Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 9 (2015) 90 – 105

Transportation Research Procedia

www.elsevier.com/locate/procedia

21st International Symposium on Transportation and Traffic Theory

Incorporating departure time choice into high-occupancy/toll (HOT) algorithm evaluation

Stephen D. Boyles [a;*], Lauren M. Gardner [b], Hillel Bar-Gera [c]

[a]The University of Texas at Austin, 301 E. Dean Keeton St. Stop C1761, Austin, TX, 78712, USA
[b]University of New South Wales, Sydney, NSW, 2052, Australia
[c]Ben-Gurion University of the Negev, Beer-Sheva, Israel

Abstract

This paper presents an evaluation framework for high-occupancy/toll (HOT) lanes, extending previous frameworks by incorporating departure time choice alongside stochasticity in demand. Drivers are divided into two classes, strategic and captive drivers. Strategic drivers time their departures to minimize expected generalized cost and can choose either the free lanes or HOT lane. Captive drivers have random departure times, and must choose the free lanes. Generalized cost for strategic drivers takes the form of the well-known schedule delay penalty function combined with travel time. The traffic flow model involves two parallel bottlenecks, one for each lane group. Four toll algorithms are compared: a fixed toll which is constant in time and unresponsive; a dynamic, but unresponsive toll based on expected demand; a dynamic and responsive toll which adapts based on congestion in the HOT lane; and a "full-information" toll which is aware of all captive driver inflows.

Peer-review under responsibility of the Scientific Committee of ISTTT21

Keywords: HOT lanes; departure time choice; single bottleneck pricing

1. Introduction

High-occupancy/toll (HOT) lanes attempt to manage congestion by permitting low-occupancy vehicles to use a designated lane or set of lanes by paying a toll, while high-occupancy vehicles can use these lanes for free. HOT lanes are parallel to a toll-free set of general purpose (GP) lanes. Typically, this toll is dynamic, responding to the congestion level, time of day, and other factors. Designing dynamic pricing algorithms is nontrivial. Several objectives can plausibly be considered. More importantly, there are many factors influencing the choices of drivers and the performance of the facility. Understanding and addressing the various factors is a critical challenge, not yet fully addressed in the literature.

In this paper we use an equilibrium evaluation framework, as it offers tractability and thus insights into the underlying processes. In the current framework, two types of choices are modeled. First, each driver must choose a departure time; second, upon arriving at the split between the GP and HOT lanes, drivers must choose which lane to take. The departure time decision is habitual, and this paper assumes that drivers departure times are the same each day and do

* Corresponding author. Tel.: +1-512-471-3548; fax: +1-512-475-8744.
E-mail address: sboyles@mail.utexas.edu

Peer-review under responsibility of the Scientific Committee of ISTTT21
doi:10.1016/j.trpro.2015.07.006

not depend on specific travel conditions on that day. We assume that drivers are provided information on the current toll and travel times on GP and HOT lanes right before making the lane choice decision. This decision is "real-time" rather than habitual, and can vary from one day to the next based on current travel conditions, which in our model are stochastic.

This paper follows most directly from Gardner et al. (2013) and Gardner et al. (2015), which are discussed here; relationships with other HOT lane literature are described in the literature review. Gardner et al. (2013) focused on alternative choice models for HOT lane choice, and compared three simple algorithms. Under the relatively simple assumptions in that paper, the principle of "full utilization" works well, in which the number of vehicles entering the HOT lane is maximized subject to a capacity constraint. When there is uncertainty in demand, however, it is not generally possible to achieve full utilization at all times, but algorithms based on this principle still seem to perform well. Gardner et al. (2015) made such a comparison while introducing stochasticity in arrival rates. Both papers demonstrate the advantages of equilibrium frameworks.

However, both of these papers assume that the arrival time profile of vehicles is fixed and independent of the toll algorithm. In reality, drivers may adjust their departure times based on the toll algorithm and HOT lane performance, particularly if it offers a more reliable travel alternative. The objective of this paper is to incorporate departure time choice effects into the HOT evaluation framework of Gardner et al. (2013) and Gardner et al. (2015).

We use a fairly standard departure time choice model, in which the generalized cost is a linear combination of travel time, early arrival penalty and late arrival penalty. However, equilibrium models of departure time choice usually do not include stochasticity, which is essential for the evaluation of HOT facilities. Therefore, the key challenge is how to integrate departure time choice and stochasticity. This can be done in many different ways, reflecting additional complexities of the reality. Our goal is to find a relatively simple integrated scenario, leading to a tractable model, which can be useful for the understanding of the primary effects of departure time choice on HOT facility performance. The novelty of this paper stems from the proposed scenario/model.

In reality, typically, all drivers choose their departure time, all drivers can choose whether to use the HOT lane or not, and the moment of arrival to the HOT lane entrance has a stochastic component for all drivers. Representing this complete mixture in an equilibrium framework can be very challenging, and to the best of our understanding does not lead to a tractable model. Therefore, we decided to separate drivers into two classes: strategic and captive. This framework is used to compare the performance of several toll algorithms, and evaluate HOT alternatives against an alternative where all lanes are free. Strategic drivers are further distinguished by their heterogeneous value of time (VOT). They choose their departure times based on the schedule delay concept. They also choose whether to use the HOT lane or not. Their travel time from departure to the HOT lane entrance is assumed constant. Captive drivers always use the GP lane. Their departure rate has a pre-specified time-dependent expected value, while actual departures can be stochastic or deterministic. Their role is to introduce day-to-day stochasticity of GP queue length, and thus travel times. The proposed separation lead to a tractable equilibrium model, described in detail in section 3, that incorporates departure time choice, HOT lane choice, and stochasticity. It can therefore serve as a useful tool for studying the effects of departure time choices in the context of HOT facilities.

In this paper we used the proposed framework to compare the performance of several toll algorithms, and evaluate HOT alternatives against an alternative where all lanes are free. In particular, four toll algorithms are compared. Two are relatively extreme cases which aim to serve as performance bounds: a time-invariant, unresponsive toll (which may nevertheless be set at an arbitrary value), and a "perfect-information" toll which can achieve full utilization by having complete information on departure rates and the stochastic realization. The other two are considered more realistic: an unresponsive, but time-varying toll which would achieve full utilization at mean demand rates, and a responsive toll which adjusts the latter based on current occupancy in the HOT lane. These are demonstrated on a test facility similar to that in Gardner et al. (2013) and Gardner et al. (2015) to facilitate comparison and identify the impact of departure time choice modeling.

The remainder of the paper is organized as follows. Section 2 discusses past literature on HOT lanes, particularly toll algorithms and evaluation methods, and also reviews departure time choice models. Section 3 presents the model components, including the facility and traffic flow, strategic and captive drivers, toll algorithms, and the equilibrium concept used. Section 4 discusses the heuristic devised for the equilibrium model and discusses other implementation details. Section 5 presents the case study and results, and Section 6 concludes and identifies future research directions.

2. Literature Review

HOT facilities have been constructed in countries around the world, and show promise in presenting a reliable alternative to congested freeways. Since their introduction in 1995 on SR-91 in California, HOT lanes have been implemented in Atlanta, Denver, Houston, Los Angeles, Miami, Minneapolis, Salt Lake City, San Diego, and Seattle in the United States, and in Tel Aviv, Israel. Several different toll algorithms are used to determine the price, and a division can be drawn between facilities where tolls are dynamic, but pre-scheduled, and facilities where tolls are responsive to real-time conditions (Chung and Recker, 2011). Responsive tolls can be set in several ways. In Minneapolis, the toll is set based on a lookup table given density in the HOT lane (Halvorson et al., 2006). Researchers have also devised optimization-based toll algorithms, such as Michalaka et al. (2011) who propose a scenario-based robust optimization model. Reactive self-learning approaches have been proposed by Yin and Lou (2009) and Lou et al. (2011) based on observed traffic arrivals and drivers' willingness to pay. For particular network configurations, the optimal tolls can be found analytically from the optimization problem (Chu, 1995; Yang and Huang, 1997; Arnott et al., 1998; Liu and McDonald, 1999; Kuwahara, 2001)

The objective of "full utilization" is an intuitive criterion for determining HOT tolls, and was described at least as early as Dahlgren (2002). Under particular assumptions on the traffic flow model — namely, a single-bottleneck model with homogeneous vehicles and driving behavior — Gardner et al. (2013) show that this principle does indeed minimize average passenger travel time. These assumptions are nontrivial, though; for instance, Lou et al. (2011) describe how moving bottlenecks may prevent full utilization from occurring in practice. Furthermore, the toll-setting algorithm is unlikely to have all of the information needed to ensure full utilization in field conditions, since the value of time (VOT) distribution of arriving drivers cannot be directly observed. Nevertheless, even under stochasticity in arrival rates and in ignorance of the specific VOTs of arriving vehicles, algorithms based on the full utilization concept still perform well, as long as the population VOT distribution is known (Gardner et al., 2015).

Researchers have proposed several evaluation schemes for comparing the performance of different algorithms. In terms of the traffic flow model, the simplest assumption is the well-known single-bottleneck queuing model of Vickrey (1969). Microscopic simulation (e.g., Morgul and Ozbay, 2011) and mesoscopic simulation (e.g., Ozbay et al., 2011) have also been used, and provide more realism at the expense of analytical tractability. A number of case studies have also been developed based on field data in locations where HOT facilities have been implemented (Parkany, 1998; Gordon et al., 2004; Burris et al., 2009; Chung and Choi, 2010; Burris et al., 2011; Cao et al., 2011; Fosgerau, 2011; Kuhn et al., 2011; Bar-Gera, 2012). In this paper, the single-bottleneck model is used for traffic flow due to its simplicity, in order to focus attention on the impacts of departure time choice.

While Gardner et al. (2015) specifically considered the impact of stochastic arrival rates of HOT lane operation, there is a large body of literature on more general toll-setting problems under uncertainty in demand or driver choice behavior (Lam and Tam, 1998; Yang, 1999a; Gardner et al., 2008), in roadway capacity (Boyles et al., 2010), or both (Yang, 1999b; Gardner et al., 2011). In such problems, the information available to drivers and to the toll setter plays a major role in determining the potential effectiveness of tolling strategies. As described in Section 3, this paper assumes that drivers learn the HOT lane toll, and travel times on both lanes, only at the moment they reach the diverge point. The four toll algorithms tested in this paper correspond to different assumptions on the toll setter's knowledge (or, under an alternate interpretation, to the flexibility the toll-setter has in implementing a responsive toll).

Vickrey's bottleneck has formed the basis for much research into departure choice modeling. Drivers are assumed to have a preferred arrival time at their destination, and earlier or later arrival is penalized with an appropriate function. The resulting equilibrium in departure time choices in different contexts has been studied extensively (Hendrickson and Kocur, 1981; Smith, 1984; Daganzo, 1985; Arnott et al., 1993; Liu and Nie, 2010; Gonzales and Daganzo, 2012). In particular, this model has been used to analyze parallel free and toll roads (Arnott et al., 1990; de Palma and Lindsey, 2000). Target arrival times with penalties have been applied in other dynamic traffic assignment problems as well, often involving a larger network (Li et al., 1999; Ziliaskopoulos and Rao, 1999; Friesz et al., 2001; Bellei et al., 2006; Bliemer and van Amelsfort, 2010). We use here the concept of target arrival times with penalties, modeling the HOT and GP lanes as parallel bottlenecks, incorporating stochasticity in arrivals of one class of drivers while allowing the other to choose departure times to minimize a generalized cost function.

The model presented in this paper builds on this literature by combining departure time choice and stochasticity in demand, using Vickrey's bottleneck model to compare four alternative toll algorithms. While each of these models has been extensively used in prior literature, we believe this particular combination is novel, and useful for comparing the performance of these alternative algorithms. In particular, the combination of departure time choice and stochastic demand appears to be highly relevant for understanding the effectiveness of real-time pricing algorithms, but not previously studied in the literature.

3. Model

This section explains each component of the model developed for this paper, along with the supporting assumptions. Section 3.1 introduces the facility geometry under consideration, and describes the two classes of drivers, strategic and captive. Section 3.2 describes the traffic flow model and the HOT facility in greater detail. Section 3.3 presents the four toll algorithms evaluated. Finally, Section 3.4 presents the equilibrium framework linking all of these components. The model is discrete in time ($t \in \left\{0, 1, \cdots, \bar{T}\right\}$) and has discrete driver classes, but the flow variables themselves are continuous.

3.1. Strategic and captive drivers

Our model has two driver classes: *strategic* and *captive* drivers. Strategic drivers are willing to change their departure time in order to minimize their generalized cost, which is based on the schedule delay concept. Each strategic driver has a preferred arrival time t^*, and their generalized cost depends on the arrival time t_A at the destination, and the departure time t_D from the origin:

$$g(t_A, t_D | t^*, \alpha, \beta, \gamma) = \alpha(t_A - t_D) + \beta[t^* - t_A]^+ + \gamma[t^A - t^*]^+ \tag{1}$$

where α, β, and γ are the value of travel time and penalty factors for early and late arrival, and $[x]^+ = \max\{x, 0\}$. We assume K classes of drivers, and that all drivers within each class have identical values of these parameters. While one may imagine these parameters are continuously distributed across the population, the solution methods presented below require a discretization, so we choose to present the model in this way. The notation $g_k(t_A, t_D)$ is used to represent the generalized cost function parameterized by the t^*, α, β, and γ values for class k. The number of drivers in class k is N_k, and the number of such drivers who depart at time t is $d^k(t)$. This latter value is called the *demand profile* for class k, and is endogenous to the model. A demand profile is feasible if its elements are nonnegative and $\sum_{t=0}^{\bar{T}} d_k(t) = N_k$. The set of feasible profiles for class k is given by D_k.

By contrast, captive drivers do not choose their departure time, but instead enter the facility according to a random process. In addition, we assume that captive drivers never choose the HOT lane. There are several ways to interpret this distinction. Captive drivers may represent drivers with negligible value of time (perhaps because of trip purpose) who do not attempt to time their departure. Alternately, captive drivers may come from routes that merge onto the mainline downstream of the HOT diverge point, so they have no option to choose the HOT lane; the stochasticity in arrival time arises from uncertainty in travel times between their origins and the merge point (Figure 1). Strict separation between strategic and captive drivers, as depicted by our proposed evaluation framework, is not necessarily reflective of reality. Yet it leads to a tractable model that captures certain key phenomena as we demonstrate in this paper, providing insights about HOT lane tolling algorithms. Let the random variable $\xi(t)$ denote the number of captive drivers entering the GP lane at time t; these distributions are exogenous and independent of strategic drivers' choices. Furthermore $\xi(t_1)$ and $\xi(t_2)$ are assumed independent whenever $t_1 \neq t_2$, but need not be identically distributed.

Strategic drivers, upon reaching the diverge point, learn travel times on the GP and HOT lanes (τ_{GP} and τ_{HOT}) and the current toll c. Based on the current time and their behavioral parameters (t^*, α, β, and γ) the lane offering a lower generalized cost is selected. This generalized cost is not known before departing, because of the presence of captive drivers. In the language of stochastic optimization, the departure time choice is a first-stage variable (made in ignorance of the actual travel times and toll), while the choice of GP or HOT lane is a recourse variable (made with knowledge of the travel time and toll realization). If the lanes are tied in generalized cost for a particular class of strategic drivers, they split between both lane groups proportionate to their capacity Q. That is, if $p^k(t)$ is the proportion of class k drivers entering the HOT lane at time t, and g_{HOT} and g_{GP} give the generalized costs for each lane group based on equation (1), then $p^k(t) = 1$ if $g_{HOT} < g_{GP}$, $p^k(t) = 0$ if $g_{HOT} > g_{GP}$, and $p^k(t) = Q_{HOT}/(Q_{HOT} + Q_{GP})$.

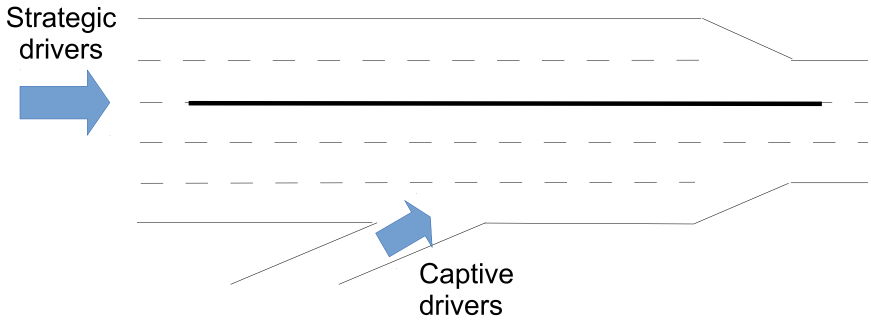


Fig. 1: One interpretation of strategic and captive drivers.

3.2. Facility and traffic flow

As indicated in Figure 1, the HOT and GP lanes both have a downstream bottleneck with capacities Q_{HOT} and Q_{GP} and free-flow times τ^0_{HOT} and τ^0_{GP} (the free flow times are assumed to be integers when measured in the same unit as the time discretization). Inflow rates to the lanes are not restricted, but bottlenecks exist at the downstream ends of the facilities. This implies that upstream of the bottleneck, the capacity is large enough to accommodate all strategic and captive drivers wishing to enter. Under the assumption that inflows are never large enough for a queue to occupy the entire length of the facility, a point queue model is a reasonable and simple way to model delay.

The facility operations can then be modeled as a time-inhomogeneous Markov chain with $\tau^0_{HOT} + \tau^0_{GP}$ state variables $n^1_{HOT}, n^2_{HOT}, \cdots n^{\tau^0_{HOT}}_{HOT}, n^1_{GP}, n^2_{GP}, \cdots n^{\tau^0_{GP}}_{GP}$, representing recent inflows. Specifically, the initial condition is $n^i_{HOT}(0) = n^j_{GP} = 0$ for all $i \in \{1, \cdots, \tau^0_{HOT}\}$ and $j \in \{1, \cdots, \tau^0_{GP}\}$, and the transition equations are $n^i(t) = n^{i-1}(t-1)$ for $i = 2, \cdots, \tau^0 - 1$, and $n^i(t) = n^{i-1}(t-1) + [n^i(t-1) - Q]^+$ for $i = \tau^0$ (these equations hold for both HOT and GP lanes, subscripts omitted for brevity). For $i = 1$, $n^1_{HOT} = \sum_k d^k(t-1)p^k(t-1)$ and $n^1_{GP} = \xi(t-1) + \sum_k d^k(t-1)(1 - p^k(t-1))$. This can be thought of as a "cell-based" representation of a point queue model, cf. Nie et al. (2008).

The values of these state variables at any time t are sufficient to reconstruct travel times τ_{HOT} and τ_{GP} at that instant, by performing a "forward simulation" of the state variables given their current values and identifying the time at which an infinitesimal vehicle entering the upstream end would leave the downstream end of the facility. Specifically, define the recursion $v(0) = n^{\tau^0}(t)$ and $v(k) = [v(k-1) - Q]^+ + n^{\tau^0 - k}(t)$ with the convention that $n_i(t) = 0$ whenever $i < 1$. Let $\bar{\tau}$ be the smallest integer such that $v(\bar{\tau}) = 0$ and $\bar{\tau} \geq \tau^0$. Then the travel time is

$$\tau(t) = \max\left\{\tau^0, \bar{\tau} - \frac{Q - v(\bar{\tau} - 1)}{Q}\right\} \tag{2}$$

where the subtraction in the second term accounts for queue clearance in the middle of a discrete time interval. The introduction of this latter factor makes τ continuous in the state variables, which is important for showing equilibrium existence in Section 3.4.

An example of how the travel times are calculated can be found in Figure 2. In the left panel, the current **n** values are $[n_1, n_2, n_3] = [5, 8, 18]$. Then $v(0) = 18$, $v(1) = [18 - 10]^+ + 8 = 16$, $v(2) = [16 - 10]^+ + 5 = 11$, $v(3) = 1$, and $v(4) = 0$ so $\bar{\tau} = 4$. The travel time is then calculated as in equation (2), using the second term in the the maximum, as $4 - (10 - 1)/10 = 3.1$. The fraction is obtained based on the number of vehicles in the downstream before clearance (1) and the capacity (10) which assumes that there is a queue discharging when the last vehicle exits. In the right panel, the current **n** values are $[n_1, n_2, n_3] = [5, 8, 2]$. Then $v(0) = 2$, $v(1) = [2 - 10]^+ + 8 = 8$, $v(2) = 5$, and $v(3) = 0$, so $\bar{\tau} = 3$. So, in equation (2), the first term in the maximum applies and the travel time is 3, corresponding to the "free-flow" case when there is no queue at the downstream end.

Time (future)	n_1	n_2	n_3
0	5	8	18
1	0	5	16
2	0	0	11
3	0	0	1
4	0	0	0

(a)

Time (future)	n_1	n_2	n_3
0	5	8	2
1	0	5	8
2	0	0	5
3	0	0	0

(b)

Fig. 2: Examples of travel time calculations by showing hypothetical evolution of cell occupancies into future time intervals (ignoring new arrivals).

3.3. Toll algorithms

This paper compares four toll algorithms: (1) a fixed toll (constant across time), (2) pre-scheduled tolls based on the full utilization principle, using the demand profiles for strategic drivers and expected inflows for captive drivers, without adjustments to realized values (FU-MEAN); (3) real-time density-modified full-utilization tolls, where the toll in (2) is modified according to the density in the HOT lane to move closer to full utilization (FU-DM); and (4) perfect information full-utilization tolls, which are set with full knowledge of the captive drivers' departure time profile (FU-PI). Algorithms 1 and 4 are seen more as bounds on performance than algorithms likely to be applied; unlike algorithm 1, HOT lanes are generally dynamically priced, and the information required by algorithm 4 is not generally available in advance of toll setting. This section explains each of these algorithms, and how they may be calculated.

The fixed-toll Algorithm 1 is self-evident; the values $c(t)$ are constant in time and specified *a priori* as an input for testing; an "optimal" level for the fixed toll can be determined experimentally. Algorithm 2 (FU-MEAN) can also be specified *a priori*, but involves some calculation. FU-MEAN tolls vary with time, and are set as if the inflow rates of captive drivers were deterministically equal to their mean values. To achieve full utilization at time t, the toll $c(t)$ must be set such that

$$\sum_k d^k(t-1)p^k(t-1) = Q_{HOT} \tag{3}$$

assuming $\sum_k d^k(t) > Q_{HOT}$; otherwise there is no harm in setting $c = 0$ under the deterministic assumption used to derive FU-MEAN. This toll can thus be calculated inductively: at $t = 0$, $\tau_{HOT} = \tau^0_{HOT}$ and $\tau_{GP} = \tau^0_{GP}$ and (3) can be solved for $c(0)$. For $t = 1$, travel times are updated based on inflows in time interval 0, calculated from the transition equations in the previous subsection with $\xi(1)$ replaced by $E[\xi(1)]$, and equation (3) is again solved to find $c(1)$, and so forth. The FU-MEAN tolls $c_{FU-MEAN}(t)$ can thus be calculated ahead of time since it only relies on expected values for captive drivers. It is well-known in stochastic programming that optimizing based on mean values of parameters is suboptimal, so we allow the FU-MEAN toll to be multiplied by a fixed constant, chosen *a priori*.

Algorithm 3 (FU-DM) is adaptive based on observed flows into the HOT lane, and thus is allowed to depend on the current values of $n^1_{HOT}, n^2_{HOT}, \cdots n^{\tau^0_{HOT}}_{HOT}$. Define the total HOT lane occupancy $O_{HOT}(t) = \sum_{i=1}^{\tau^0_{HOT}} n^i_{HOT}(t)$ to be the number of vehicles currently on the HOT facility at time t. If full utilization has been maintained at all time intervals in the past, then $O_{HOT}(t)$ equals the "target occupancy" $O^*_{HOT}(t) = Q_{HOT} \min\{t, \tau^0_{HOT}\}$. If this is not the case, FU-DM modifies the FU-MEAN toll accordingly, increasing the toll if too many vehicles have entered the lane:

$$c_{FU-DM}(t) = c_{FU-MEAN}(t) + \phi\,[O_{HOT}(t) - O^*_{HOT}(t)]^+ \tag{4}$$

where ϕ is a scaling factor which can be calibrated experimentally.

Algorithm 4 (FU-PI) is also adaptive, and makes use of the realized value of the current captive driver inflow rate $\xi(t)$, as opposed to FU-MEAN which uses only its mean. That is, FU-PI is calculated using (3), but with $\xi(t)$ used in the calculation of p^k. FU-PI is thus able to achieve full utilization at all times when it is possible to do so. FU-MEAN attempts to approximate FU-PI by replacing the realized values of $\xi(t)$ with their means, and FU-DM attempts to adjust FU-MEAN based on current occupancy (without knowledge of the actual captive driver inflow rate, which may be difficult to measure in real time).

Furthermore, given a demand profile $d^k(t)$ and toll algorithm, we must also calculate the joint distribution of $\tau_{HOT}(t)$, $\tau_{GP}(t)$, and $c(t)$ at all points in time. The need for this will become apparent in the next subsection. In practice, this is most easily accomplished by sampling different $\xi(t)$ and repeated simulation. However, in theory this joint distribution can be calculated explicitly, by casting the evolution of the system as a Markov chain with acyclic transition graph. The states of this system are vectors $(t, \xi, n^1_{HOT}, n^2_{HOT}, \cdots n^{\tau^0_{HOT}}_{HOT}, n^1_{GP}, n^2_{GP}, \cdots n^{\tau^0_{GP}}_{GP})$. Let S denote such a state, and $\mathcal{S}$ the set of all states. As described above, this state description provides sufficient information for calculating travel times and tolls for all algorithms. The notation $\tau_{HOT}(S)$, $\tau_{GP}(S)$, and $c(S)$ is used to reflect the travel times and tolls corresponding to state S. The system starts in state $(0, \xi, \mathbf{0}, \mathbf{0})$ with probability $P(\xi(0) = \xi)$. The transition probability from state $S = (t, \xi, \mathbf{n_{HOT}}, \mathbf{n_{GP}})$ to $S' = (t', \xi', \mathbf{n'_{HOT}}, \mathbf{n'_{GP}})$ is zero unless $t' = t + 1$, and $\mathbf{n'_{HOT}}$ and $\mathbf{n'_{GP}}$ are calculated as in Section 3.2 given $\tau_{HOT}(S)$, $\tau_{GP}(S)$, and $c(S)$. Otherwise, the transition probability is given by $P(\xi(t') = \xi')$. Given these initial and transition probabilities, the probability of observing any joint values of $\tau_{HOT}(t)$, $\tau_{GP}(t)$, and $c(t)$ at all points in time can in principle be calculated.

3.4. Equilibrium

With the model components described above, the joint distribution of travel times and tolls depends on the strategic drivers' departure profile. However, strategic drivers will also adjust their departure profile based on the joint distribution of travel times and tolls. An equilibrium framework is introduced to resolve this mutual dependence. At an equilibrium solution, strategic drivers have no incentive to change departure times given the distribution of travel times and tolls induced by their current profiles. Given the joint distribution of $\tau_{HOT}(t)$, $\tau_{GP}(t)$, and $c(t)$ for all time intervals, one can also calculate the joint distribution of the generalized costs drivers from each class will experience from departing at a particular time: let $g^*_k(t) = E[\min\{g_k(t + \tau_{GP}(t), t), c(t) + g_k(t + \tau_{HOT}(t), t)\}]$ represent this value. Note the order of expectation and minimization: drivers learn the actual travel times and tolls upon reaching the diverge point and can pick the lane group with lower disutility.

At equilibrium, class k drivers depart during a time interval only if $g^*_k(t)$ is minimal over all t; that is, $d^k(t) > 0$ implies $g^*_k(t) = \min_{t'}\{g^*_k(t')\}$. Note that $g^*_k(t)$ implicitly depends on the demand profiles $\mathbf{d}$ by affecting the joint distribution of $\tau_{HOT}(t)$, $\tau_{GP}(t)$, and $c(t)$, and this dependence can be made explicit by writing $g^*_k(t, \mathbf{d})$. In this way, an equilibrium solution $\mathbf{d^*_k(t)}$ can be characterized by the variational inequalities

$$\mathbf{g^*_k}(\mathbf{d^*_k(t)}) \cdot (\mathbf{d^*_k(t)} - \mathbf{d_k}) \le 0 \tag{5}$$

for all k and $\mathbf{d_k} \in D_k$, where the vectors are dimensioned by the time interval. Since the D_k are compact and convex sets, existence of an equilibrium solution follows from Corollary 2.2.5 of Facchinei and Pang (2003) if $\mathbf{g}^*(\mathbf{d})$ is a continuous function. This function is indeed continuous, because it is the composition of continuous functions, including the lane choice, toll algorithms (any of the four), flow propagation, and travel time functions. However, it is not clear whether this equilibrium is unique.

4. An Equilibrium Heuristic

Due to the complexity of the g^* mapping, an exact algorithm for solving the variational inequalities (5) is not apparent. Thus, we propose a heuristic for finding an approximate equilibrium solution, based on the method of successive averages (MSA). In this heuristic, one provides initial feasible demand profiles for each class, then iterates between (1) simulating these demand profiles repeatedly to obtain a sampling distribution for travel times and tolls; (2) identifying "target" departure times for each class based on the distributions obtained from simulation; and (3) shifting the demand profiles for each class toward the target times.

In the first step, each simulation run involves loading the given demand profiles onto the facilities, sampling to generate captive vehicle entries, calculating tolls based on the chosen algorithm and lane travel times based on equation (2), and splitting strategic drivers based on generalized costs in increasing order of time. This is used to generate a sampling distribution for $\tau_{HOT}(t)$, $\tau_{GP}(t)$, and $c(t)$. From an implementation standpoint, it is enough to calculate the sample mean of $g^*_k(t)$, which can be done with minimal memory overhead (cf. Knuth, 1997, section 4.2.2, equation 15).

In the second step, the target departure time for each class is selected as $t_k^* \in \arg\min_t \{g_k^*(t)\}$. In the third step, a convex combination is taken between the current demand profile, and a target profile in which all class k travelers depart at t_k^*. That is, $d_k(t)$ is replaced by $d_k(t) + \lambda(N_k - d_k(t))$ if $t = t_k^*$ and $(1-\lambda)d_k(t)$ otherwise, for some $\lambda \in [0, 1]$. In the implementation below, the standard method of successive averages is applied, where λ is inversely proportional to the iteration number.

Specifically, the following steps are performed:

1. Initialize departure profiles $d_k(t)$ for all strategic classes k. (A reasonable choice for initialization is timing departures to arrive at the target arrival time at free-flow.)
2. Repeat the following steps N times, where N is the number of samples for each simulation:
   (a) Initialize the time index $t \leftarrow 0$ and facility state variables: $\mathbf{n_{HOT}}, \mathbf{n_{GP}} \leftarrow \mathbf{0}$.
   (b) Calculate the number of captive drivers $\xi(t)$, and the travel times for each lane group τ_{HOT} and τ_{GP}.
   (c) Calculate the toll c as a function of $\mathbf{n_{GP}}$, $\mathbf{n_{HOT}}$, t, and the selected toll algorithm.
   (d) For each strategic class k, determine the proportion $p^k(t)$ entering the HOT lane as in Section 3.1.
   (e) Load vehicles and propagate flow according to the transition equations in Section 3.2.
   (f) Increment time index $t \leftarrow t + 1$ and return to step 2b until the time horizon is exceeded.
3. For each class k, calculate sample mean of expected generalized cost for each departure time t based on sampling distribution: $g_k^*(t) = E[\min\{g_{HOT}^k(t), g_{GP}^k(t)\}]$
4. Identify "all-or-nothing" profiles for each class: $d_k^*(t_k^*) = N_k$ for some $t_k^* \in \arg\min g_k^*(t)$ and 0 for all other times.
5. Calculate gap by comparing expected generalized costs for current and target profiles; terminate if sufficiently small.
6. Perform MSA updating: $\mathbf{d} \leftarrow \lambda \mathbf{d}^* + (1-\lambda)\mathbf{d}$
7. Return to step 2.

5. Case Study

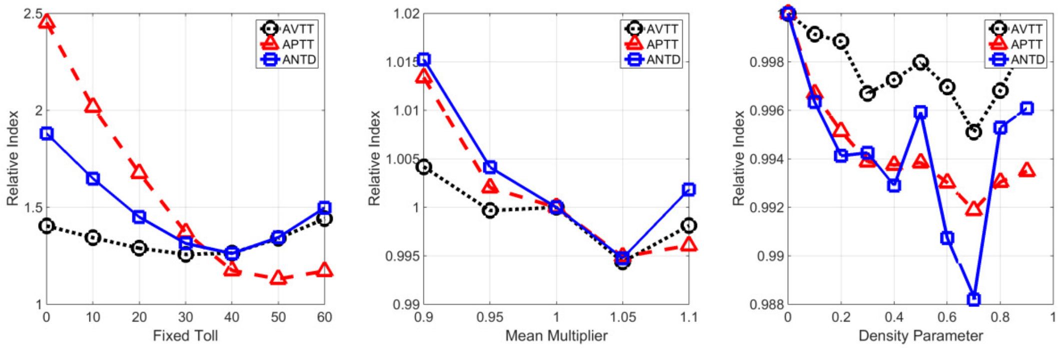
The results presented in this section are based on a facility of the type mentioned above. The HOT lane has a capacity of 1800 vph and the two GP lanes have a total capacity of 4200 vph. The length is 10 km and the free flow speed is 100 km/h. The expected demand profiles for the captive and strategic driver classes are provided in Table 1. The captive demand is assumed to follow a normal distribution with a standard deviation equal to 40% of the mean captive demand (in the base case). The expected demand profile is chosen such that the vehicle demand exceeds the GP lane capacity and a queue is formed at the bottleneck. Comparisons are also made to the all GP case, where all lanes are available to all drivers, to serve as a do-nothing comparison for alternatives analysis.

For the numerical analysis the time step is set to one minute, and the demand profiles are uniformly distributed within each hour. Both the captive group demand departures and first target arrival time for the strategic demand begin at t=60. The target arrival profile is set such that the first strategic drivers depart after t=0, thus removing any boundary effects. For the strategic demand profile there are a total of 2160 classes, with 12 classes per target arrival minute, with one class assigned to high-occupancy vehicles (HOVs), one class assigned to transit vehicles, and the remaining ten assigned to single-occupant vehicles (SOVs). For strategic SOV vehicles the VOT follows the Burr distribution with parameter $\gamma = 2$, and the median VOT is assumed to be \$15 per hour. The all-or-nothing lane choice model is implemented. Finally, the early and late arrival penalties are set such that the ratios α/β and α/γ are 0.5 and 1, respectively, following Lam and Small (2001). In this case users prefer early arrival to travel time, and equally weight travel time and late arrival.

Four metrics are considered, average vehicle travel time (AVTT), average person travel time (APTT), average non-transfer disutility (ANTD), and revenue. The metric AVTT is calculated by dividing the total travel time by the total number of strategic vehicles, where the metric APTT is calculated by dividing the total travel time by the total number of persons in strategic vehicles. ANTD is defined as the total generalized cost (including travel time and arrival penalties) excluding tolls (which are assumed to be transfer payments). Captive drivers are excluded from all of these metrics because their decisions are independent of any potential toll control. Both the mean and standard deviation

Table 1: Peak period travel demand profile per vehicle class

Group	Average occupancy	7:00–8:00	8:00–9:00	9:00–10:00	Total
Captive	1.2	3150	2550	1950	7650
Strategic	1.2	3150	2550	1950	7650
HOV	4.0	600	600	600	1800
Transit	40	300	300	300	900
Total		7200	6000	4800	18000



(a) Fixed toll (b) FU-MEAN (c) FU-DM

Fig. 3: Sensitivity analysis under base conditions for the three parameterized algorithms.

of each metric are calculated. Each metric is calculated for each sample, and $E[\cdot]$ and $\sigma[\cdot]$ are used to represent the sample means and standard deviations calculated across 50 samples.

5.1. Parameter calibration

Three of the tolling policies presented in this work are parameterized: fixed, FU-MEAN and FU-DM tolls. Therefore, prior to comparing the performance of each tolling policy, a sensitivity analysis is conducted for each parameterized tolling schemes in order to identify the optimal respective parameter values. In each case the optimal value will be chosen such that $E[ANTD]$ is minimized under the base case demand profile. For the fixed, FU-MEAN and FU-DM toll policies, sensitivity analysis is conducted regarding the toll value, mean multiplier and scaling factor, respectively. Figure 3 illustrates the performance in terms of the three metrics $E[AVTT]$, $E[APTT]$ and $E[ANTD]$, for a range of fixed toll values between \$0 and \$60, a mean toll multiplier ranging between 0.9 and 1.1, and a DM parameter ranging between 0 and 0.9. For the mean toll multiplier the vertical axis represents the ratio of performance under the chosen mean multiplier relative to a mean multiplier of 1.0, computed as the ratio of each expected metric to the value of the expected metric when the multiplier equals 1.0. (The reference values are $AVTT = 13.46$, $APTT = 7.83$, and $ANTD = 14.21$.) The sensitivity to the DM parameter is presented in a similar manner, where the vertical axis represents the ratio of performance for each DM parameter relative to using a DM parameter of 0, which is also equivalent to the FU-Mean toll (with a multiplier of 1.0).

From the figures it is evident that the optimal fixed toll is \$40, the optimal mean multiplier is 1.05 and the optimal DM parameter is 0.7. In contrast to the fixed toll value, the performance is highly robust to the parameters chosen for the DM and mean tolls. Relative to the reference case, the optimal mean multiplier reduces E(ANTD) by 0.5%, while the optimal density parameter reduces E(ANTD) only by 0.2%.

Table 2: Sample means of performance metrics under alternative toll schemes.

Metric	All GP	Fixed (40)	Mean (1.05)	DM (0.7)	PI
AVTT	18.9	17.0	13.4	13.4	13.0
APTT	19.2	9.20	7.79	7.77	7.72
ANTD	26.7	17.9	14.1	14.0	13.8
Revenue ($\times 10^3$)	0	184	146	144	132

Table 3: Sample standard deviations of performance metrics under alternative toll schemes.

Metric	All GP	Fixed (40)	Mean (1.05)	DM (0.7)	PI
AVTT	17.2	9.96	7.84	8.05	10.3
APTT	43.9	5.73	6.93	5.18	5.84
ANTD	26.2	6.75	6.22	5.76	7.46
Revenue	0	291	426	938	3412

5.2. *Facility performance under various tolling schemes*

In the analysis that follows the parameters for the fixed, FU-MEAN and FU-DM tolling policies are those identified as optimal from the previous section: \$40, 1.05, and 0.7, respectively. Tables 2 and 3 provide the expected value and standard deviation for each of the tolling policies.

The FU-MEAN and FU-DM tolls perform almost identically in terms of the expected value of each metric, while the FU-DM tolls results in slightly lower variability. As expected, the FU-PI tolls perform best with respect to minimizing $E[APTT]$ and $E[ANTD]$, however they only slightly outperform the MEAN-based tolls, which are significantly simpler to implement in practice. The fixed tolls perform the worst of the four tolling schemes, but still offer improvement in $E[APTT]$ and $E[ANTD]$ compared with scenario where all lanes are GP (a "do-nothing" alternative for HOT lane construction). For example, the fixed tolls result in a 32.8% decrease in E[ANTD], while the FU-DM tolls result in a 47.4% decrease. All four tolling polices result in significant decreases in travel time variability as well, relative to the All-GP scenario.

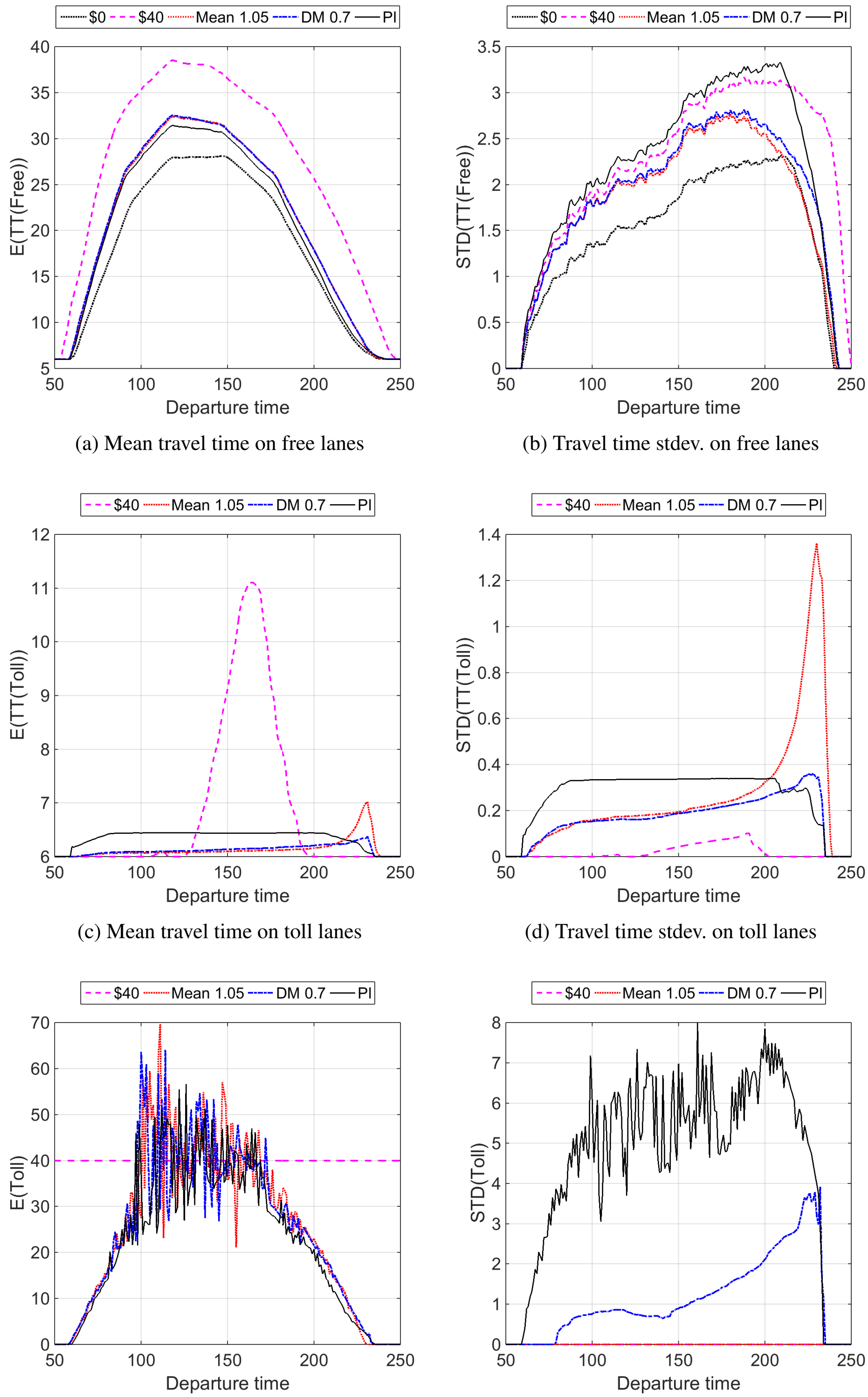
Figure 4 illustrates the change in the expected value and standard deviation of the toll and travel times on the GP and HOT lanes throughout the peak period for the four tolling schemes and the All GP case. The three FU tolling schemes are shown to mimic similar behavior over the course of the peak period, with a toll that increases sharply at the beginning of the peak period, remains high and highly variable during most of the peak period, and slowly decreases during the third and final hour (until the demand reduces to the point that the HOT lane does not need to be fully utilized to further decrease congestion). The variability of each toll is also illustrated. By design, the FU-MEAN toll has no variability, while the FU-DM toll varies increasingly during the second and third hour in the peak period, and the FU-PI toll is highly variable over nearly the entire peak period.

In terms of expected travel time, the All GP case represents the optimal facility performance. The highest GP travel times are under fixed tolls, while the three FU tolling schemes result in similar expected GP lane travel times. Similarly, the fixed tolls result in significantly higher HOT lane travel times during the middle of the peak period, which suggest that tolls which minimize ANTD are too low to completely prevent congestion on the HOT lane. In contrast, all three FU tolling schemes are illustrated to keep the HOT lane flowing near free flow conditions throughout the entire peak period.

The highest travel time variability on the GP lanes occurs under the FU-PI tolls (which are also the most variable), reaching 10% of the expected GP lane travel time. The higher variability under FU-PI tolls is because the HOT lane is continually kept at free flow capacity therefore the GP lane absorbs all of the variability resulting from the stochastic demand.

5.3. *Departure demand profile by VOT*

The focus of this work is the impact of departure time choice in combination with lane choice. Detailed analysis of the results provide useful insights, as illustrated here for the PI tolling scheme. Figure 5 shows the overall distribution



(a) Mean travel time on free lanes

(b) Travel time stdev. on free lanes

(c) Mean travel time on toll lanes

(d) Travel time stdev. on toll lanes

(e) Mean toll

(f) Toll standard deviation

Fig. 4: Performance of various tolling schemes. Horizontal axis is departure time, vertical axis is tolls in dollars or travel time in minutes.

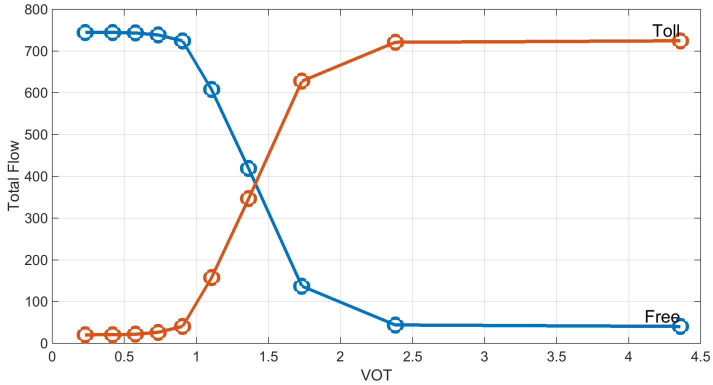


Fig. 5: Total flows into GP and toll lanes by VOT class.

between lane groups by VOT, aggregated over the entire modeling horizon. It can be seen that low VOT drivers (up to 0.9) use the GP lane almost always, while high VOT drivers (2.38 and 4.36) use the toll lane almost always. Therefore, the remainder of the analysis focuses on the three middle VOT values, 1.11, 1.36 and 1.73.

Figure 6 shows the proportion of GP users by departure time, for each of these three VOTs. It shows that during most late times, $t \in [100, 200]$, drivers with VOT of 1.73 primarily use the toll lane. The same is true for drivers with VOT of 1.36, but starting only at $t = 150$. On the other hand, during most early times, $t \in [60, 170]$ drivers with VOT 1.11 primarily use the GP lane. The same is true for drivers with VOT of 1.36, but only until $t = 100$. The periods when drivers split between the GP and toll lanes are $t \in [170, 220]$ for VOT 1.11; $t \in [100, 150]$ for VOT 1.36; and $t \in [60, 100]$ for VOT 1.73. So it seems as if at any given departure time only drivers with a specific VOT are actively choosing between the two alternative lanes.

To examine this conjecture, we show in Figure 7 the lead time drivers choose as a function of departure time for the same three VOT values. In the case of VOT 1.36, the lead time for $t \in [60, 90]$ increases rapidly, in a similar manner to the GP travel time. The model predicts that these travelers realize that they will probably use the GP lane, and therefore they plan their departure accordingly. On the other hand, for $t > 170$, the lead time of drivers with VOT 1.36 is about 6 minutes, which is the travel time on the toll lane. Again, this is in agreement with the observation that these drivers primarily use the toll lane, as shown in Figure 6. Similar connections can be seen for the other VOT values.

Figure 7 also shows (in dotted lines) the lead time needed by travelers choosing the toll lane (about 6 minutes during all departure times) and by travelers choosing the free lanes. Model lead time results for high VOT drivers (2.38 and 4.36) follow closely the toll lane curve, while lead time results for low VOT drivers (up to 0.9) follow closely the free lanes curve. (The curves of these other VOT drivers are not shown in the figure to maintain clarity.)

To analyze the behavior when drivers split between the lanes, it is useful to examine the mismatch between the expected arrival time by departure time and lane choice, and the target arrival time. This is shown in Figure 8. For example, for drivers with VOT 1.36, departing at time 75, if they choose the toll lane their expected arrival is approximately 12 minutes before their target arrival time, and if they choose the GP lane their expected arrival is approximately 12 minutes after their target arrival time. These drivers are likely to make choices in real time, in response to the actual travel time and toll prevailing on the specific day. The figure shows that for drivers with VOT 1.36 departing early ($t < 80$) the mismatch is nearly zero if they use the GP lanes, and for those departing late ($t > 170$) the mismatch is nearly zero if the use the toll lane. The same logical relationship can be observed for the other VOT values.

The different choices of lead time affect the departure profiles, as shown in Figure 9. For drivers with VOT 1.73 the departure profile follows the target arrival profile fairly closely. For drivers with VOT 1.11 the departure profile is higher than the target arrival profile during $t \in [60, 90]$ as a result of the increasing lead time, and the departure profile is lower than the target arrival profile during $t \in [150, 170]$ as a result of the decreasing lead time.

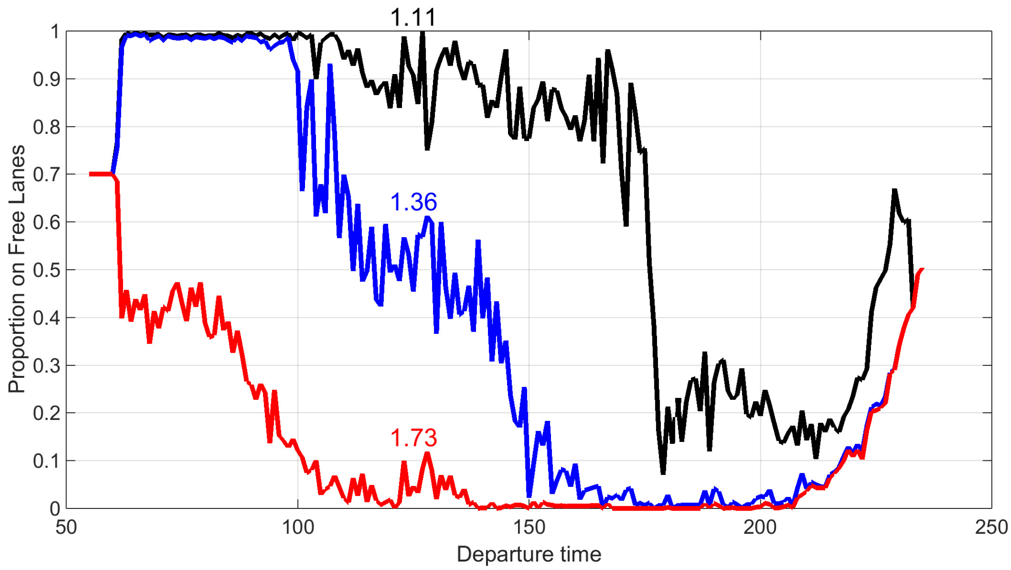


Fig. 6: Proportion of travelers selecting GP lane for selected VOTs with PI tolls.

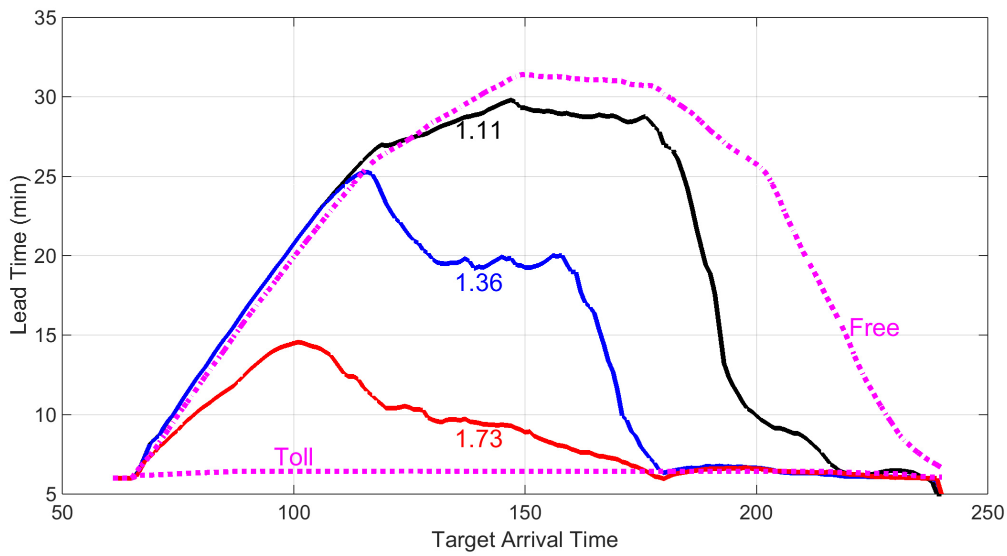


Fig. 7: Average lead time for selected VOTs with PI tolls.

6. Conclusions

This paper presented a HOT lane evaluation framework incorporating departure time choice and stochasticity in demand. Four toll algorithms were chosen and compared as input demand varied. Over the range of demand levels considered in this paper, FU-MEAN and FU-DM tolls perform almost exactly as well as the full information FU-PI toll scheme, which is encouraging. Furthermore, FU-DM often resulted in greater travel time reliability (lower standard

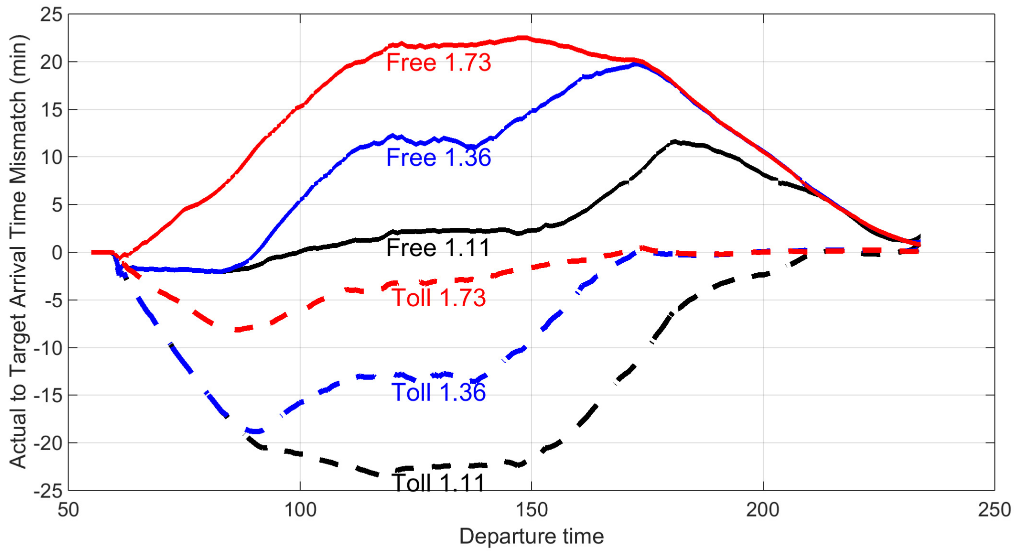


Fig. 8: Mismatch for selected VOTs with PI tolls.

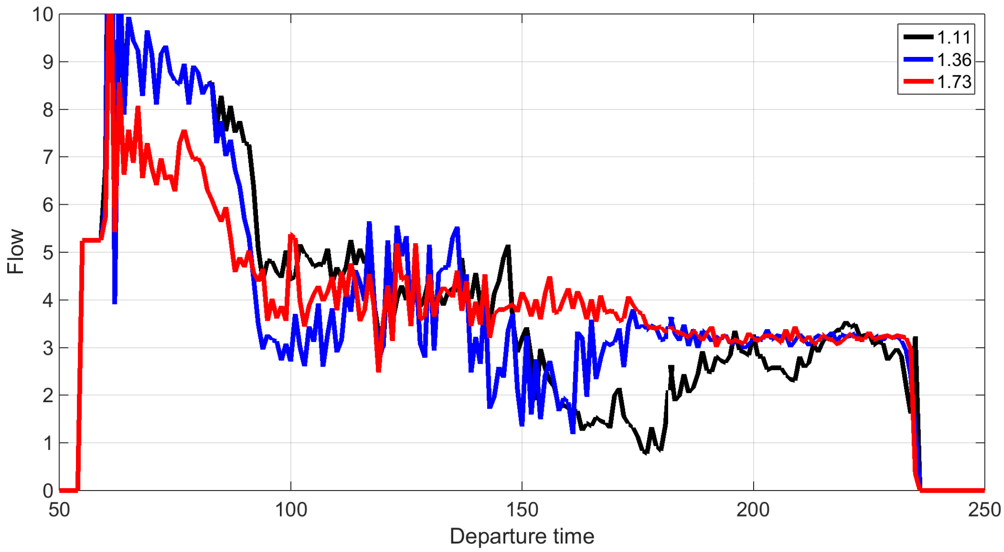


Fig. 9: Departure demand profile for selected VOTs with PI tolls.

deviation) than the other algorithms, perhaps because of its responsive nature. Any of the tolled lane configurations or algorithms tends to result in lower average travel time than the untolled setup.

Future research can extend this model in several ways. First, and most importantly, the network topology should be generalized to allow multiple entry and exit points of the lanes, and perhaps also to include other routes, origins, and destinations explicitly. In this way, route choice can be integrated with departure time choice, and the distinction between captive and strategic drivers can be traced to a more basic principle based on which drivers choose routes which present the option of the HOT lane. The equilibrium problem can also be studied in more detail, and more efficient algorithms can be developed. Finally, tailoring the case studies to specific field data would allow the model to either be validated or point to other directions for future work.

Acknowledgements

This research was partially supported by the U. S. Department of Transportation through the Data-Supported Transportation Operations and Planning (D-STOP) Tier 1 University Transportation Center.

References

Arnott, R., de Palma, A., Lindsey, R., 1990. Economics of a bottleneck. Journal of Urban Economics 27, 111–130.

Arnott, R., de Palma, A., Lindsey, R., 1993. A structural model of peak-period congestion: a traffic bottleneck with elastic demand. The American Economic Review 83, 161–179.

Arnott, R., de Palma, A., Lindsey, R., 1998. Recent developments in the bottleneck model, in: Button, K.J., Verhoef, E.T. (Eds.), Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility. Edward Elgar Pub, pp. 79–110.

Bar-Gera, H., 2012. The fast lane to Tel-Aviv: High-occupancy toll project with a Pareto package. Presented at the 91st Annual Meeting of the Transportation Research Board, Washington, DC.

Bellei, G., Gentile, G., Meschini, L., Papola, N., 2006. A demand model with departure time choice for within-day dynamic traffic assignment. European Journal of Operational Research 175, 1557–1576.

Bliemer, M.C.J., van Amelsfort, D.H., 2010. Rewarding instead of charging road users: a model case study investigating effects on traffic conditions. European Transport 44, 23–40.

Boyles, S.D., Gardner, L.M., Waller, S.T., 2010. Congestion pricing for transportation networks under supply and demand uncertainty. Presented at the 89th Annual Meeting of the Transportation Research Board, Washington, DC.

Burris, M.W., Geiselbrecht, T.C., Goodin, G.D., MacGregor, M.E., 2011. Predetermining performance-based measures for managed lanes. Presented at the 90th Annual Meeting of the Transportation Research Board, Washington, DC.

Burris, M.W., Ungemah, D.H., Mahlawat, M., Pannu, M.S., 2009. Investigating the impact of tolls on high-occupancy-vehicle lanes using managed lanes. Transportation Research Record 2099, 113–122.

Cao, X., Xu, Z., Huang, A., 2011. Safety benefits of converting HOV lanes to HOT lanes: case study of I-394 MnPASS. Presented at the 90th Annual Meeting of the Transportation Research Board, Washington, DC.

Chu, X., 1995. Endogenous trip scheduling: the Henderson approach reformulated and compared with the Vickrey approach. Journal of Urban Economics 37, 324–343.

Chung, C., Recker, W., 2011. State-of-the-art assessment of toll rates for high-occupancy and toll lanes. Presented at the 90th Annual Meeting of the Transportation Research Board, Washington, DC.

Chung, Y., Choi, K., 2010. Dynamic toll concept to assess feasibility of high-occupancy vehicle lane on Kyungbu Freeway, South Korea. Tranpsortation Research Record 2178, 162–169.

Daganzo, C.F., 1985. The uniqueness of a time-dependent equilibrium distribution of arrivals at a single bottleneck. Transportation Science 19, 29–37.

Dahlgren, J., 2002. High-occupancy/toll lanes: where should they be implemented? Transportation Research Part A 36, 239–255.

de Palma, A., Lindsey, R., 2000. Private toll roads: competition under various ownership regimes. The Annals of Regional Science 34, 13–35.

Facchinei, F., Pang, J.S., 2003. Finite-dimensional variational inequalities and complementarity problems. Springer Verlag.

Fosgerau, M., 2011. How a fast lane may replace a congestion toll. Transportation Research Part B 45, 845–851.

Friesz, T.L., Bernstein, D., Suo, Z., Tobin, R.L., 2001. Dynamic network user equilibrium with state-dependent time lags. Networks and Spatial Economics 1, 319–347.

Gardner, L., Bar-Gera, H., Boyles, S.D., 2013. Development and comparison of choice models and tolling schemes for high-occupancy/toll (HOT) facilities. Transportation Research Part B 55, 142–153.

Gardner, L., Boyles, S.D., Bar-Gera, H., Tang, K., 2015. Robust tolling schemes for high-occupancy/toll (HOT) facilities under variable demand. Transportation Research Record 2450, 152–162.

Gardner, L., Boyles, S.D., Waller, S.T., 2011. Quantifying the benefit of responsive pricing and travel information in the stochastic congestion pricing problem. Transportation Research Part A 45, 204–218.

Gardner, L., Duthie, J., Unnikrishnan, A., Waller, S.T., 2008. Robust pricing for networks with demand uncertainty. Presented at the 87th Annual Meeting of the Transportation Research Board, Washington, DC.

Gonzales, E.J., Daganzo, C.F., 2012. Morning commute with competing modes and distributed demand: user equilibrium, system optimum, and pricing. Transportation Research Part B 46, 1519–1534.

Gordon, P., Moore II, J.E., Richardson, H.W., Williamson, C., 2004. The Economic Effects of Highway Widening: Tolled Lanes vs. General Purpose Lanes — Using an Integrated Impact Model. Technical Report FHWA/CA/OR-2005/12. Federal Highway Administration, US Department of Transportation.

Halvorson, R., Nookala, M., Buckeye, K.R., 2006. High occupancy toll lane innovations: I-394 MnPASS. Public Works Management & Policy 10, 242–255.

Hendrickson, C., Kocur, G., 1981. Schedule delay and departure time decisions in a deterministic model. Transportation Science 15, 62–77.

Knuth, D.E., 1997. The Art of Computer Programming, Volume 2: Seminumerical Algorithms. 3rd ed., Addison-Wesley.

Kuhn, B., Goodin, G., Ballard, A., Brewer, M., Brydia, R., Carson, J., Chrysler, S., Collier, T., Eisele, W., Fenno, D., Fitzpatrick, K., Ullman, B., Ullman, G., Venglar, S., 2011. Findings from Texas: Five Years of Research on Managed Lanes. Technical Report FHWA/TX-06/0-4160-25. Federal Highway Administration, United States Department of Transportation.

Kuwahara, M., 2001. A theoretical analysis on dynamic marginal cost pricing, in: Proceedings of the Sixth Conference of Hong Kong Society for Transportation Studies, pp. 28–39.

Lam, W.H.K., Small, K., 2001. The value of time and reliability: measurement from a value pricing experiment. Transportation Research Part E 37, 231–251.

Lam, W.H.K., Tam, M.L., 1998. Risk analysis of traffic and revenue forecasts for road investment projects. Journal of Infrastructure Systems 4, 19–27.

Li, I.Y., Ziliaskopoulos, A.K., Waller, S.T., 1999. Linear programming formulations for system optimum dynamic traffic assignment with arrival time based and departure time based demands. Transportation Research Record 1667, 52–59.

Liu, L.N., McDonald, J.F., 1999. Economic efficiency of second-best congestion pricing schemes in urban highway systems. Transportation Research Part B 33, 157–188.

Liu, Y., Nie, Y.M., 2010. Morning commute problem considering route choice, user heterogeneity and alternative system optima. Transportation Research Part B 36, 421–443.

Lou, Y., Yin, Y., Laval, J., 2011. Optimal dynamic pricing strategies for high-occupancy/toll lanes. Transportation Research Part C 19, 64–74.

Michalaka, D., Lou, Y., Yin, Y., 2011. Proactive and robust dynamic pricing strategies for high-occupancy-toll (HOT) lanes. Presented at the 90th Annual Meeting of the Transportation Research Board, Washington, DC.

Morgul, E.F., Ozbay, K., 2011. Simulation-based evaluation of a feedback based dynamic congestion pricing strategy on alternate facilities. Presented at the 90th Annual Meeting of the Transportation Research Board, Washington, DC.

Nie, Y.M., Ma, J., Zhang, H.M., 2008. A polymorphic dynamic network loading model. Computer-Aided Civil and Infrastructure Engineering 23, 86–103.

Ozbay, K., Morgul, E.F., Ukkusuri, S., Iyer, S., Yushimoto, W.F., 2011. Mesoscopic simulation evaluation of dynamic congestion pricing strategies for New York City crossings. Presented at the 90th Annual Meeting of the Transportation Research Board, Washington, DC.

Parkany, E., 1998. Can HOT lanes encourage carpooling? A case study of carpooling behavior on the 91 express lanes. Technical Report UCI-ITS-WP-98-5. UC Irvine Institute of Transportation Studies.

Smith, M.J., 1984. The existence of a time-dependent equilibrium distribution of arrivals at a single bottleneck. Transportation Science 18, 385–394.

Vickrey, W.S., 1969. Congestion theory and transport investment. American Economic Review 59, 251–261.

Yang, H., 1999a. Evaluating the benefits of a combined route guidance and road pricing system in a traffic network with recurrent congestion. Transportation 26, 299–322.

Yang, H., 1999b. System optimum, stochastic user equilibrium, and optimal link tolls. Transportation Science 33, 354–360.

Yang, H., Huang, H.J., 1997. Analysis of the time-varying pricing of a bottleneck with elastic demand using optimal control theory. Transportation Research Part B 31, 425–440.

Yin, Y., Lou, Y., 2009. Dynamic tolling strategies for managed lanes. Journal of Transportation Engineering 135, 45–52.

Ziliaskopoulos, A.K., Rao, L., 1999. A simultaneous route and departure time choice equilibrium model on dynamic networks. International Transactions in Operational Research 6, 21–37.