2000/2000-0.08.pdf>. Date accessed: March 2008.
- Joksimovic, D., Bliemer, M.C.J., and Bovy, P.H.L., 2005a. Optimal Toll Design Problem in Dynamic Traffic Networks with Joint Route and Departure Time Choice. Transportation Research Record, Volume 1923 / 2005, pp. 61 – 72.
- Joksimovic, D., Bliemer, M.C.J., and Bovy, P.H.L., 2005a. Optimal Toll Design Problem in Dynamic Traffic Networks with Joint Route and Departure Time Choice. Transportation Research Record, Volume 1923 / 2005, pp. 61 – 72.
- Joksimovic, D., Bliemer, M.C.J., Bovy, P.H.L., and Verwater –Lukszo, Z. 2005b. Dynamic road pricing for optimizing network performance with heterogeneous users. Networking, Sensing and Control, 2005. Proceedings. 2005 IEEE Volume, Issue, 19-22 March 2005, Pages: 407 – 412. Available from: <http://www.ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/9887/31421/01461225.pdf>. Date accessed: January 2008.
- Joksimovic, D., Bliemer, M.C.J., Bovy, P.H.L., and Verwater –Lukszo, Z. 2005b. Dynamic road pricing for optimizing network performance with heterogeneous users. Networking, Sensing and Control, 2005. Proceedings. 2005 IEEE Volume, Issue, 19-22 March 2005, Pages: 407 – 412. Available from: <http://www.ieeexplore.ieee.org/Xplore/login.jsp?url=/iel5/9887/31421/01461225.pdf>. Date accessed: February 2008.
- Knight, F. H. 1924. Some fallacies in the interpretation of social cost. Quarterly Journal of Economics, 38, 582-606.
- Laval, J.A. and Daganzo, C.F. 2006. Lane-changing in traffic streams. Transportation Research, Part B, 40, 251-264.
- Lin, T.K. 2006. Dynamic Pricing with real-time demand learning. European Journal of Operational Research, 174, 522-538.

- Lin, W.H. and Ahanotu, D., 1995. Validating the basic cell transmission model on a single freeway link. University of California, Berkeley, Technical Note UCB-ITS-PATH-TN-95-3.
- Lin, W.H. and Daganzo, C.F., 1994. Technical description of NETCELL. General framework and data structure. University of California, Berkeley, Technical Note UCB-ITS-PATH-TN-95-3.
- Lindsey, R. and Verhoef, E. 2000. Traffic Congestion and Congestion Pricing. [Tinbergen Institute Discussion Papers 00-101/3](http://www.tinbergen.nl/discussionpapers/00101.pdf), Tinbergen Institute. Available from: <http://www.tinbergen.nl/discussionpapers/00101.pdf>. Date accessed: January 2009.
- Liu L.N., and McDonald J.F. 1999 Economic efficiency of second-best congestion pricing schemes in urban highway systems. *Transportation Research, Part B*, 33, 157-188.
- Lo, H. 1999 A novel traffic signal control formulation. *Transportation Research, Part A*, 33, 433-448.
- Lo, H. 2001 A cell-based traffic control formulation: strategies and benefits of dynamic timing plans. *Transportation Science*, 35, 148-164.
- Lo, H. and Y.W. Szeto 2002 A cell-based variational inequality formulation of the dynamic user optimal assignment problem. *Transportation Research, Part B*, 36, 421-443.
- Lou, Y., Yin, Y. and Laval J. A. 2007. Optimal Dynamic Pricing Strategies for High-Occupancy/Toll Lanes. World Conference on Transport Research Society, TRIS online. Available from: <http://ntlsearch.bts.gov/tris/record/tris/01122366.html>. Date accessed: March 2008.
- Morrison, S. A., 1986. A survey of road pricing. *Transportation Research A*, Vol. 20A, No. 2, pp. 87-97.
- Obesberger, J. 2004. Managed lanes: combining access control, vehicle eligibility, and pricing strategies can help mitigate congestion and improve mobility on the Nation's busiest roadways. *Journal of Public Roads*, November 2004.
- Pigou, A. C. 1920. *Wealth and Welfare*, Macmillan, London.
- Pockafellar, R.T. and Uryasev, S. 2000. Optimization of conditional value-at-risk. *Journal of risk* 2:21-41.
- Press Release, Congestion Eroding Interstate System's Safety and Economic Benefits. The Road Information Project, January 16, 2003.
- Small, K. A., 1983. The incidence of congestion tolls on urban highways. *J. Urban Econ.* 13, 90-110.
- Small, K.A. 1982. The scheduling of consumer activities: work trips. *American Economic Review*, 72, 467-79.

- Smith, L. 2007. Congestion Pricing. Available from:
http://www.calccit.org/itsdecision/serv_and_tech/Congestion_pricing_report_print.htm.
Date accessed: February 2008.
- Sullivan, E. 2000. Continuation Study to Evaluate the Impacts of the SR 91 Value-Priced Express Lanes. Final Report. Available from:
http://ceenve3.civeng.calpoly.edu/sullivan/SR91/final_rpt/FinalRep2000.pdf. Date accessed: April 2008.
- U.S. Department of Transportation, 2006. National Strategy to Reduce Congestion on America's Transportation Network.
- Verhoef, E., 1997. Time-Varying Tolls in a Dynamic Model of Road Traffic Congestion with Elastic Demand. Tinbergen Institute Discussion Papers from Tinbergen Institute, No 97-028/3. Available from: <http://www.tinbergen.nl/discussionpapers/97028.pdf>. Date accessed: February 2008.
- Vickrey, W.S. 1969. Congestion theory and transport investment. *American Economic Review* 59, 251-260.
- Yin, Y. 2007. A scenario-based Model for Fleet Allocation of Freeway Service Patrols. *Netw Spat Econ*. Available from:
<http://www.springerlink.com/content/a82374380x714302/fulltext.pdf>. Date accessed: March 2008.
- Yin, Y. and Lou, Y. 2006. Dynamic Tolling strategies for managed lanes. The 86th Annual Meeting of the Transportation Research Board, Compendium of Papers CD-ROM, No 07-1806, January 21 – 25, 2007.
- Yperman, I., Logghe, S., and Immers, B., 2005. Dynamic congestion pricing in a network with queue spillover. In *Proc. 12th World Congress on Intelligent Transportation Systems*, San Francisco.
- Ziliaskopoulos, T. 2000 A linear programming model for the single destination system optimum dynamic traffic assignment problem. *Transportation Science*, 34, 1-12.

BIOGRAPHICAL SKETCH

Dimitra Michalaka was born in Lesvos, Greece, in 1984. In 2001, after she passed the national general exams, she enrolled as a student at the National Technical University of Athens, where she received the Bachelor of Science in civil engineering in 2006. In 2009 she earned the Master of Science in civil engineering from University of Florida. During her graduate studies, she was a research assistant for her advisor, Dr. Yafeng Yin.