APPLICATION OF BAYESIAN STOCHASTIC LEARNING AUTOMATA FOR MODELING LANE CHOICE BEHAVIOR IN SR-167 HOT LANES

Ender Faruk Morgul, Ph.D. (Corresponding Author)
Research Fellow, Department of Civil & Urban Engineering,
Polytechnic School of Engineering,
Center for Urban Science + Progress (CUSP), New York University (NYU)
One MetroTech Center, 19th Floor, Brooklyn, NY 11201, U.S.A.
Tel: 1-(646)-997-0531
Email: efm279@nyu.edu

Kaan Ozbay, Ph.D.
Professor, Department of Civil & Urban Engineering,
Polytechnic School of Engineering,
Center for Urban Science + Progress (CUSP), New York University (NYU)
One MetroTech Center, 19th Floor, Brooklyn, NY 11201, U.S.A.
Tel: +1 (646) 997 0552
Email: kaan.ozbay@nyu.edu

Abdullah Kurkcu, M.Eng.
Graduate Research Assistant, Department of Civil & Urban Engineering,
Polytechnic School of Engineering,
Center for Urban Science + Progress (CUSP), New York University (NYU)
One MetroTech Center, 19th Floor, Brooklyn, NY 11201, U.S.A.
Tel: 1-(646)-997-0538
Email: ak4728@nyu.edu

Word count: 5,843 + 1 Table + 5 Figures = 7,343 words
Abstract: 135 words

Re-submission Date: November 15, 2015

Paper Submitted for Presentation and Publication at the
Transportation Research Board's 95th Annual Meeting, Washington, D.C., 2016

ABSTRACT

This paper investigates learning behavior in SR-167 HOT lanes using a six-month toll tag reading data. Bayesian-Stochastic Learning Algorithm (SLA) theory is employed to model drivers' sequential lane choice occasions. Reward and penalty parameters are used to update users' lane choice probabilities. The results show that the effect of reward parameters which increases selecting probability of an alternative after a satisfactory experience is more obvious than penalty parameters that decrease the probability of selecting an unfavorable choice. Low magnitudes of learning parameters might indicate strong habit formation among the users. Moreover, posterior distribution of learning parameters indicates there exist user perception heterogeneity when evaluating the outcomes of choices. Finally, user familiarity is investigated with a less experienced subsample and it is shown that learning rates of more familiar users are lower than less familiar users.

BACKGROUND

A widely used implementation of congestion pricing in the U.S. is Managed Lanes (ML), where one or more lanes are dedicated to certain user groups to maintain free-flowing traffic conditions. High-occupancy vehicle (HOV), High-Occupancy toll (HOT), Express lanes (EL) and busways are among the most common forms of MLs (*1*). In HOT lanes drivers can carpool and experience a faster and more reliable travel for free or they can pay a toll when driving alone. Most of the existing HOT lanes in the U.S. employ certain kind of a dynamic pricing that aims to allocate capacity in an efficient way by adjusting toll rates in real-time depending on the congestion level. Several existing HOT lane facilities and pilot projects have shown that obtained benefits to society include more reliable travel times, less number of accidents, increased ridesharing and transit usage.

Existing literature on driver behavior in MLs mainly focused on savings in average travel times and travel time reliability (*2-8*). Moreover, recent empirical work from two different ML facilities showed that that the choice of using MLs cannot solely be dependent on travel time savings since operating speeds in dedicated lanes are mostly similar with the general purpose lanes (*5; 9*). For example, Burris et al. (*9)* showed that 35 percent of paid trips in I-394 ELs took place when travel time savings were less than one minute. Some other recent literature shifted their focus on other underlying factors of lane choice behavior such as individual risk related attributes. Burris et al. (*10)* showed evidence for ML users' having significantly higher risk tolerance and also having tendency towards risky driving which is associated with higher travel speeds (*10*).

Estimation of willingness-to-pay (WTP) for travel time and travel time reliability savings of ML users has been generally carried out using discrete choice models. Data used for model estimations are usually collected from stated preference (SP) surveys where ML users' responses are evaluated for hypothetical choice situations. Alternative data sources that are employed in the literature include combined SP and revealed preference (RP) data (*7; 11*) and loop detector data (*12; 13*). Alternative to discrete choice modeling, some studies employ methodologies to calculate aggregated mean or median cost per hour savings from revealed toll reader data (*5; 9*).

An interesting observation based on two ML facilities, i.e. Katy MLs and I-394 HOT lanes, is that the majority of the ML users only infrequently choose to travel on tolled lanes for rare situations such as when there is a risk of being late to work (*10*). Unfortunately, occasional usage of MLs cannot be captured by random samples that are used in stated preference (SP) surveys (*2*). Therefore WTP estimations from SP data are often found to underestimate the actual values when compared with revealed preference (RP) data (*5*).

One of the major goals of MLs is to generate revenue while improving mobility and relieving traffic congestion (*14*). Although empirical studies strongly agree that individual decision-making towards ML usage is mainly motivated by savings in travel time and its reliability, Cao et al. (*15)* found that in I-394 ELs only 7 percent of total costs are compensated by travel time savings, whereas travel time reliability savings contribute an additional 23 percent to cover total costs. Improvements in safety are stated as the main sources of benefits in I-394 ML program (*15*).

Although factors that affect ML usage have been extensively investigated in the literature, the research question whether observed behavior of ML users reveals that they learn from their previous experiences remains largely unexplored. This study aims to add to the literature by investigating HOT lane users' lane choice behavior within a stochastic learning framework by utilizing large-size observed toll transaction data. In other words, this research models learning characteristics of a large group of HOT lane users by taking advantage of this rich revealed preference data

LEARNING BEHAVIOR OF TOLL ROAD USERS

Understanding day-to-day evolution of user decisions in MLs is of utmost interest for many aspects of operational efficiency. Until recently, lane choice behavior in MLs is usually evaluated based on SP surveys which can only provide a snapshot of actual case in a tolled road. Therefore, from SP data it is not possible to observe whether frequent users are seeking an equilibrium state over time as they continue to use the ML facilities. Some studies also used laboratory experiments to get responses of participants for consecutive time periods with varying travel conditions (*16; 17*).

Day-to-day learning models in transportation literature can be broadly grouped into four major categories. The first and most commonly used method is the weighted average of previous choices. In this method experienced travel times are given weights and a cumulative weighted average is calculated as a learning measure to incorporate in a model framework. Several theoretical and empirical studies can be found using different forms of weighted average approach in transportation decision-making context (*18-23*). The second main avenue of learning models is the Bayesian learning which can be considered as a special form of weighted average model. The distinctive difference in Bayesian learning is that the previous travel times are represented as distributions and variance is considered as the driver's confidence about past information. Therefore in Bayesian models perceived quality of travel time information is used as an attribute in the utility associated with the route choice (*24-26*). The third methodology is the reinforcement learning which addresses habitual behavior in decision-making. In particular, in reinforcement models as long as a choice of route does not result in very inconvenient travel times, perceived relative attractiveness of this choice increases (*21; 27; 28*). The final approach is stochastic learning automata (SLA) in which drivers' day-to-day choice probabilities are updated based on their actual experiences in the system. The primary idea in SLA framework is that actions take place in an unknown random environment in which automation operates. Actions are categorized as favorable and unfavorable where rewards and penalties are used to update choice probabilities. Being very similar to Bayesian learning, the main difference in SLA is that probability updating procedure is not only limited to Bayes' rule (*29*). Several implementations of the SLA based learning models have been proposed in the transportation literature including route and departure time choice (*17; 30; 31*).

Although there are many papers dealing with-day-to day learning behavior of drivers on a regular traffic network, only a limited number of studies attempted to empirically model day-to-day learning of drivers in MLs. Yin et al. (*23)* developed a macroscopic simulation based tool to investigate HOT lane users' learning behavior for two cases; pre-trip departure time choice and en-route lane choice. The reason for using a simulation framework was stated by the authors as *"the lack of adequate observed data"(23)*. They have developed a weighted average methodology in which learning dynamics associated with travel time and toll are assumed to be independent. In their formulation weights are discounted exponentially with the most recent experience having the highest weight. Route costs for General Purpose (GP) and HOT lanes are defined as a function of departure time including attributes associated with travel time, toll, early and late arrival delay, all of which were assumed to have fixed coefficients. Each traveler was assumed to have a desired arrival time, which is calculated from a distribution, and any kind of delay, early or late, is assumed to be not preferable. It was also assumed that drivers choose optimal departure time to minimize their route cost. After departure time choice, en-route choice is considered with the newly available projected travel time information and users are allowed to change their lane options. Different user classes are considered in the simulation with various constraints for example some vehicles that are not equipped with toll tags could not use HOT lanes. Simulation results showed that depending on demand level system might exhibit hysteresis-like behavior. In particular, in case of an overloaded demand, users were observed to constantly change their departure time and lane choices on a day-to-day basis.

Alvarez (*32)* investigated learning process in MLs based on aggregated lane choice observations from I-95 HOT lanes. First, a correlation analysis is conducted between HOT lane user proportion and difference between travel times of HOT and GP lanes for different time windows. The results implied that prior experiences up to 60 days could be correlated with the lane choice. Second analysis used weighted averages for earlier travel times and it was found that users tend to give the same weight to all experiences within a given period of time. This results is not consistent with average weight approaches in the literature where most recent experiences are assumed to have higher weights (*33*). Finally, no evidence was found that worst experiences within a time window has a significant impact on learning process.

Xiong (*34)* developed a departure time choice model for MLs considering Bayesian learning. The model assumed that agents continuously search for optimal departure times until every agent in the systems stops searching and system reaches user equilibrium. Moreover, prior and posterior beliefs of users were assumed to follow a Dirichlet distribution from which random draws were taken. The approach was then applied in a hypothetical microsimulation framework of a toll road. The results showed that under the developed model system equilibrium was reached after adequate number of iterations and lower number of changes in departure times was observed over time.

This study aims to develop a better understanding of the effect of day-to-day experience in decision making of drivers using a longitudinal toll tag data. Observed individual decision making information is extracted from toll data over a six-month period. Bayesian-SLA approach proposed by Yanmaz-Tuzel et al. (*31)* is then applied for modeling the lane choice behavior of selected agents which are ensured to be the frequent users in SR-167 HOT lanes.

METHODOLOGY

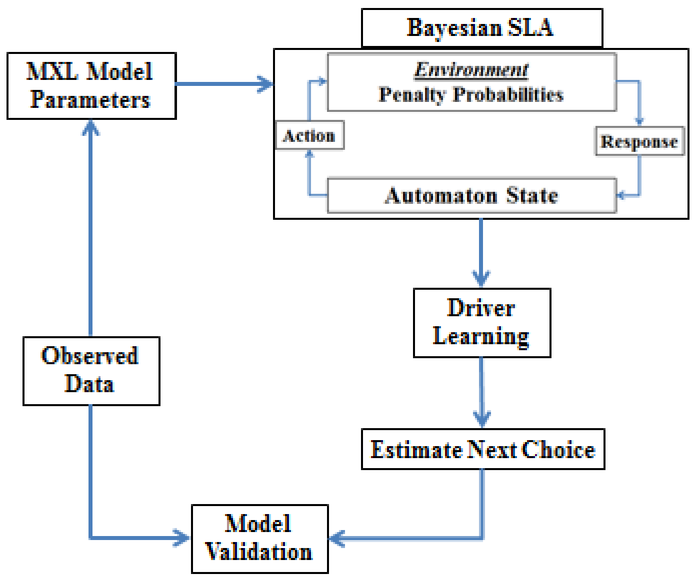

FIGURE 1: Modeling Flowchart

Modeling flowchart followed in this paper is shown in FIGURE 1. Route choice behavior modeling using SLA was first introduced by Ozbay et al. (*17)*. In a following study, SLA model was extended to address combined departure time and route choice (*30*). Yanmaz-Tuzel et al. (*31)* further improved the SLA methodology to Bayesian-SLA model that also incorporates bounded rationality (BR) and analyzed the impacts of a network disruption using aggregated observed demand data as an empirical example.

In this paper we adopt Bayesian-SLA methodology using an empirical setup with a longitudinal dataset where day-to-day disaggregated user behavior can be investigated. Therefore "agents" are real individuals who frequently use MLs and their observed behaviors are available over a certain time period for model parameters' estimation.

Bayesian Stochastic Learning Automata

Environment

Two major elements of Bayesian-SLA are the environment and the stochastic automaton. The automaton basically operates in a random environment and it can perform a finite number of actions. In this study, transportation system is considered as the random environment and traveler as the automaton. When automaton performs a specific action, for example choosing to travel on the MLs, random environment responds by producing an output which is stochastically related to the action. The goal is to design an automaton that can determine the best action for itself based on this feedback. Throughout the process automaton is guided by its previous actions and corresponding responses by the environment which is basically the process of learning.

Environment may be time varying and actions of other agents in the system, which are unknown to the automaton, may affect the responses of an environment. An environment is defined by a triplet $\{\underline{\alpha}, \underline{c}, \underline{\beta}\}$. The term $\underline{\alpha}$ represents the action set in which the traveler decisions are defined. The term $\underline{c}$ represents the set of probabilities of receiving a penalty from the environment as a result of an action. Each element c_i in penalty probability set corresponds to an action α_i. If probability c is assumed to be constant for a specific action, the environment is called stationary. In the case at hand we assume non-stationary environments. The last term $\underline{\beta}$ is the output set which gives the utility obtained as a result of a choice. In a P-model the output values are considered to only take one of two values. In this study output takes value of 1 for the unfavorable outcomes and 0 for the favorable ones. Alternatively a continuous range of values may be defined for the outcome values, which is called S-model (*29*).

Mathematically, the case when automaton is applied to the environment at time $t = n$, for sets of size r, the relationship between the three components of environment can be given as:

$$\Pr(\beta(n) = 1 \mid \alpha(n) = \alpha_i) = c_i \qquad \text{(i=1,2,...,r)} \tag{1}$$

where α_i is one of the possible actions (e.g. selecting GP or HOT) which is unfavorable. Action probabilities are updated based on the response from the environment at every stage using a reinforcement scheme.

Reinforcement Schemes

Penalty probabilities are updated based on linear reinforcement functions which are conditioned on whether the action is favorable or not. A general reinforcement scheme is defined as $p(n+1) = T[p(n), \alpha(n), \beta(n)]$ where T is the mapping function. Optimality is achieved in the case when an action α_m with the minimum penalty probability c_m is chosen asymptotically with probability one. However, for most of the cases it is not possible to achieve optimality, therefore sufficiently close asymptotic behavior is generally considered as an acceptable solution (*29*).

In this paper, following linear reinforcement scheme is used:

$$
\begin{aligned}
&if \ \ \alpha(\text{n})=\alpha_i \\
&\beta(n)=0 \Rightarrow \begin{cases} p_j(n+1)=(1-a)\,p_j(n) & \forall_j \neq i \\ p_i(n+1)=p_i(n)+a[1-p_i(n)] \end{cases} \\
&\beta(n)=1 \Rightarrow \begin{cases} p_j(n+1)=\dfrac{b}{r-1}+(1\text{-}b)\,p_j(\text{n}) & \forall_j \neq i \\ p_i(n+1)=(1-b)\,p_i(n) \end{cases}
\end{aligned}
\tag{2}
$$

In above formulation a and b are reward and penalty parameters both of which are defined in [0,1]. This set of equation is also called linear reward-penalty learning ($L_{R\text{-}P}$) scheme (*29*). The estimation of parameters a and b is generally performed with trial and error (*17; 30*). The effect of penalty parameter b Following Yanmaz-Tuzel et al. (*31)*, heterogeneity in learning parameters is incorporated using Bayesian Inference theory. In other words joint posterior distribution of learning parameters are calculated instead of using a unique set of learning parameters for the entire population. Parameter r represents the choice set which generalizes model for multi-action choice frameworks. In case at this paper r is equal to two since only choice considered is selection between HOT and GP lanes.

Interpretation of Equation 2 in route choice context is straightforward. If travel time of first alternative (i.e. HOT lane) at day n is less than travel time of second alternative (i.e. GP lane) and traveler selects the first alternative, the outcome becomes favorable and using reward parameter a the probability of selecting first alternative increases at day $n+1$. For the reverse case, outcome becomes unfavorable and probability of choosing the same alternative next day decrease using penalty parameter b .

Posterior Distribution of Learning Parameters

The main motivation for using Bayesian analysis for the learning parameter estimation is to address the differences in user perceptions in the population. Earlier studies used fixed learning parameters, usually by selecting an average of a range of acceptable values after several iterations (*17*). In Bayesian approach, given the likelihood function of observations and assumed prior information of learning parameters, a joint posterior distribution of learning parameters is estimated. Based on Equation (2), likelihood function $p(D \mid a,b)$ of the observations D given learning parameters (a,b) is given as follows:

$$
p(D \mid a,b)=\prod_{k=1}^{K}\prod_{n=1}^{N}\prod_{i=1}^{r} p_{ki}(n), \tag{3}
$$

$$p(D \mid a,b)=\prod_{k=1}^{K}\prod_{n=1}^{N}\prod_{i=1}^{r}\begin{cases}[\,p_{ki}(n-1)+a(1-p_{ki}(n-1))]^{\alpha_{ki}(n-1)(1-\beta_k(n-1))}\\ [(1-a)\,p_{ki}(n-1)]^{(1-\alpha_{ki}(n-1))(1-\beta_k(n-1))}\\ [(1-b)\,p_{ki}(n-1)]^{\alpha_{ki}(n-1)\beta_k(n-1)}\\ [\dfrac{b}{r-1}+(1-b)\,p_{ki}(n-1)]^{(1-\alpha_{ki}(n-1))\beta_k(n-1)}\end{cases} \tag{4}$$

where k and n are the index numbers of the user and choice occasion i.e. days, respectively. Probability of choosing alternative i for user k on choice occasion $(n-1)$ is represented by $p_{ki}(n-1)$. Parameters α and β are binary variables which are defined as:

$$\alpha_{ki}(n-1)=\begin{cases}1 \text{ if user } k \text{ selects alternative } i \text{ on choice occasion } (n\text{-}1)\\ 0 \text{ otherwise}\end{cases} \tag{5}$$

$$\beta_k(n-1)=\begin{cases}1 \text{ if user } k \text{ experiences a favorable action on choice occasion } (n\text{-}1)\\ 0 \text{ otherwise}\end{cases} \tag{6}$$

In Bayesian framework posterior distribution is simply proportional to the product of likelihood function and prior distribution. Joint prior distribution of learning parameters, $p(a,b)$, is assumed to follow Normal distribution with mean μ and covariance matrix Σ. Mathematically,

$$p(a,b)=\frac{1}{2\pi\,|\,\Sigma\,|^{1/2}}\exp\left\{\frac{1}{2}(x-\mu)^T\,\Sigma^{-1}(x-\mu)\right\}$$
$$\text{where } \mu=\begin{bmatrix}\mu_a\\ \mu_b\end{bmatrix} \quad \Sigma=\begin{bmatrix}\sigma_a^2 & 0\\ 0 & \sigma_a^2\end{bmatrix} \quad x=(a,b) \tag{7}$$

Finally, joint posterior distribution of learning parameters given the observations is defined as:

$$p(a,b \mid D)\sim P(D \mid a,b)\,p(a,b) \tag{8}$$

Algorithms such as Metropolis-Hasting (M-H) (*35*) or Gibbs Sampling (*36*) are employed to generate samples from the complex multidimensional joint posterior distributions. In this study M-H algorithm is used which basically generates samples using Markov Chain Monte Carlo (MCMC) procedure. Convergence of Markov chains is tested with Heidelberger et al.

(*37)*'s convergence diagnostic which uses Cramer-von-Mises statistic to compare observed sequence of samples to a hypothetical stationary distribution. The null hypothesis is that sequence of samples is similar to the stationary distribution. First, stationary test iteratively eliminates first 10% of the chain until either the null hypothesis is not rejected or 50% of the original chain remains. A second test, half-width test, by calculating half-width of the confidence interval, $(1-c)\times 100\%$ in which c is a level for confidence, is then applied for the samples which pass the stationary test. If the ratio of the mean of the sample to the half-width of confidence interval is smaller than some threshold then the second test succeeds. Final estimated group of parameters are the ones which pass both tests (*37*).

Action Outcome Evaluation

Outcomes of actions are classified as favorable and unfavorable depending on experienced schedule delays associated with the selected alternative route. A utility framework is used to evaluate outcomes of each action. In proposed Bayesian-SLA, users are assumed to seek for satisfying choices. A satisfying choice is not necessarily the best choice. Therefore proposed framework does not assume that users have cognitive ability to process all the available information simultaneously. The coefficients of parameters included in the utility function are estimated using Mixed Logit (MXL) model which allows heterogeneity among population by assigning random coefficients to different users (*38; 39*). Random coefficients of selected attributes can be assumed to have any kind of distribution, however to avoid unrealistic WTP estimations (e.g. negative valuation of travel time savings) bounded distributions are recommended (*40*). In this study we use one-sided triangular distribution for the behavioral profile for preference heterogeneity for model attributes travel time, toll cost and schedule delays. Consistent with literature, early and late delays are treated differently in model structure and separate coefficients are estimated for each

In MXL model specification attributes that are assumed to have random coefficients are travel time (in minutes), toll rate (dollars), early and late schedule delays (in minutes). Non-random attributes include binary variables for date and time of trips, a lagged variable which holds the immediate previous decision as a variable. The reason for including a lagged variable is to control for state dependency and the effect of earlier choices. Since lagged variables are not available for the very first choices, first observation of each user is excluded in the estimation sample. Utility function of an alternative i for user k on choice occasion n is described as:

$$U_{ikn} = \alpha_i + \beta_k X_{ikn} + \theta_k I_{ikn} \tag{10}$$

where α_i is alternative-specific constant, β_k is the individual-specific random parameter vector, X_{ikn} are attributes that are assumed to have random coefficients, θ_k is the vector of non-random parameters and I_{ikn} are the attributes associated with non-random parameters. After investigating several constrained distribution specifications, the values of random parameters are drawn from a one-sided triangular distribution due to superior model fit performance. Maximum simulated likelihood method is used for estimation by setting 1500 Halton draws.

DATA

Data Source

Proposed methodology is implemented based on observed data from SR-167 HOT lanes in Washington, USA. Toll tag reader records of approximately six months, between January 15, 2013 and June 30, 2013, are used for the analysis. HOT lanes include six northbound segments with a total length of 10.76 miles and four southbound segments with a total length of 7.69 miles. HOT lanes operate all days of the week between 5 a.m. to 7 p.m (*41; 42*). During the time frame of available data, daily average traffic flow detected by loop detectors was about 2500 vehicles per hour (*43*). According to the total number of trips in toll reader data, about 20 per cent of all travelers employ toll tag readers in their cars. The user group analyzed in this study are the paying HOT lane users who are detected by the tag readers when traveling in either HOT and GP lanes. Over 1.2 million trip records are available in the raw dataset.

Toll tag reader data has several advantages for investigating observed learning behavior in MLs. First and foremost, none of the previous studies were able to investigate actual lane choice behavior at individual level. This brings a major improvement in external validity of results over hypothetical simulation based analysis that works with aggregated data (*33*). Second, time-varying trip attributes such as actual toll rate paid and experienced schedule delays can be incorporated in a utility framework to evaluate user satisfaction as a result of the trip. This may also bring a better understanding of habit formation of users based on route characteristics perceptions (*27*). Third, calibration of learning parameters of employed methodologies can be improved with large-sized, heterogeneous and more granular data which leads to overall better model performance (*33*).

Data Processing

Raw dataset includes toll tag identification numbers (IDs) for individual vehicles. Toll tag IDs are anonymized for privacy reasons but it is still possible to track behaviors of users over the time-frame of available data. The lists of traversed segments for each trip are extracted. Since there exist multiple segment records for most of the trips (i.e. if a user is traveling in more than one segment in a single trip occasion), these records are reduced to single trips are by merging segment records for each trip occasion. Final processed data included the choice of lane, travel time information for both alternatives and toll rate when the trip took place.

Since true desired arrival time information of users is not available, schedule delays are calculated based on the assumption that every user has their specific reference travel speeds for each alternative based on their own experiences. For each user median travel speeds for HOT and GP lanes are calculated based on the six-month data. If experienced travel time deviates from the base travel time calculated using reference speeds it is assumed that traveler is delayed.

Sample Selection

In order to observe temporal evolution of lane choice behavior in MLs, only toll tag IDs that regularly use the facility are included in the analysis. Northbound (NB) and southbound (SB) trips are modeled separately since they show significantly different traffic flow patterns throughout the day (*43*). In particular, northbound traffic volumes make peak hours during morning hours (6-8 am) and southbound volumes are highest during afternoon peak hours (4-6 pm). One implication could be the impact of work trips but there is no solid evidence for trip purposes in the dataset at hand.

Number of drivers that repeatedly traveled throughout the six-month time window is depicted by travel direction in FIGURE 2, for northbound and southbound respectively. It should be noted that not all users regularly appear on similar days and choice sequences differ for every traveler. Therefore it is not possible to obtain adequate amount of daily sample to show day-to-day choice behavior. Instead, choice occurrences or consecutive choices of travelers are considered which do not necessarily take place day-by-day.

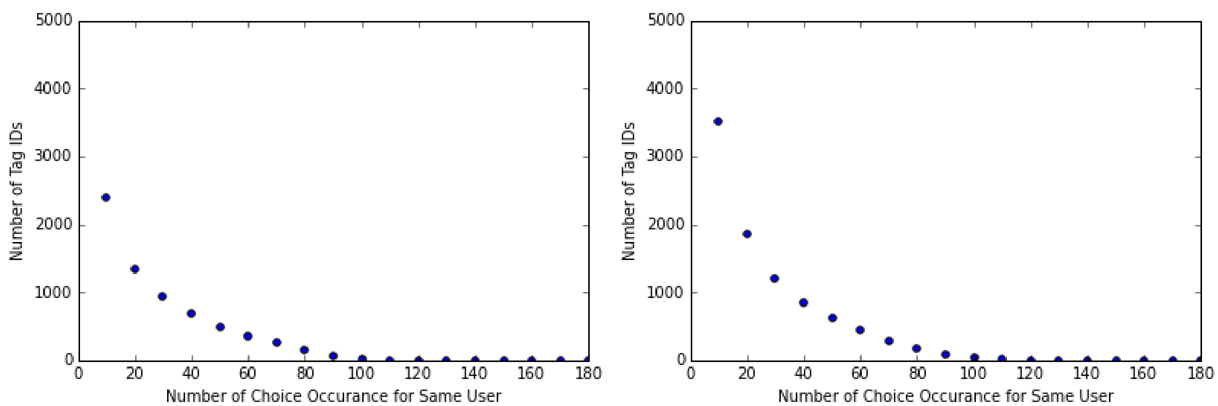

FIGURE 2: Sample Size Selection a) Northbound b) Southbound

Finally, 490 tag IDs in northbound and 630 tag IDs in southbound for 50 choice occurrences each are considered as balanced in terms of sample size and number of occurrences. In particular, every user is evaluated separately starting from their first choice up to their 50th choice during six month period. Based on the responses they receive feedback from the environment and they update their alternative-specific choice probabilities for their next choice occasion which may take place in different days for each user.

Calibration and Validation

Model calibration is performed to ensure the consistency of probability estimations with the observed data. Set of learning parameters, (a,b), iteratively modified to find the minimized difference between observed and estimated lane choices. Mathematically,

$$\min\ F = \Sigma_{\theta}\Sigma_{n}\left[\left(\frac{q_n^{est} - q_n^{obs}}{q_n^{obs}}\right)^2\right] \tag{9}$$

where q_n^{est} is the link flows estimated in the model and q_n^{obs} observed link flows in choice occasion n. After calibration of model parameters, mean standard errors are calculated for validation of travel volumes in the two lane alternatives for each choice occasion.

RESULTS

MXL model Estimation results are given in Table 1. The variables used in this model are selected with regard to best model fit and in consistent with an earlier study of valuation of reliability with the same dataset (*43*). A satisfactory model fit is obtained considering the McFadden pseudo r-squared statistic of 0.544. Signs of estimated random parameters are all negative as expected since increasing toll, travel time and delays would decrease overall utility. It is seen that among random parameters only early schedule delay is not a significant estimator for lane choice for the sampled group. Other random parameters, toll, travel time and late delays, both means and standard deviations are all significant at 1% level. Significance of standard deviations indicates unobserved user heterogeneity associated with random parameters is important in explaining driver behavior in MLs. Another interesting finding is the significance of state dependency which is captured by the lagged variable. This indicates that recency effect (i.e. emphasis on the most recent experience) or stickiness for experienced users has considerable impact on lane choice probability. Similar findings were reported in the literature for route choice behavior from SP experiments (*33*). Alternative specific constants for both HOT and GP lane options, day- and month-specific binary variables are found to have no explanatory power on choice behavior for the sample group.

TABLE 1: MXL Model Results

	Coefficient	Standard Error	Standard Deviation of Random Parameters
Random Parameters			
Toll Price	**-2.179*****	**0.699**	**-2.179*****
Travel Time	**-1.409*****	**0.197**	**-1.409*****
Schedule Delay Early	**-0.201**	**0.240**	**-0.201**
Schedule Delay Late	**-1.164*****	**0.235**	**- 1.164*****
Non-Random Parameters			
ASC GP	**-0.746**	**0 (Fixed Parameter)**	
ASC HOT	**0.746**	**0 (Fixed Parameter)**	
Lagged Variable	**0.321*****	**0.087**	
Day			
Monday	**0.092**	**1.073**	
Friday	**0.0012**	**0.101**	
Month			
January	**2.649**	**0 (Fixed Parameter)**	
February	**2.504**	**0 (Fixed Parameter)**	
March	**2.398**	**0 (Fixed Parameter)**	
April	**2.442**	**0 (Fixed Parameter)**	
May	**2.510**	**0 (Fixed Parameter)**	
June	**2.452**	**0 (Fixed Parameter)**	
Goodness-of-Fit Statistics			
Log-likelihood	**-11558.562**		
AIC	**23147.1**		
McFadden Pseudo R^2	**0.544**		
Number of observations	**36,598**		

******: Significance at 1% level***

Estimation of Learning Parameters

Action probability updating algorithm given in Equation 2 is used to estimate the next choice of individuals in SR-167 HOT lanes. Since $p_j(n)$ is a discrete-time Markov process whose state space is the unit interval [0,1], asymptotic behavior of action probabilities can be calculated for assumed learning reward-penalty parameters *(a,b) (29)*. Instead of single point estimates for learning parameters, Bayesian Inference approach is combined with SLA methodology and joint posterior distributions of learning parameters are estimated. Prior distribution is approximated by a normal distribution with the following parameters chosen on a trial-error basis:

$$\mu = \begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix} = \begin{bmatrix} 0.02 \\ 0.02 \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_a^2 \end{bmatrix} = \begin{bmatrix} 0.06 & 0 \\ 0 & 0.06 \end{bmatrix} \tag{11}$$

Joint posterior distribution is computed for a case where traveler behavior is mimicked as accurately as possible by minimizing the difference between observed share of travel volumes in GP and HOT lanes. A pool of random learning parameters is generated using Metropolis-Hasting algorithm. Then, learning parameters are iteratively simulated to achieve an acceptable mean standard error (i.e. lower than 0.03) in traffic volume percentages. FIGURE 3 gives scatter plots of converged reward and penalty parameters for northbound and southbound directions. One observation is that there is no certain concentration of parameters for both directions. As also pointed out by earlier research the effect of reward parameter on the learning rate is much larger than the penalty parameter (*17*). This indicates a choice that yields a satisfactory result increases a repetition probability by a higher magnitude compared to a decrease in repetition probability as a result of a not satisfactory outcome.

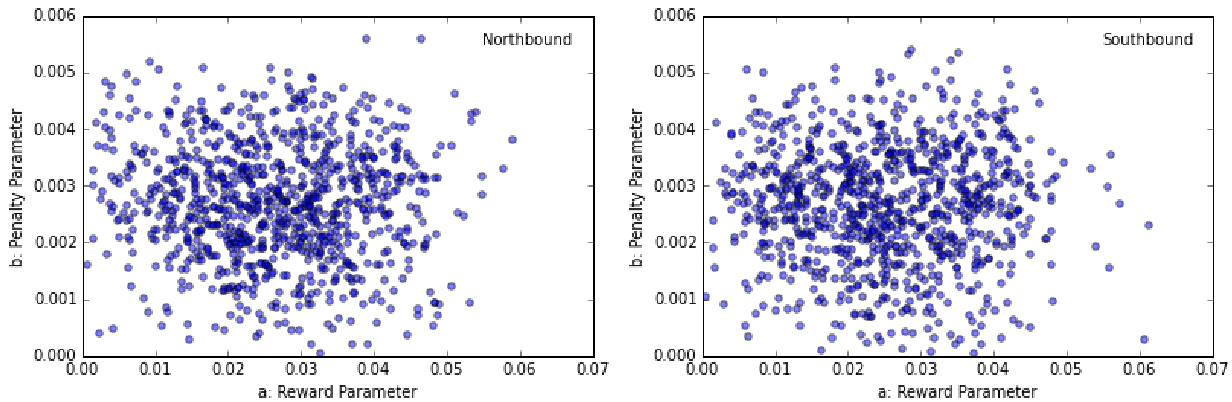

FIGURE 3: Sample Plots for Posterior Distributions of Learning Parameters

Mean values for learning parameters *(a,b)* are estimated as (0.026,0.003) for northbound and (0.025,0.003) for southbound directions. These estimations are considerably lower than other empirical findings using SLA methodology. For example, Yanmaz-Tuzel et al. (*31)* found an average set of values of (0.062, 0.0067) for frequent users in New Jersey Turnpike which is a tolled freeway. Sample group used in study showed comparably slower learning rates which might support Burris et al. (*10)*'s finding that the majority of users prefer to pay tolls only for only rare occasions. In other words, for the case in SR-167 HOT lanes, tendency to frequently change behavior to find the best alternative is not as strong as it was found in other tolled facilities. More generally, other factors such as habits and inertia in choices as discussed in Bogers et al. (*27)* might have a strong influence on user decisions for the case at hand. Such factors are also supported by the significance of lagged variable in MXL model estimated from the same sample.

Histograms of estimated learning parameters are given in FIGURE 4. Posterior distribution of each parameter can be approximated with Beta distribution as depicted. Learning parameter constraints are naturally met for all cases since Beta distribution is defined in domain [0,1].

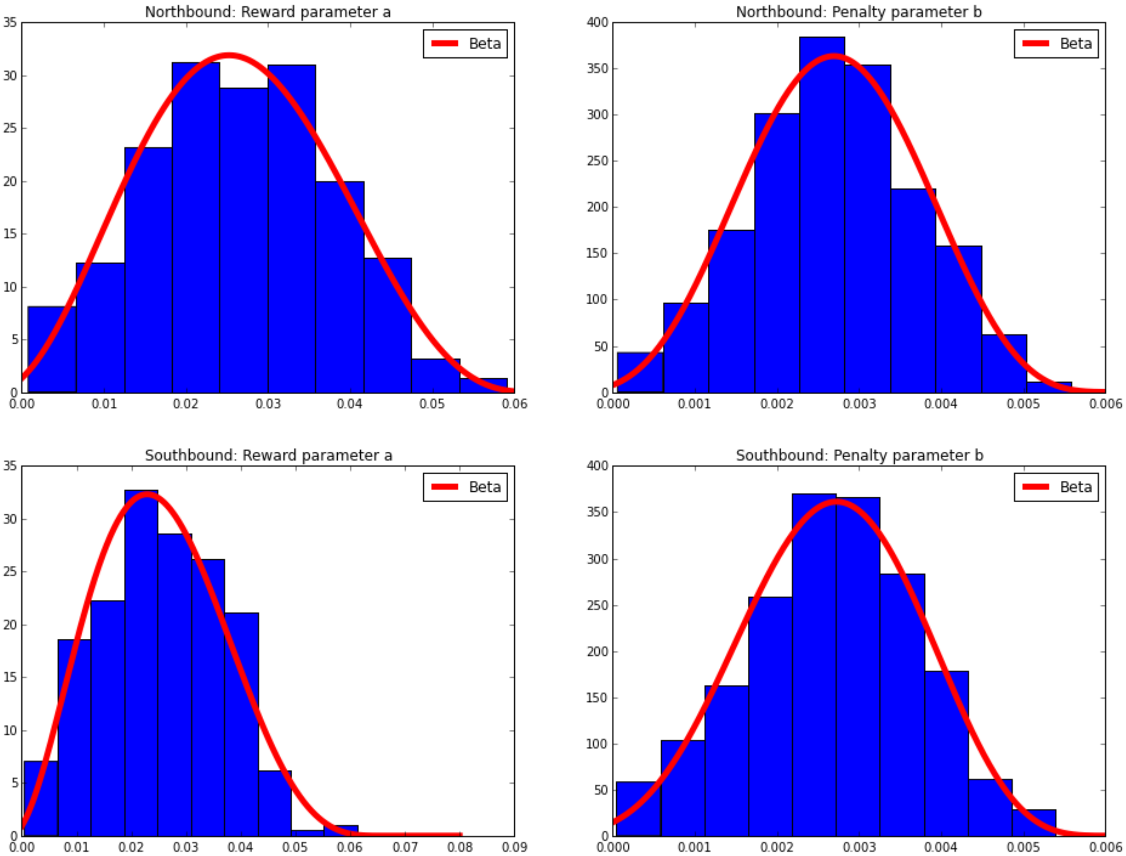

FIGURE 4: Histograms for Posterior Distributions of Learning Parameters

Probability functions using Beta Distribution, *B(.)*, are given as follows:

$$
\begin{aligned}
&\text{NB Reward Parameter:} \quad && P(a) = \frac{1}{B(3.41, 4.09)} \frac{(a+0.0041)^{2.41}(0.0668-a)^{3.09}}{0.0709^{4.5}} \\
&\text{NB Penalty Parameter:} \quad && P(a) = \frac{1}{B(5.34, 5.35)} \frac{(a+0.0008)^{4.34}(0.007-a)^{4.35}}{0.0078^{61.79}} \\
&\text{SB Reward Parameter:} \quad && P(a) = \frac{1}{B(3.08, 4.58)} \frac{(a+0.0022)^{2.08}(0.0682-a)^{3.58}}{0.0704^{6.66}} \\
&\text{SB Penalty Parameter:} \quad && P(a) = \frac{1}{B(6.74, 5.66)} \frac{(a+0.0015)^{5.74}(0.0077-a)^{4.66}}{0.0092^{11.4}}
\end{aligned}
\qquad (12)
$$

The results show that proposed Bayesian-SLA framework can successfully model frequent users' lane choice behavior in SR-167 HOT lanes. The distributions of learning parameters reveal the fact that users are less reactive in adapting their behavior. This could either mean or users do not consider actively changing their behavior or the system is operating close to equilibrium conditions that majority of the users have already learnt from their past behavior and found their quasi-optimal behavior. To test the latter hypothesis, Bayesian-SLA analysis is repeated with a subsample group. Since SR-167 HOT lanes pilot project had started five years prior[1] to time period that available data was collected, for long-time users there is a possibility that the effects of learning may be less apparent.

Modeling of New Users

A subsample which can also be called as "new users" is extracted using the timestamp records for toll tag ID readings. Among the earlier user sample with 50 choice occasions, tag IDs for whom at least 75 percent of their trips took place in the last three months are determined. In other words, during the first three month period these users only showed up occasionally and then they became frequent users within the last three month period. Moreover, starting conditions of these users are clearer since their experience with the facility is comparably limited before they started traveling regularly. Extracted subsamples constitute only 4% of previously used sample (19 out of 490 users) in southbound and about 2% (15 out of 630 users) in northbound.

Histogram plots show the distribution of learning parameters for this subsample of users in FIGURE 5. Set of mean values of reward and penalty parameters are (0.03,0.003) for northbound and (0.028,0.003) for southbound. Mean of reward parameters are slightly higher compared to the overall sample meaning that users in subsample group are more sensitive to positive changes in action outcomes. Penalty parameters, on the other hand, are the same.

Two-sample Kolmogorov-Smirnov statistics are calculated to test whether estimated learning parameter distributions are identical for the entire sample and subsample. The null hypothesis that two samples follow same continuous distribution is strongly rejected for reward parameters[2]. For the penalty parameters null hypothesis is still rejected[3], however the calculated p-value is higher in magnitude. These findings provide support to the notion that user familiarity with the facility is one of the key factors to take into account when modeling learning effects in route/lane choice (*44; 45*).

[1] SR-167 HOT lanes Pilot project started in 2008 (*41*). Available dataset is from the first half of year 2013.
[2] Kolmogorov-Smirnov P-values reward parameter distributions for NB: 1.5e-06, and for SB: 7.2e-06
[3] Kolmogorov-Smirnov P-values penalty parameter distributions for NB: 0.0012 for SB: 0.0728

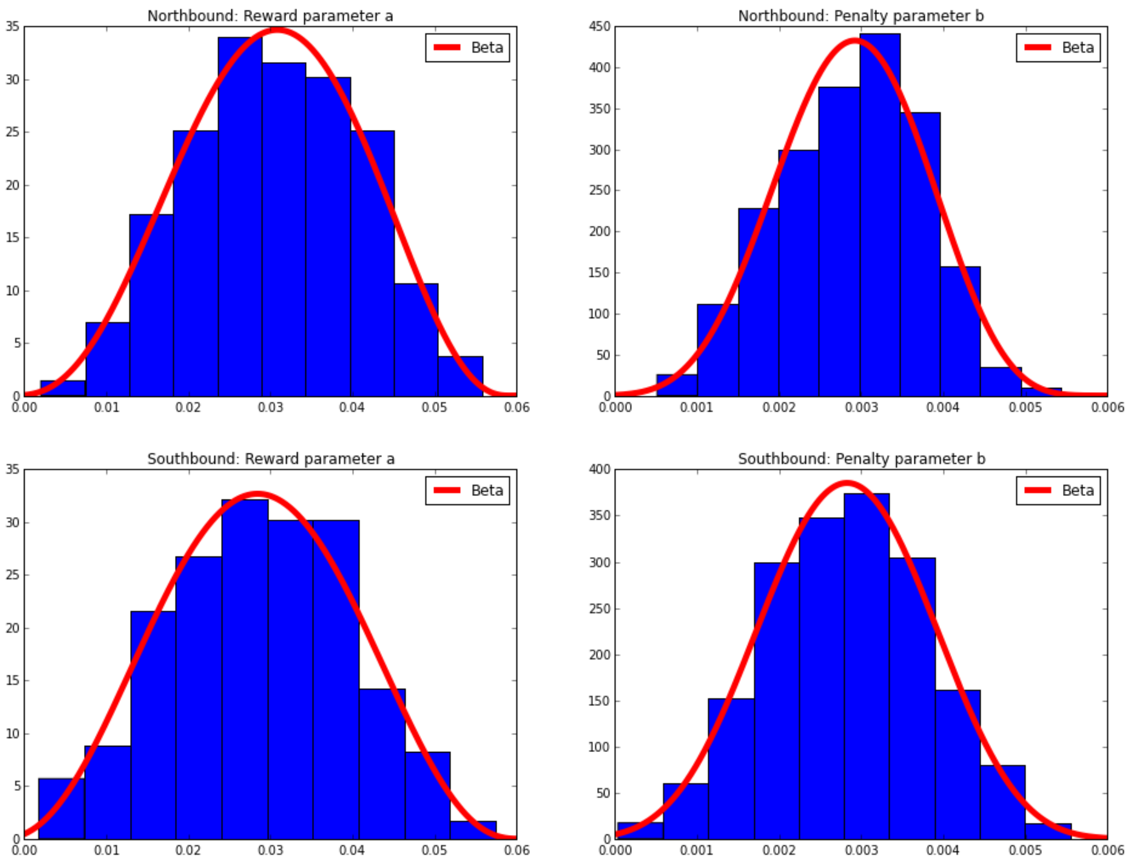

FIGURE 5: Subsample Group Histograms for Posterior Distributions of Learning Parameters

User heterogeneity can be investigated in more detail when socio-demographic factors such as gender, age, income information can be imputed into toll tag reader data. The lack of such information is one of the drawbacks of current analysis. For example a number of studies showed that women are more risk-averse than male and they are more reluctant to change their behavior in response to provided traffic information (*16; 46*). It is a fact that obtaining socio-economic data for such a large sample is quite difficult due to consumer privacy issues. It is also expected that data size would be significantly lower than the size used in current study in a case of voluntary participation which might possibly result in a biased sample. Although the effect of learning is clearly seen with action probability updating with Bayesian-SLA model, several other factors which are not observable with the existing data are yet to be explored.

CONCLUSIONS

This paper presents a modeling framework for investigating learning effects in lane choice behavior in MLs. Toll tag reader data for an approximately six-month period are used to analyze sequential choice behavior of users. Bayesian-SLA model is implemented with a MXL model and reward/penalty parameters are estimated for updating action probabilities of users. Developed model successfully simulates the observed traffic flow conditions. Estimated learning parameters indicate that ML users are less reactive to the feedback from previous experiences compared to the earlier estimations for other tolled facilities. Two possible explanations for this might be 1) Effects of habits and inertia in choices of users or 2) Quasi-equilibrium conditions in the facility in which most users are well informed and familiar with the traffic conditions. To test the second explanation, a subsample of non-familiar users is used to re-estimate learning parameters and reward parameter is found to be slightly higher for this subsample of new frequent users. Therefore it can be concluded that with increasing experience, users have less tendency to change their behavior.

Present study also supports the evidence of the usefulness of RP data in exploring driver behavior in MLs. One caveat of this data source, however, is the lack of information about differences in user socio-demographic characteristics. Thus, there is room for improvement in the provided approach in terms of implications of learning behavior, for example revealing possible causes for state-dependency in choices which is found in the MXL model.

For future directions, a longer time-span may be considered and more user groups can be identified depending on their familiarity with the system. Larger longitudinal data expectedly improve statistical explanatory power of similar learning models with more data points for the same users. An interesting case might be the effect of learning after a network disruption, i.e. addition of new ML segments, which can show the adaption process for various user groups. In such a case habit formation as a supporting factor and learning duration after which driver behavior will not change dramatically can also be examined.

ACKNOWLEDGMENTS AND DISCLAIMER

This research was partially funded by Polytechnic School of Engineering and Center for Urban Science and Progress (CUSP) of New York University. The authors acknowledge Mr. Tyler Patterson and Mrs. Michelle La Bolle from Washington State Department of Transportation for sharing SR-167 HOT lanes tolling and loop detector data.

REFERENCES

[1] Burris, M. W., and S. Patil. Estimating the benefits of managed lanes.In, 2009.
[2] Brownstone, D., A. Ghosh, T. F. Golob, C. Kazimi, and D. Van Amelsfort. Drivers' willingness-to-pay to reduce travel time: evidence from the San Diego I-15 congestion pricing project. *Transportation research part A: policy and practice,* Vol. 37, No. 4, 2003, pp. 373-387.
[3] Steimetz, S. S., and D. Brownstone. Estimating commuters'"value of time" with noisy data: a multiple imputation approach. *Transportation Research Part B: Methodological,* Vol. 39, No. 10, 2005, pp. 865-889.
[4] Lam, T. C., and K. A. Small. The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation Review,* Vol. 37, No. 2, 2001, pp. 231-251.

[5] Devarasetty, P. C., M. Burris, and W. Douglass Shaw. The value of travel time and reliability-evidence from a stated preference survey and actual usage. *Transportation research part A: policy and practice*, Vol. 46, No. 8, 2012, pp. 1227-1240.

[6] Small, K. A., C. Winston, and J. Yan. Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, Vol. 73, No. 4, 2005, pp. 1367-1382.

[7] Brownstone, D., and K. A. Small. Valuing time and reliability: assessing the evidence from road pricing demonstrations. *Transportation research part A: policy and practice*, Vol. 39, No. 4, 2005, pp. 279-293.

[8] He, X., H. X. Liu, and X. J. Cao. Estimating value of travel time and value of reliability using dynamic toll data.In *Transportation Research Board 91st Annual Meeting*, 2012.

[9] Burris, M., S. Nelson, P. Kelly, P. Gupta, and Y. Cho. Willingness to Pay for High-Occupancy Toll Lanes. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2297, No. 1, 2012, pp. 47-55.

[10] Burris, M., W. Arthur Jr, P. C. Devarasetty, J. McDonald, and G. J. Munoz. Understanding traveler behavior: The psychology behind managed lane use.In, 2012.

[11] Small, K. A., C. Winston, J. Yan, N. Baum-Snow, and J. A. Gómez-Ibáñez. Differentiated Road Pricing, Express Lanes, and Carpools: Exploiting Heterogeneous Preferences in Policy Design [with Comments]. *Brookings-Wharton Papers on Urban Affairs*, 2006, pp. 53-96.

[12] Liu, H. X., W. Recker, and A. Chen. Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. *Transportation research part A: policy and practice*, Vol. 38, No. 6, 2004, pp. 435-453.

[13] Liu, H. X., X. He, and W. Recker. Estimation of the time-dependency of values of travel time and its reliability from loop detector data. *Transportation Research Part B: Methodological*, Vol. 41, No. 4, 2007, pp. 448-461.

[14] Dahlgren, J. High-occupancy/toll lanes: where should they be implemented? *Transportation research part A: policy and practice*, Vol. 36, No. 3, 2002, pp. 239-255.

[15] Cao, X. J., L. Munnich, H. Liu, X. S. He, Z. Xu, and Y. A. Huang. Benefit and Cost Analysis of the I-394 MnPASS Program. 2012.

[16] Knorr, F., T. Chmura, and M. Schreckenberg. Route choice in the presence of a toll road: The role of pre-trip information and learning. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 27, 2014, pp. 44-55.

[17] Ozbay, K., A. Datta, and P. Kachroo. Modeling route choice behavior with stochastic learning automata. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1752, 2001, pp. 38-46.

[18] Horowitz, J. L. The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research Part B: Methodological*, Vol. 18, No. 1, 1984, pp. 13-28.

[19] Oh, J.-S., C. Cortés, and W. Recker. Effects of less-equilibrated data on travel choice model estimation. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1831, 2003, pp. 131-140.

[20] Mahmassani, H. S., and G.-L. Chang. Experiments with departure time choice dynamics of urban commuters. *Transportation Research Part B: Methodological*, Vol. 20, No. 4, 1986, pp. 297-320.

[21] Bogers, E., M. Bierlaire, and S. Hoogendoorn. Modeling learning in route choice. *Transportation Research Record: Journal of the Transportation Research Board*, 2007.

[22] Avineri, E., and J. N. Prashker. Sensitivity to travel time variability: travelers' learning perspective. *Transportation Research Part C: Emerging Technologies,* Vol. 13, No. 2, 2005, pp. 157-183.
[23] Yin, Y., S. Washburn, S., D. Wu, A. Kulshrestha, V. Modi, D. Michalaka, and J. Lu. Managed Lane Operations: Adjusted Time of Day Pricing vs. Near-Real Time Dynamic Pricing Volume 1: Dynamic Pricing and Operations of Managed Lanes. *Final Report to Florida Department of Transportation*, 2012.
[24] Jha, M., S. Madanat, and S. Peeta. Perception updating and day-to-day travel choice dynamics in traffic networks with information provision. *Transportation Research Part C: Emerging Technologies,* Vol. 6, No. 3, 1998, pp. 189-212.
[25] Zhang, L. *Search, Information, Learning, and Knowledge in Travel Decision-making: A Positive Approach for Travel Behavior and Demand Analysis*. ProQuest, 2006.
[26] Nakayama, S. Bayesian learning, day-to-day adjustment process, and stability of Wardrop equilibrium.In *Transportation and Traffic Theory 2009: Golden Jubilee,* Springer, 2009. pp. 425-440.
[27] Bogers, E., F. Viti, and S. Hoogendoorn. Joint modeling of advanced travel information service, habit, and learning impacts on route choice by laboratory simulator experiments. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1926, 2005, pp. 189-197.
[28] Mahmassani, H. S., and Y.-H. Liu. Dynamics of commuting decision behaviour under advanced traveller information systems. *Transportation Research Part C: Emerging Technologies,* Vol. 7, No. 2, 1999, pp. 91-107.
[29] Narendra, K. S., and M. A. Thathachar. Learning automata.In, Prentice Hall, 1989.
[30] Ozbay, K., A. Datta, and P. Kachroo. Application of stochastic learning automata for modeling departure time and route choice behavior. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1807, 2002, pp. 154-162.
[31] Yanmaz-Tuzel, O., and K. Ozbay. Modeling learning impacts on day-to-day travel choice.In *Transportation and Traffic Theory 2009: Golden Jubilee*, Springer, 2009. pp. 387-401.
[32] Alvarez, P. A. A methodology to estimate time varying user responses to travel time and travel time reliability in a road pricing environment. 2012.
[33] Ben-Elia, E., and Y. Shiftan. Which road do I take? A learning-based model of route-choice behavior with real-time information. *Transportation research part A: policy and practice,* Vol. 44, No. 4, 2010, pp. 249-264.
[34] Xiong, C. All about congestion: Modeling departure time dynamics and its integration with traffic models.In, 2011.
[35] Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *The journal of chemical physics,* Vol. 21, No. 6, 1953, pp. 1087-1092.
[36] Geman, S., and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* No. 6, 1984, pp. 721-741.
[37] Heidelberger, P., and P. D. Welch. Simulation run length control in the presence of an initial transient. *Operations Research,* Vol. 31, No. 6, 1983, pp. 1109-1144.
[38] Hensher, D. A., and W. H. Greene. The mixed logit model: the state of practice. *Transportation,* Vol. 30, No. 2, 2003, pp. 133-176.

[39] McFadden, D., and K. Train. Mixed MNL models for discrete response. *Journal of applied Econometrics,* Vol. 15, No. 5, 2000, pp. 447-470.
[40] Hess, S., M. Bierlaire, and J. W. Polak. Estimation of value of travel-time savings using mixed logit models. *Transportation research part A: policy and practice,* Vol. 39, No. 2, 2005, pp. 221-236.
[41] WSDOT. *SR 167 HOT Lanes Pilot Project: Fourth Annual Performance Summary*. http://www.wsdot.wa.gov/NR/rdonlyres/2DA072CA-7D4A-4876-A58B-64DBE152D48D/0/SR167_AnnualPerformanceSummary_080812_web.pdf.
[42] ---. *WSDOT Toll Division Annual Report FY 2014*. http://www.wsdot.wa.gov/NR/rdonlyres/9F45D2AE-4DB8-4AC0-A3F3-89ECDB970772/0/TollDivisionAnnualReport_2014_web.pdf. Accessed 8/24/2015.
[43] Morgul, E. F., and K. Ozbay. The Value of Schedule Delays by Time-of-day: Evidence from Actual Usage Data from SR 167 HOT Lanes.In *Transportation Research Record (forthcoming)*, 2015.
[44] Adler, J. L., W. W. Recker, and M. G. McNally. Using interactive simulation to model driver behavior under ATIS. *University of California Transportation Center*, 1992.
[45] Adler, J. L. Investigating the learning effects of route guidance and traffic advisories on route choice behavior. *Transportation Research Part C: Emerging Technologies,* Vol. 9, No. 1, 2001, pp. 1-14.
[46] Mannering, F., S.-G. Kim, W. Barfield, and L. Ng. Statistical analysis of commuters' route, mode, and departure time flexibility. *Transportation Research Part C: Emerging Technologies*, Vol. 2, No. 1, 1994, pp. 35-47.