MODELLING HOT LANE OPERATION UNDER MULTIPLE POLICY GOALS

Wenbo Fan
School of Transportation and Logistics, Southwest Jiaotong University. National United Engineering Laboratory of Integrated and Intelligent Transportation. 111 Erhuan Road, Beiyiduan, Chengdu, 610031, China; National Center for Smart Growth Research and Education, University of Maryland, College Park, USA; wbfan@swjtu.edu.cn, wbfan@umd.edu

Xinguo Jiang, Corresponding author
School of Transportation and Logistics, Southwest Jiaotong University; National United Engineering Laboratory of Integrated and Intelligent Transportation. 111 Erhuan Road, Beiyiduan, Chengdu, 610031, China. xjiang@swjtu.edu.cn;

Word Count: 4574 Texts+5Figures*250+ 5 Tables*250= 7074 words
Abstracts: 173 words
Submission Date: August 1 2015
Submitted for Presentation at the 95th Annual Meeting of the Transportation Research Board and Publication in Transportation Research Record.
ABSTRACT

Understanding the interaction between the high occupancy/toll (HOT) lane operators and the roadway users is crucial to design and succeed in a HOT project. The paper proposes a bilevel programming framework to model the HOT operations: the upper-level model simulates how the operators develop the pricing strategies under various policy goals (e.g., stable flow speeds and social welfare maximization) and the lower-level model is related to the travelers’ mode- and lane-choice, which can be developed in the form of a multinomial logit model, a nested logit model, or a sequential model (i.e., a Logit and a deterministic user equilibrium). The paper compares the above-stated choice models and conducts the scenario analyses to evaluate the highway performances (e.g., optimal toll level and average travel speed) of various HOT scenarios. Results show that: (i) the Logit-form choice models may not fully reflect the impacts of varying tolls; (ii) the homogeneity assumption may lead to biased results of the effectiveness of HOT lanes; and (iii) different policy goals have significant influences on the HOT lane operations.

Keywords: congestion pricing; high occupancy/toll lane; bilevel programming model; managed lane; equilibrium model; heterogeneous commuters
INTRODUCTION
In the past two decades, the U.S. Department of Transportation (DOT) has been promoting a national Value Pricing Pilot (VPP) Program to encourage the pricing measures for traffic congestion management (1). As one of the most widely applied VPP programs, the high occupancy/toll (HOT) lane concept was firstly implemented in 1995 on State Route 91 in Orange County (California) (2). Since then, HOT lanes have been successfully employed in eight cities (e.g., San Diego, Houston, and Miami) with additional fourteen cities having the HOT plans (3).

The concept of HOT lane is actually a combination of congestion pricing and high occupancy vehicle (HOV) lane (4). Although HOV lanes had been rapidly constructed, some projects were criticized due to its low utilization (5). It is found that it is very difficult to optimize the utilization of HOV lanes if only targeting the occupancy requirement (i.e., vehicles with two or more occupants) (6). Thus, HOT lane is proposed to allow single occupancy vehicles (SOV) to pay for the use of the exclusive lane, while it remains free of charge to the HOVs. The tolls are adjustable to maintain a desirable level of service (LOS) on the HOT lane. Modeling HOT operation, however, becomes more complicated as opposed to individual strategies of HOV lanes and traditional CP, due to the complex interactions between the HOT lane operator and heterogeneous roadway users (e.g., SOV and HOV users with different value of time (VOT)).

The objective of the paper is to model and examine the performance of the HOT lane operation under various conditions with respect to the policy goals, the right-of-way of the HOT lane, the heterogeneity of commuters, etc. The paper is organized as follows: firstly the literature review is presented and a bilevel model of HOT lane operations is established with an optimization model at the upper level and an equilibrium model at the lower level; and then the paper compares three forms of travelers’ choice models and proposes eighteen scenarios on an experimental highway segment and main conclusions are drawn in the final section.

LITERATURE REVIEW
A majority of existing studies evaluated the HOT operation based on travelers’ lane choice models. For instance, Abdelghany et al. (7) compared the performance (e.g., lane utilization and pricing structure) of various designs and operations of HOT lanes in a highway network in Texas; Dahlgren (5) examined the congestion reduction effects on three approaches of adding an additional highway lane (i.e., an HOV lane, an HOT lane, and a regular lane) to an existing freeway; Zhang et al. (8) developed a feedback-control tolling algorithm to optimize the HOT lane operation in response to varying travel demand; and Lou et al. (9) introduced a self-learning approach in the HOT lane operations to estimate the travelers’ willingness-to-pay and accordingly optimize the pricing strategies.

In recent years, the static HOT operation studies had been extended into the dynamic context. For instance, Yin and Lou (10) developed two dynamic tolling strategies (i.e., a feed-back control and a reactive self-learning approach) to maximize the throughput of the HOT facilities; Jang and Chung (11) examined the dynamic pricing strategy considering difference in travelers’ value of time (VOT); and Gardner et al. (2) compared the impacts of static and dynamic tolling on travelers’ lane choices. In addition, the lane-choice models had also been embedded in simulation tools (e.g., CORSIM, VISSIM, and DYNASMART) to enhance the applicability (8,12).
To capture the demand shifting among travel modes (e.g., SOV, HOV and transit), travelers’ mode-choice models had been introduced into the HOT models. Burris and Xu (13), Chum and Burris (14), and Burris et al. (15) examined the potential SOV demand for HOT lanes in Texas. Erhardt et al. (16) simultaneously modeled the travelers’ mode and lane choices and evaluated the impacts of converting HOV lanes to HOT lanes. Murray et al. (17) combined the mode-choice model with a dynamic traffic-simulation method to assess the HOT lane usage.

Most of the above-mentioned studies, however, assumed that the travelers were “identical” in their social-economical characteristics and thus made the travel choices in the exact same manner. According to the surveys conducted by Brownstone and Small (18), Lam and Small (19) and recently Wood et al. (20), there is a substantial heterogeneity in the commuters’ value of time on the HOT lane projects. Small and Yan (21) found that accounting for the heterogeneous travelers was important in evaluating and improving the HOT operations. Recently, a few HOT lane studies have emerged to address the heterogeneous travelers either by dividing them into discrete groups according to their VOTs or by employing a continuous distributed VOT across the population (2,11,12,22).

In addition, few attentions have been directed to explicitly examine the impacts of policy goals, which may significantly influence the pricing strategies of the HOT operators. In practice, HOT lane projects often adopt a policy goal as simple as to provide a free-flow traffic service or a stable travel speed on the managed lanes (4). This level-of-service (LOS) related goal is easy to implement and evaluate, but may not reflect the interests of different stakeholders. For example, a private operator would pursue to maximize the profit during its franchise period; and a local transportation agency may expect the HOT project to minimize the congestion cost on the highway of interest. From the perspective of the whole society, social welfare maximization represents a theoretical system optimization (21). Other possible objectives include delay minimization (11) and traffic volume maximization (2). Under the various policy goals, understanding the HOT operations is crucial to design and succeed in a HOT project.

MODEL FORMULATION

The operation of HOT lane involves two parties: the HOT operators and the commuters. They interact in a way that the operator pursues policy goals via the operational activities (e.g., pricing), and the commuters make and/or adjust their travel decisions (e.g., mode and lane choices) according to the toll level and the traffic conditions. The hierarchical interaction is depicted in Figure 1 and formulated as a bilevel programming model.
FIGURE 1 The hierarchical interaction between HOT operator and commuters.

Table 1 summarizes the notations used in the paper.

**TABLE 1 Notation List.**

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Descriptions</th>
<th>Symbols</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Cc$</td>
<td>Congestion costs of highway users (dollar/year)</td>
<td>$Sw$</td>
<td>Social welfare obtained by all agents of the highway (dollar/year)</td>
</tr>
<tr>
<td>$C_l$</td>
<td>Capacity of highway lane (vehicles/hour)</td>
<td>$s_0, s_l$</td>
<td>Free-flow speed and average speed on lane $l$ (km/h)</td>
</tr>
<tr>
<td>$C_p$</td>
<td>Unit cost of HOT project ($/km$)</td>
<td>$Tr$</td>
<td>Toll revenues collected on HOT lane ($/year$)</td>
</tr>
<tr>
<td>$c$</td>
<td>Fuel cost and tax of vehicles ($/delay$ hour)</td>
<td>$T_{m,l}^i$</td>
<td>Generalized travel costs of user $i$ on lane $l$ ($/dollar$)</td>
</tr>
<tr>
<td>$D^{-1}()$</td>
<td>Reverse demand function</td>
<td>$t_0$</td>
<td>Travel time at free flow speed on lane $l$ (hour)</td>
</tr>
<tr>
<td>$G()$</td>
<td>Objective function</td>
<td>$t_l$</td>
<td>Actual travel time on lane $l$ (hour)</td>
</tr>
<tr>
<td>$i$</td>
<td>Type of commuters, $i=[1, 2, 3]$</td>
<td>$\nu_{m,l}^i$</td>
<td>Volume of commuters on lane $l$ (person/hour)</td>
</tr>
<tr>
<td>$L$</td>
<td>Length of the highway (km)</td>
<td>$\alpha_l$</td>
<td>Parameter in demand function</td>
</tr>
<tr>
<td>$l$</td>
<td>Index of the highway lane, $l=[1, 2]$</td>
<td>$\beta$</td>
<td>Average occupancy of HOV (persons/vehicle)</td>
</tr>
<tr>
<td>$m$</td>
<td>Travel mode, $m=[SOV, HOV]$</td>
<td>$\lambda_l^i$</td>
<td>VOT of $i$th commuter ($$/hour$$)</td>
</tr>
<tr>
<td>$Q'$</td>
<td>Potential demand of the $i$th type of commuters</td>
<td>$\tau$</td>
<td>The toll on HOT lane ($$/vehicle$$)</td>
</tr>
</tbody>
</table>
Assumptions

In order to facilitate the model formulations, the following assumptions are made:

A1. An experimental highway segment suffers from a chronic traffic congestion during the study period (e.g., the morning peak hour) and there are two types of highway lanes (indexed by \( l = [1, 2] \)): the HOT and general-purpose (GP) lanes.

A2. Commuters are heterogeneous due to their social-economic characteristics, and the heterogeneity is quantified with a discrete set of VOT (indexed by \( i = [1, 2, 3] \)) and correspondingly commuters are classified into high-, median-, and low-VOT groups (23).

A3. Commuters have two options of travel modes to accomplish their trips: the solo-driving mode (or single-occupied vehicle, SOV) and the carpool mode (also referred as high-occupied vehicle, HOV). Solo driver in the SOV has to pay for the use of the HOT lane, while carpoolers in the HOV are free of charge on all highway lanes.

A4. The actual demand of commuters is not fixed but elastic to their travel costs in the highway (24). The elasticity reflects the fact that commuters may switch from/to alternative routes, modes, and departure times in response to various toll levels and traffic conditions.

Upper-level objective function

Four objective functions are formulated in equations (1)-(4) to capture the operating behaviors of HOT operators under different policy goals, i.e., the LOS maintenance, the toll revenue maximization (TRM), the congestion cost minimization (CCM), and the social welfare maximization (SWM):

First, the LOS objective criterion, which is the most commonly adopted policy goal in practice (25), usually aims to maintain a stable traffic flow at a desirable speed (e.g., \( s_0 = 60 \) km/h) on the HOT lane (\( l = 1 \)), and can be given by:

\[
\min_{\tau} S(\tau) = |v^{\prime}_i(\tau) - s_0|, \quad l = 1
\]

Second, the TRM criterion tends to maximize the annual net revenue, which equals to the total toll revenue minus the annualized construction and operation costs of the HOT lane projects:

\[
\max_{\tau} Tr(\tau) = 250 \cdot 6 \cdot \tau \cdot \sum v^{\prime}_{SOV,i}(\tau) - C_p \cdot L, \quad l = 1
\]

where 250 is the total number of weekdays in a year and 6 is the average peak hours during a weekday.

Third, the congestion cost (\( Cc \)) is the total costs of time delay experienced by the commuters on the congested highway compared to the state of free flow speed (FFS). And the CCM criterion is given by

\[
\min_{\tau} Cc(\tau) = 250 \cdot 6 \cdot \sum v^{\prime}_{m,i}(\tau) \cdot \lambda \cdot (t_i - t_0)
\]

And fourth, the social welfare (\( Sw \)) is the sum of consumer surplus (CS) and producer surplus (PS). According to the work by Wichiensin et al. (26), CS is defined
as a sum of all commuters’ marginal surplus minus their total travel costs, and PS
equals to the HOT operators’ net revenues. The SWM criterion represents an efficient
allocation of the economic resources (the highway capacity in this case) via the
interaction between the supply and demand (27), and is expressed as
\[
\max_{\tau} \text{Sw}(\tau) = 250 \cdot 6 \sum_{j} \int_{0}^{\tau} D^{-1}(x) dx - \sum_{m,i} v_{m,i}(\tau) \cdot T_{m,i} + Tr(\tau)
\] (4)

Note that the actual demand \( (q') \), the volumes \( (v_{m,i}) \), and the generalized travel
costs \( (T_{m,i}) \) of the commuters are obtained from the following lower-level models.

**Lower-level travelers’ choice models**

According the toll level and traffic conditions, commuters make their travel decisions
in terms of mode choices (i.e., SOV and HOV modes) and lane choices (HOT and GP
lanes). Most of the above-mentioned studies simulated the travelers’ mode and lane
choices with the Logit-formula models, e.g., binomial logit models (8,9,19), nested
logit models (15), and mixed logit models (13,16,28). These models had been well
applied to the HOT lane evaluation on a single corridor. The applicability of the Logit
models is, however, limited for the network-scale studies, because numerating all the
alternative travel options (e.g., modes and routes/lanes) may lead to a considerable
complexity in the Logit formulation and the computation. To enhance the HOT lane
modelling in the network scale, a few studies made use of the existing travel demand
forecasting/simulation tools (e.g., the Maryland Statewide Transportation Model and
DTALite) (28,29), where the travelers’ lane choices are treated in a deterministic
manner (e.g., all-or-nothing assignment) and the mode choices still follow a Logit
manner. For the sake of comparisons, the paper presents three typical types of
commuters’ choice models, i.e., the multinomial logit model, the nested logit model,
and the Logit & user equilibrium (Logit & UE) model.

**Multinomial logit model**

The multinomial logit model (MNL) assumes that (i) commuters simultaneously
make their mode and lane choices; (ii) the error components (e.g., commuters’
perception errors) of the generalized travel costs (GTC) follow Gumbel distribution
and are identically and independently distributed across alternatives (30). And the
MNL is a function of commuters’ GTC, given by:
\[
v_{m,i} = q' \cdot \exp\left(-T_{m,i}\right) / \sum_{m,i} \exp\left(-T_{m,i}\right)
\] (5a)

According to the assumption A4, the actual demand of commuters is
hypothized to be a continuous monotone decreasing function of the GTC, given as:
\[
q' = D(T_{m,i}) = q' \cdot \exp\left(-\alpha' \cdot \ln\left(\sum_{m,i} \exp\left(-T_{m,i}\right)\right)\right),
\] (5b)

where parameters \( \alpha' \) reflects the elasticity of the \( i \)th type of commuters’ demand with
respect to the GTC.

**Nested logit model**

The MNL model has been widely criticized for its Independence of Irrelevant
Alternatives (IIA) property. The NL model classifies alternative travel options that are
similar to each other with respect to the system characteristics (e.g., travel pattern of
SOV and HOV modes) as opposed to other alternatives (e.g., HOT and GP lanes).
And thus, the mode- and lane-choice models are given by equations 6(a-b),
respectively:

\[ v'_m = \begin{cases} 
  v'_{\text{ov}} = q' \exp\left(-\ln\left(\sum_l \exp\left(T_{\text{ov},l}^i\right)\right)\right) / \sum_l \exp\left(-\ln\left(\sum_l \exp\left(T_{\text{ov},l}^i\right)\right)\right) \\
  v'_{\text{ov}} = q' - v'_{\text{ov}} 
\end{cases} \]  

(6a)

\[ v'_{m,l} = v'_m \cdot \exp\left(-T_{m,l}^i\right) / \sum_l \exp\left(-T_{m,l}^i\right) \]  

(6b)

According to the property of the Logit model, the elastic demand function is converted to:

\[ q' = D(T_{m,l}^i) = Q \cdot \exp\left(-\alpha' \cdot \ln\left(\sum_m \exp\left(\ln\left(\sum_l \exp\left(T_{m,l}^i\right)\right)\right)\right)\right) \]  

(6c)

Logit model and UE model

For a long-term planning, it can be assumed that through the day-to-day commuting experience and appropriate information channels, commuters are well familiar with the toll levels and traffic conditions on the highway, and thus make their lane choices in a deterministic manner to minimize their GTC. An equilibrium state can be reached when each commuter has no alternative to minimize his/her GTC, and consequently the chosen highway lanes have the minimum GTC. Such a state is the user equilibrium (UE) condition as given below:

\[ v'_{m,l} = \begin{cases} 
  v'_m & \text{if } T_{m,l}^i = \min_j(T_{m,l}^i) \\
  0 & \text{if } T_{m,l}^i > \min_j(T_{m,l}^i) 
\end{cases} \]  

(7a)

Due to the random errors in commuters’ preferences and perceptions, their mode choices may still follow a stochastic manner in the following Logit form:

\[ v'_m = \begin{cases} 
  v'_{\text{ov}} = q' \cdot \exp\left(-\min_j(T_{\text{ov},l}^i)\right) / \sum_j \exp\left(-\min_j(T_{\text{ov},l}^i)\right) \\
  v'_{\text{ov}} = q' - v'_{\text{ov}} 
\end{cases} \]  

(7b)

And correspondingly, the elastic demand function is rewritten by:

\[ q' = D(T_{m,l}^i) = Q \cdot \exp\left(-\alpha' \cdot \ln\left(\sum_m \exp\left(\min_j(T_{m,l}^i)\right)\right)\right) \]  

(7c)

Generalized travel cost

Commuters’ GTC is mainly comprised of the travel time (\(t_i\), convertible into the monetary unit), fuel and tax costs, additional toll (\(\tau\)) for the solo drivers using the HOT lanes, and extra time (\(\Delta\)) for carpoolers who have to pick up or drop off the passengers. The GTC is given by

\[ T_{\text{SOV},l}^i = \begin{cases} 
  \lambda^i \cdot t_i + c \cdot t_i + \tau & l = \text{HOT lane} \\
  \lambda^i \cdot t_i + c \cdot t_i & l = \text{GP lanes} 
\end{cases} \]  

(8a)

\[ T_{\text{HOV},l}^i = \lambda^i \cdot t_i + \frac{c}{\beta} \cdot t_i + \lambda^i \cdot \Delta, \]  

(8b)

where the toll level (\(\tau\)) on the HOT lanes is set by the upper-level operators with the objectives in equations (1-4); the unit fuel and tax cost, \(c\), can be obtained from the empirical studies; and the travel time (\(t_i\)) on lane \(l\) is expressed in the Bureau of Public Roads (BPR) form of the volume-delay function (31):

\[ t_i = t_0 \left(1 + 0.15 \cdot \left(\sum_j (v_{\text{SOV},j}^i + v_{\text{HOV},j}^i) / \beta \right) / C_j \right)^4, \]  

(9)
where the average occupancy $\beta$ of HOVs can be calibrated by previous traffic surveys on the highway, and $v_{\text{SOV},i} + v_{\text{HOV},/i} / \beta$ converts the passenger volume into vehicles.

**Bilevel programming model and solution algorithm**

Based on the above-given models (equations (1-9)), a bilevel programming model of HOT lane operation is given:

$$\text{Upper level: } G(\tau) = \begin{cases} 
\text{min } S(\tau), & \text{LOS criterion} \\
\text{max } T_r(\tau), & \text{TRM criterion} \\
\text{min } C_c(\tau), & \text{CCM criterion} \\
\text{max } S_w(\tau), & \text{SWM criterion}
\end{cases} \quad (10a)$$

Subject to:

$$\tau \geq 0, \quad (10b)$$

Lower level:

$$\begin{cases} 
\text{Multinomial model: } & \text{Equations(5a-b)} \\
\text{Nested model: } & \text{Equations(6a-c)} \\
\text{Logit & UE: } & \min \sum_{i,m} \int_0^{v_i} T_{wi}(v_i(\tau))dv - \sum_{i} \int_0^{q_i} D^{-1}(w)dw \\
& + \sum_{i,m} v_m^i \left( \ln \left( \frac{v_m^i}{q_i} \right) - 1 \right) \quad (10c)
\end{cases}$$

Subject to:

$$q^i = \sum_m v_m^i, \quad (10d)$$

$$v_m^i \geq 0, \quad (10e)$$

where in equation (10c) the minimization problem is equivalent to the user equilibrium with modal choice and elastic demand (30); $v_i(\tau) = \sum_n v_{m,i} + \sum_{m} v_{m,i} / \beta$; $v_m = \sum_n v_{m,i}$; and $D^{-1}(\cdot)$ is the inverse functions of the elastic demand functions (i.e., equations (5b, 6c, 7c)).

The above bilevel programming model contains only one decision variable (i.e., the toll level), and thus can be solved by the following numerical increment method embedded with the method of successive averages (MSA), which is used to find solutions for the lower-level equilibrium problem. The solution algorithm is programed in the following steps:

1. Initialize the decision variable toll level $\tau = 0$ and set the incremental step (e.g., 0.1) and the loop variable $n=1$.
2. Substitute the current toll level $\tau^{(n)}$ into the lower-level algorithm.
   2.1 initialize the input basic data (e.g., $Q^i, \lambda^i, \beta$, and $\alpha^i$). Set inner loop variable $k=1$.
   2.2 run the elastic demand functions (5b, 6c, 7c) and compute actual travel demand.
   2.3 calculate travel times on the highway lanes ($t_i$) and commuters’ GTC ($T_{wi}$) by equations (8-9).
   2.4 compute the auxiliary $v_{m,i}$ by commuters’ mode and lane models (equations (5-7)).
2.5 update $v_{m,j}^{(k)}$ by the following MSA equation (11).

$$v_{m,j}^{(k+1)} = v_{m,j}^{(k)} + \left( v_{m,j}^{(k)} - v_{m,j}^{(k-1)} \right) / k$$

(11)

2.6 break the inner algorithm if the following equation (12) is satisfied or the inner loop reaches the maximum iteration (e.g., 3000), output $v_{m,j}^{(n)}, T_{m,j}^{(n)}, \lambda_{j}^{(n)}, t_{i}^{(n)}$ and then go to Step 3; otherwise go to Step 2.1.

$$\sum_{i,m,j} v_{m,j}^{(k)} - v_{m,j}^{(k-1)} \leq \varepsilon,$$

(12)

where $\varepsilon$ is a pre-specified minimum value, e.g., 0.001.

3. Evaluate the fitness of the current solution with equations (1-4), if $G^{(n)} = \max_{n=1} G$, update the optimal toll $\tau^{(n)}$ and the highway performance indicators (e.g., travel speed and modal split).

4. Generate the new toll level with the numerical increment equation, $\tau^{(n+1)} = \tau^{(n)} + 0.1$.

5. Terminate the algorithm if the specified iteration limit (e.g., 100) is reached; otherwise, set $n=n+1$ and go to Step 2.

**NUMERICAL STUDIES**

**Basic settings**

Suppose that on an experimental highway (as shown in Figure 2), the HOT project is planned on a segment with length of $L=20$ km (i.e., the average length of HOT facilities in US (I)), three lanes per direction, free flow speed of $t_0=80$ km/h, and the capacity of $C=2000$ vehicles/hour-lane. According to the available right-of-way (ROW) of the HOT lane, there are generally two ways to implement the HOT project as depicted in Figure 2: (i) adding a new lane as the HOT lane (denoted as HOT_NL) and (ii) making use of one existing highway lane (HOT_EL). Note that in practices most of HOT projects are converted from existing HOV lane, which had been often built on newly added right-of-way. Thus HOV-to-HOV projects are treated as HOT_NL scenario with consideration of additional investment cost. Although HOT_EL scenario has not been seen in reality due to low public acceptance, it is still worthwhile to study the economical feasibility.

![Diagram of HOT lane introduction](image_url)

**FIGURE. 2 Two ways of introducing a HOT lane**

During the peak hours, a one-way potential travel demand is $Q=[4,000, 4,000, 4,000]$ persons/h for three groups of commuters, whose VOTs are $\lambda = [25, 15, 5]$ dollar/h. Note that actual travel demand should be obtained by equations (5b, 6c, 7c) and will not be so high as compared to the highway capacity (i.e., 2000 vehicles/hour-lane). According to the estimation of DeCorla-Souza (32), the unit fuel and tax cost is $c=4.2$ dollar per delayed vehicle hour and the construction and operation costs ($C_p$) of typical HOT_EL and HOT_NL projects are approximately
0.18 and 0.47 million dollars per km, respectively. Based on the FHWA’s report (33),
the average occupancy of carpools is approximately $\beta=2$ (implying very few
HOV3+). And the extra time for carpoolers is assumed to be $\Delta_s = 20$ min.

Comparison of travelers’ choice models
The above-given choice models (i.e., MNL, NL, and Logit & UE) are applied to the
do-nothing case, respectively. Table 1 summarizes the equilibrium results and it is
observed that the equilibrium results of MNL and NL are almost the same. Due to the
property of Logit model (30), the travel speeds on the highway lanes are not equal at
the equilibrium state of MNL and NL. With a distinct lane choice model, the Logit &
UE model leads to different assignment results (i.e., travel speed and lane flow
distribution) and consequently divergent congestion cost from the MNL and NL
models.

**TABLE 1 Results of different choice models under do-nothing case.**

<table>
<thead>
<tr>
<th>Items</th>
<th>Travel speed</th>
<th>Mode split (%)</th>
<th>Lane flow distribution (%)</th>
<th>Heterogeneous users distribution (%)</th>
<th>Congestion cost ($\text{million}$$\delta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNL</td>
<td>[45,49]</td>
<td>[93,7]</td>
<td>[35,65]</td>
<td>[36,30,34]</td>
<td>34</td>
</tr>
<tr>
<td>NL</td>
<td>[46,49]</td>
<td>[92,8]</td>
<td>[35,65]</td>
<td>[36,30,34]</td>
<td>34</td>
</tr>
<tr>
<td>Logit &amp; UE</td>
<td>[46,46]</td>
<td>[92,8]</td>
<td>[33,67]</td>
<td>[35,30,34]</td>
<td>39</td>
</tr>
</tbody>
</table>

Note: \(^1\)values in bracket are for the median lane (potential HOT lane) and other highway lanes; \(^2\)values in bracket are for SOV and HOV modes, respectively; \(^3\)values in bracket are for high-, median-, and low-VOT commuters, respectively.

Furthermore, to compare the results under the four objective functions, the
above-proposed bilevel programming model is employed on HOT scenarios (i.e.,
HOT_EL and HOT_NL), and the optimal solutions are presented in Table 2. It is
found that the optimal tolls differ significantly under various scenarios (i.e.,
combinations of HOT scenarios, objective functions, and choice models). The toll
levels tend to be less in HOT_NL cases than HOT_EL cases due to additional
highway capacity of the new lane. Comparatively, the results imply that the selection
of choice models (as well as policy goals) may have larger influence to the
performance of HOT_EL scenarios than HOT_NL. In addition, it is noticeable that
the MNL and NL models result in unreasonable high level tolls (more than 10 dollar)
under the CCM criterion.

**TABLE 2 Optimal tolls of HOT schemes with different choice models.**

<table>
<thead>
<tr>
<th>Items</th>
<th>LOS criterion</th>
<th>TRM criterion</th>
<th>CCM criterion</th>
<th>SWM criterion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HOT_EL</td>
<td>HOT_NL</td>
<td>HOT_EL</td>
<td>HOT_NL</td>
</tr>
<tr>
<td>MNL</td>
<td>3.2</td>
<td>0.6</td>
<td>4.4</td>
<td>3</td>
</tr>
<tr>
<td>NL</td>
<td>2.8</td>
<td>0.5</td>
<td>3.4</td>
<td>2.3</td>
</tr>
<tr>
<td>Logit &amp; UE</td>
<td>4.1</td>
<td>0</td>
<td>3.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Figure 3 depicts the relationship between the congestion cost and the toll
levels for MNL, NL, and Logit & UE models, respectively. It is seen that as the toll
rises, the congestion cost of HOT_EL with Logit & UE model first decreases and
then increases with the inverse point corresponding to the optimal toll ($2.8$). The result is intuitive that the tolling releases the traffic congestion on HOT lane at the
first stage; however, the shifting traffic from HOT lane ultimately enhances the
congestion state on GP lanes. While the toll continues to rise, the Logit & UE curve
tends to be flat, because most of the SOV demand is driven off the HOT lane and the
pricing strategy becomes ineffective in adjusting traffic flows. The MNL and NL
models, however, demonstrate a strikingly different phenomenon due to Logit-type
equations (5,6). Although the part of concave curve is also observed for the MNL and
NL models, their congestion costs continue to drop with the toll levels. The unrealistic
results are caused by the fact that the rising tolls remain functional in the elastic
demand equations (5b, 6c), resulting in continuously decreasing actual demand as
well as the congestion cost. This explains the optimal toll rises unreasonable high
under the CCM criterion. Therefore, it is suggested that the selection of different
choice models should be made with caution, the Logit & UE model may be more
reliable to apply to HOT lane simulation and evaluation.

FIGURE 3 Congestion costs with three different choice models.

Scenarios analysis with heterogeneity/homogeneity assumptions
Based on the Logit & UE model, the HOT scenarios (with heterogeneity/homogeneity
assumptions) are compared between the do-nothing cases and the HOV scenarios.
The comparability between heterogeneity and homogeneity scenarios is guaranteed by
assigning homogeneous users with the VOT equal to the average of discrete VOTs.
Table 3 summarizes the highway performances in terms of optimal toll, travel speed,
and mode split. It is demonstrated that the homogeneity assumption leads to biased
results due to the inability to capture the distinct choices of commuters: specifically,
the optimal toll on HOT lane tends to be underestimated, but the travel speed appears
to be overestimated as compared with the scenarios with heterogeneous users. The
estimation of HOV-mode share is conspicuously lower than that under the
heterogeneity assumption.

TABLE 3 Highway performances with heterogeneity/homogeneity assumptions.
The optimal results of HOT scenarios with heterogeneity/homogeneity assumptions are shown in Table 4, including the net toll revenue (TR), social welfare (SW), congestion cost (CC). It is seen that the homogeneity assumption results in considerably lower values of TR, SW, and CC. It is also noted that although HOV_NL scenarios yield relatively higher SW than HOV_EL scenarios, they appear to be financially unsustainable due to the negative TR under all policy goals. Furthermore, Table 5 summarizes the feasibilities of HOV and HOT scenarios from the financial, economical, and congestion-related perspectives. It is observed that the heterogeneity/homogeneity assumptions may lead to the opposite conclusions of the feasibilities. For instance, the HOT_EL scenario shows good financial and economical potentials with positive congestion mitigation under the heterogeneity assumption, while the results state otherwise under the homogeneity assumption. The results suggest that the heterogeneity/homogeneity assumptions play a critical role in the evaluation of HOT/HOV lane projects, which should be validated with justified surveys and studies to avoid biased or even delusive results.

### Table 4 Optimal results with heterogeneity/homogeneity assumptions.

<table>
<thead>
<tr>
<th>Scenarios Objectives</th>
<th>Heterogeneous users</th>
<th>Homogeneous users</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TR ($million)</td>
<td>SW ($million)</td>
</tr>
<tr>
<td>Do-nothing</td>
<td>0</td>
<td>622</td>
</tr>
<tr>
<td>On an existing lane</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOV_EL</td>
<td>0</td>
<td>642</td>
</tr>
<tr>
<td>HOT_EL_TRM</td>
<td>6</td>
<td>639</td>
</tr>
<tr>
<td>HOT_EL_SWM</td>
<td>4</td>
<td>641</td>
</tr>
<tr>
<td>HOT_EL_CCM</td>
<td>5</td>
<td>636</td>
</tr>
<tr>
<td>HOT_EL_LOS</td>
<td>6</td>
<td>640</td>
</tr>
<tr>
<td>On a new lane</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOV_NL</td>
<td>-6</td>
<td>676</td>
</tr>
</tbody>
</table>

Note: The VOT of homogeneous commuters is set as the average value (i.e., $15 per hour) of that of heterogeneous users; Values in bracket are for the managed lane (i.e., HOT lane or HOV lane) and the GP lanes, respectively; Values in bracket are for the SOV and HOV modes, respectively.
<table>
<thead>
<tr>
<th>Scenarios</th>
<th>Financial feasibility</th>
<th>Economical feasibility</th>
<th>Congestion relief</th>
<th>Financial feasibility</th>
<th>Economical feasibility</th>
<th>Congestion relief</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOV_EL</td>
<td>=</td>
<td>+</td>
<td>-</td>
<td>=</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>HOT_EL</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+(−)&quot;</td>
<td>+(=)&quot;</td>
<td>(=)&quot;</td>
</tr>
<tr>
<td>HOV_NL</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>HOT_NL</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Note: the symbols, "="; "+"; and "−", indicate equivalent, better, and worse conditions as compared to the do-nothing case; "" symbols in brackets represent the feasibilities of the HOT_EL_CCM.

Sensitivity analysis

The above results are dictated by the initial settings (e.g., the commuters’ VOT and travel demand), which may vary temporally and spatially under specific applications. A sensitivity analysis is conducted on the HOT_EL scenario to explore the impacts of commuters’ heterogeneity and the travel demand on the optimal toll.

The heterogeneity of commuters is quantitatively interpreted as the VOT difference between the three groups of commuters. Figure 4 shows the optimal toll changes of HOT_EL under four objective criteria. As the VOT difference increases, the optimal tolls rise in a linear manner for the TRM, SWM, and LOS criteria, while the curve of CCM performs in the shape of an inverse “U”. It is comprehensible that on a highway with larger VOT difference among commuters, higher toll is reasonable for obtaining more toll revenue and social welfare and maintaining the SOV traffic flow composed of high-VOT commuters. For the U-shape curve of CCM, it may be explained by the congestion cost computation (i.e., equations (3)), which is the product of commuters’ VOT and the lower-level traffic flow solutions to equations (7a-c). It is implied that the effectiveness of the pricing strategy on congestion mitigation of HOT-lane highway may not be equivalent for different compositions of heterogeneous commuters.

Figure 5 illustrates the impacts of the travel demand on the optimal toll. It is shown that higher level of travel demand lifts the toll levels for all objective criteria, but the optimal tolls vary in different shapes of curves. The increments of the optimal tolls on the SWM, CCM, and LOS curves appear to be much sharper than the TRM curve, which shows a diminishing sensitivity in response to the varying travel demand. The above results reveal the divergent impacts of the commuters’ heterogeneity and travel demand on optimal tolls under different objective criteria and highlight the importance of selecting different policy goals for the HOT projects.
FIGURE 4 Sensitivity analysis on the VOT difference.

FIGURE 5 Sensitivity analysis on the travel demand.

CONCLUSIONS
The paper proposes a bilevel modeling framework to simulate the interactions of the
HOT lane operators and the roadway users. Four objective functions are formulated to reflect the pricing strategies of HOT lane operators under different policy goals. Travelers’ travel choices are simulated and tested in three forms of models, i.e., the multinomial logit model, the nested logit model, and the Logit & UE model. The heterogeneity of commuters is explicitly addressed and represented by the discrete categories of their VOTs. In the numerical example, the paper compares the three forms of travel choice models and conducts scenario analysis to evaluate various HOT operations. Results demonstrate that: (i) the MNL and NL models may result in an unrealistic high toll under CCM criterion due to the overestimated impacts of the tolls on the mode split and elastic demand; (ii) the homogeneity assumption may lead to biased results of the effectiveness of HOT lanes; (iii) different policy goals have significant influences on the HOT lane operations; and (iv) although the HOT concept shows a good economic efficiency from the social perspective, it is not financially sustainable to add a new lane for the HOT project due to the considerable costs from obtaining a new right-of-way (ROW).

It is pointed out that the research has two main limitations: (i) the commuters’ mode- and lane-choice models are developed in their generic formats without specific structure design and the comparison results should not apply to the models that are calibrated with the empirical data and (ii) although settings of the numerical studies are based on the real data or justified researches, the results cannot directly be applied to guide any specific HOT projects. In future research, the study can be extended to examine the time-varying/dynamic tolls during different times of days with various travel demand.

Acknowledgements

The research is supported in part by the National Natural Science Foundation of China (No.71271176) and the youth science and technology foundation of Sichuan Province.

REFERENCES


26. Wichiensin, M., Bell, M. G. H., and Yang, H. Impact Of Congestion Charging


