Hindawi
Journal of Advanced Transportation
Volume 2018, Article ID 6809304, 15 pages
https://doi.org/10.1155/2018/6809304

Research Article

Impact of Capacity Drop on Commuting Systems under Uncertainty

Wenwei Zhang, Hui Zhao, and Rui Jiang

MOE Key Laboratory for Urban Transportation Complex System Theory and Technology, School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

Correspondence should be addressed to Hui Zhao; zhaoh@bjtu.edu.cn and Rui Jiang; jiangrui@bjtu.edu.cn

Received 10 April 2018; Revised 30 August 2018; Accepted 26 September 2018; Published 17 October 2018

Academic Editor: Eneko Osaba

This paper analyzes the impact of capacity drop on commuters' travel choice behaviors under uncertainty. For clarity, we assume that the capacity drop is triggered by the queue forming at the bottleneck under the hypercongestion circumstances, and the stochasticity of the drop could not be neglected. Considering the uncertainty of travel time, we establish a bottleneck model with commuters choosing their departure time according to the mean travel cost. From the proposed model, analytical solutions are achieved and therefore several properties are presented, including monotonicity of travel cost and departure rate, and the relationship between dispersion degree and length of peak period. To alleviate traffic congestion at the bottleneck and avoid capacity drop, we design a time-varying toll scheme and a step toll scheme. Evolution of queue length in equilibrium is discussed based on the Laih model. Numerical examples are also presented to demonstrate the established model and the effectiveness of the proposed toll schemes.

1. Introduction

Capacity plays one key role in the commuting systems, and the capacity drop phenomenon attracts much research from the field of transportation. As presented by empirical studies, the self-organized capacity drop could be observed in bottlenecks at many locations. For instance, by observation in and near Toronto, Canada, Cassidy and Bertini [1] reported that the average rate discharge from a queue could be 10% lower than the flows measured prior to queue's formation. Latterly, Zhang and Levinson [2] observed 3-12% capacity drop on freeway bottlenecks in Twin Cities area, Minnesota. In addition, self-organized capacity drop phenomenon is also observed in a bottleneck located in the Twin Cities west metropolitan region by Srivastava and Geroliminis [3].

Though most of the existing literature on capacity drop in bottlenecks is from deterministic perspective, there is some research suggested that the self-organized capacity drop does not always occur at a fixed flow rate and is intrinsically stochastic. It is even pointed out that that capacity of a bottleneck could be defined as a function of breakdown probability [4, 5]. In fact, the stochastic properties may not be limited in the self-organized capacity drop phenomenon. For example, the capacity could fluctuate from day to day and could be treated as a stochastic one [6]. Based on the consideration, this paper is to investigate the impact of capacity drop at a bottleneck in a stochastic circumstance.

The analysis of this paper is based on the classical bottleneck model, which is originally developed by Vickrey [7]. The model could formulate commuters' trip schedule and was recognized as a useful tool for modeling the queuing behaviors at a bottleneck during the morning peak. In literature, many studies have extended the basic bottleneck model from different aspects, like considering the case of elastic demand [8] and commuters' heterogeneity [9, 10]. Specifically, the influences of various impact factors on the performance of commuting systems are discussed too. For example, Lindsey [11] investigated the effects of information provision under stochastic bottleneck capacity conditions. van den Berg and Verhoef [12] studied the influence of commuters' value of time on the efficiency of the commuting system. By considering the provision of parking spaces, Arnott et al. [13] examined the effects of parking availability on the commuting systems and evaluated the efficiency of road tolls and parking fees. Latterly, Qian et al. [14, 15] considered various impact factors including parking capacity allocations, parking fees

and access times for traffic congestion mitigation, and total social costs reduction. In addition, extensions of bottlenecks could also be found like day-to-day travel behavior evolution [16], preference of autonomous vehicle users [17], and travel choice of shared-riders [18, 19]. Highly correlated with the work in this paper, Fosgerau and Small [20] investigated the bottleneck model with variable capacity in the rush hour in the hypercongestion environment. However, the authors discussed the commuters' travel choice behavior under deterministic circumstances.

Besides the works on deterministic bottleneck model, the commuting systems are also discussed in stochastic environments. Fosgerau [21] derived a closed form expression for the expected cost in a bottleneck model with stochastic capacity and demand. Latterly, Fosgerau [22] presented characterization of how the expected delay and the variance of delay are related in a congested facility with random capacity and demand. Li et al. [23] analyzed the departure time choice behavior under stochastic capacity using Vickrey bottleneck model and developed numerical methods to solve a bottleneck where daily capacity follows a uniform distribution. Siu et al. [24] considered equilibrium trip scheduling under random travel delay in commuting system at rush hours. They found that random delay plays a significant role in travel costs and introduces substantial differences in the queuing pattern, departure, and arrival times. Xiao et al. [6] studied the bottleneck model with stochastic capacity with a uniform distribution; both analytical and numerical results show that the capacity variability of bottleneck leads to significant changes in departure time patterns. Xiao et al. [25, 26] extended the stochastic bottleneck model to heterogeneous values of time for studying the heterogeneous travelers' departure time choice behavior. Furthermore, commuters depart at the same time can experience early or late arrival depending on the capacity in the stochastic circumstances. Furthermore, the impact of stochastic capacity at the downstream bottleneck after a merge is also investigated by Xiao et al. [25, 26].

As the capacity drop degrades the performance of the commuting system, researchers also tried to find approaches for alleviate the efficiency loss caused by capacity drop. Among the proposed approaches, ramp metering is proved to be effective for enhancing the dropped capacity [2, 27]. In addition, variable speed limit seems to work also [28], at least at the theoretical level. However, there are many methods to eliminate the queue and decrease the externality cost of the commuting system, which are toll based schemes. As a seed work, Arnott et al. [29] discussed the time-varying toll scheme together with the step toll scheme for higher efficiency in a commuting system. Compared with the time-varying toll schemes which are not easy to implement, the step toll schemes are much more practical and were studied by many researchers [30, 31]. In the step toll schemes, some fixed pricing toll charges are set during time intervals. Commuter in front of the queue need to pay the charges according to the time intervals. In the single-step toll scheme proposed by Arnott et al. [29] which is called ADL model, the arrival period is affected and the total private travel cost is found to be higher than that in the no-toll equilibrium. Furthermore, to achieve equilibrium, a mass of individuals depart just after the toll is lifted. Laih [32] proposed a single-step toll scheme called Laih model. Under the assumption that there are separate queuing facilities where commuters can wait and no mass departures are needed when the toll comes, the arrival times and the total travel cost may not be affected by the toll scheme. By considering the case in which drivers have an incentive to wait for the toll decrement if the waiting cost is less than the amount of toll saved, Lindsey et al. [30] presented a single-step toll scheme called braking model. In the model, the bottleneck is idle for a certain period in equilibrium; thus the peak period will be prolonged and higher total travel cost could be expected. Besides the toll scheme, there is another approach to enhance the efficiency of the commuting system from economic perspective, which is credit scheme. It is proved theoretically that the tradable credit schemes are able to gain the same efficiency that toll scheme could achieve [33–35].

Based on the fact and above studies, we investigate the impact of stochastic capacity drop on commuters' departure time choice behavior in this paper by deriving the analysis solution at equilibrium state. For convenience, it is assumed that the bottleneck capacity will drop by a stochastic amount when the queue at the bottleneck reaches a certain level, i.e., the capacity of bottleneck becomes stochastic after the queue length exceeds the certain level. Such kind of changes in capacity of bottleneck will yield different trip costs and inevitably affect the departure time choice behavior of the commuters. Furthermore, we design a time-varying toll to avoid capacity drop by eliminating queue. Because the time-varying toll is practically difficult to employ, here we also propose a step toll scheme to avoid capacity drop associated with a bottleneck.

The rest of this paper is organized as follows. Section 2 formulates the commuters' travel costs and departure time choice in a bottleneck and reports the closed form solution and properties of the proposed model. Impacts of the stochastic capacity drop are discussed analytically. Section 3 presents both the time-varying toll scheme and the step toll scheme to avoid capacity drop for efficiency improvement. In Section 4, numerical examples are presented and sensitivity analysis is conducted. Finally, Section 5 concludes the paper.

2. The Bottleneck Model with Stochastic Capacity Drop

2.1. The Basic Model. The basic model could be seen as a simplification of the work by Vickrey [7]. Assume that there is a fixed number (N) of commuters from residential district to the Central Business District (CBD) through a highway with a single bottleneck. Since the bottleneck capacity is limited in the morning rush hour, if the commuters' departure rate exceeds the capacity s, a queue forms at the bottleneck. For any commuter traveling from residential location to the CBD, the travel time contains two parts, which are queuing time and free flow travel time of highway t_{free}. For simplicity, it is assumed that $t_{free} = 0$. For all commuters, the preferred arrival time (work start time) is set to be t^*; one faces a schedule delay cost no matter he or she arrives early or late.

We assume the departure rate at time point x is $r(x)$ and the earliest departure time is t_0.

As there are a lot of symbols with various of subscripts in the paper, a table that summarizes all used notations is presented as an appendix (see the appendix).

The cumulative departures $R(t)$ could be formulated as

$$R(t) = \int_{t_0}^{t} r(x)\,dx. \tag{1}$$

In the morning rush hour, the capacity of the bottleneck is fully utilized from time t_0 and the length (in vehicles) of the queue is

$$Q(t) = \max\left\{R(t) - s\left(t - t_0\right), 0\right\}. \tag{2}$$

Therefore, the queuing time for commuters depart at time t is

$$T(t) = \frac{Q(t)}{s}. \tag{3}$$

The total cost could be seen as the sum of the travel time cost and schedule delay cost:

$$C(t) = \alpha T(t) + \begin{cases} \beta\left(t^* - t - T(t)\right), & if\ t^* \geq t + T(t); \\ \gamma\left(t + T(t) - t^*\right), & if\ t^* \leq t + T(t); \end{cases} \tag{4}$$

where α, β, and γ are the values of travel time, schedule delay early (SDE), and schedule delay late (SDL). In accordance with empirical finding by Small (1982), we assume $\gamma > \alpha > \beta$ for the existence and uniqueness of the equilibrium.

At equilibrium, no commuter can reduce his or her travel cost by unilaterally altering his or her departure time; i.e., $dC(t)/dt = 0$. Then the departure rate during the rush hour is

$$r(t) = \begin{cases} \dfrac{\alpha s}{(\alpha - \beta)} & if\ t_0 \leq t \leq t_t \\ \dfrac{\alpha s}{(\alpha + \gamma)} & if\ t_t \leq t \leq t_e \end{cases} \tag{5}$$

where t_t is the departure time at which an individual arrives at the CBD on time t^* and t_e is the ending time of the rush hour period. Because all commuters incur the same cost, and the duration of the peak is N/s, we can deduce that

$$t_0 = t^* - \frac{\gamma N}{(\beta + \gamma)s}, \qquad t_e = t^* + \frac{\beta N}{(\beta + \gamma)s}. \tag{6}$$

The commuter's travel cost at equilibrium could also be obtained by looking at the first commuter:

$$\overline{C} = \beta\left(t^* - t_0\right) = \frac{N}{s}\frac{\beta\gamma}{\beta + \gamma}. \tag{7}$$

The longest queue at equilibrium is

$$q_0 = \frac{N}{\alpha}\frac{\beta\gamma}{\beta + \gamma}. \tag{8}$$

2.2. Problem Settings and Assumptions. In this paper, we assume that the capacity of the single bottleneck will drop when the queue at the bottleneck reaches a certain level, denoted as q_c; thus there are basically two stages during the rush hour. Before the capacity drop occurs, the capacity of bottleneck is relatively high and stable. Let $\hat{s}$ be the capacity of the bottleneck during the first stage. As the queue length exceeds the critical length q_c, the second stage begins, and during the second stage, the capacity drops down stochastically. The dropped capacity of the bottleneck could be seen as a stochastic variable and follows a uniform distribution within interval $[\theta\overline{s}, \overline{s}]$, where $\overline{s} = \lambda\hat{s}$, $\lambda < 1$, representing the percentage of the remaining capacity, and $0 < \theta \leq 1$, representing the lowest rate of available capacity. Let $f(s)$ be the probability density function of the capacity; it follows that

$$f(s) = \frac{1}{\overline{s} - \theta\overline{s}}. \tag{9}$$

In addition, we present the basic assumptions listed below:

(1) Commuters are homogeneous with the same values of travel time and schedule delays.

(2) Once the capacity has dropped, it would not recover until the queue dissipates. Then the intermediate travelers who are already in the queue should keep on waiting and passing through the bottleneck.

(3) At the second stage, the capacity of the bottleneck is constant within a day but fluctuates from day to day [6].

(4) Commuters are aware of the capacity drop probability, and their departure time choice follows the user equilibrium principle in terms of mean travel cost.

2.3. Formulation of the Bottleneck Model with Capacity Drop. Under the above conditions, the commuters' travel cost is not always deterministic; commuters are assumed to make departure time decisions according to the mean trip cost. For simplicity, we set the preferred arrival time t^* equal to 0. The mean travel cost with respect to departure time t could be formulated as

$$E[C(t)] = \alpha E[T(t)] + \beta E\left[\overline{SDE}(t)\right] + \gamma E\left[\overline{SDL}(t)\right]. \tag{10}$$

Here $\overline{SDE}(t)$ and $\overline{SDL}(t)$ are the schedule delay early and late for commuters departing at time t, where

$$\overline{SDE}(t) = \max\{0, -T(t) - t\}, \tag{11}$$

and

$$\overline{SDL}(t) = \max\{0, T(t) + t\}. \tag{12}$$

We consider a very congested rush hour, in which commuters always queue behind the bottleneck. As assumed above, at the first stage, the capacity of the bottleneck is deterministic; thus once the departure time is chosen by a commuter, his or her travel time and schedule delay are both deterministic. At the second stage, due to the stochasticity of the capacity, commuters departing at the same time may

endure schedule delay early or late in different days. They may or may not encounter queuing delays randomly. For the complexity of the commuter's choice behaviors, the rush hour could be divided into five departure intervals to characterize the different travel choice behaviors as listed below:

(I) Commuters always arrive early and pass the bottleneck before the capacity drop.

(II) Commuters always arrive early and pass the bottleneck after the capacity drop.

(III) Depending on capacity commuters can arrive early or late.

(IV) Commuters always arrive late and incur a queuing delay.

(V) Commuters always arrive late and, depending on capacity, may or may not incur a queuing delay.

Here we use t_1, t_3, t_4, and t_5 to denote the time points that separate the five situations. For clarity, here we assume that commuters queuing at length q_c exactly do not suffer the capacity drop, which means the capacity drop begins after the commuters queuing at time point t_1 with length q_c passed the bottleneck. For convenience, let t_2 denote the time point at which the capacity drops and the queue length reach q_c at time point t_1. it follows that

$$t_1 + \frac{q_c}{\hat{s}} = t_2. \tag{13}$$

As the equilibrium condition for commuters' departure time choice is that no commuter can reduce their mean trip cost by unilaterally altering his or her departure time, the five situations under the user equilibrium scheme could be summarized as follows.

Situation I. Commuters always arrive early in $[t_0, t_1]$, and commuters departing at this interval will pass the bottleneck before the bottleneck capacity drops.

The departure rate is derived by differentiating (10) and setting the derivative as zero; i.e.,

$$\frac{dE\left[C\left(t\right)\right]}{dt} = 0, \quad \text{if } r\left(t\right) > 0. \tag{14}$$

The departure rate in this interval is

$$r_1\left(t\right) = \frac{\alpha\hat{s}}{\alpha - \beta}, \quad t_0 \le t \le t_1, \tag{15}$$

where t_0 is the earliest departure time. The boundary condition for this situation is that the cumulative number of departures by time t_1 reaches N_1; that means

$$R\left(t_1\right) = N_1 = \frac{\alpha}{\beta} q_c. \tag{16}$$

Situation II. Commuters always arrive early in $(t_1, t_3]$, and commuters departing at this interval will pass the bottleneck after the bottleneck capacity drops.

Using the equilibrium condition, we can obtain the departure rate in this interval:

$$r_2\left(t\right) = \frac{\alpha\overline{s}\left(1-\theta\right)}{\left(\alpha - \beta\right)\ln\theta^{-1}}, \quad t_1 < t \le t_3. \tag{17}$$

When the capacity of the bottleneck at the second stage is $\theta\overline{s}$, $\overline{SDE}(t_3) = 0$. Thus we have

$$R\left(t_3\right) = N_1 + \theta\overline{s}\left(t^* - t_2\right). \tag{18}$$

Situation III. Depending on capacity, commuters can arrive early or late in $(t_3, t_4]$.

If the capacity of the bottleneck at the second stage is $\overline{s}$, no commuters experience schedule delay late. If the capacity of the bottleneck at the second stage is $\theta\overline{s}$, no commuters experience schedule delay early. In other cases, i.e., where the capacity of the bottleneck is distributed stochastically within the interval $[\theta\overline{s}, \overline{s}]$, either schedule delay early or late may occur. If commuters departing at time point t arrive at their workplaces on time, the corresponding capacity is

$$s = \frac{R\left(t\right) - N_1}{-t_2}. \tag{19}$$

The departure rate in this interval is

$$r_3\left(t\right) = \frac{\alpha}{A + B\left(\ln\left(R\left(t\right) - N_1\right) + 1\right)}, \quad t_3 < t \le t_4, \tag{20}$$

where

$$\begin{aligned} A &= -\frac{\alpha\ln\theta + \beta\ln\left(-t_2\overline{s}\right) + \gamma\ln\left(-t_2\theta\overline{s}\right) + \left(\beta + \gamma\right)}{\overline{s} - \theta\overline{s}}, \\ B &= \frac{\beta + \gamma}{\overline{s} - \theta\overline{s}}. \end{aligned} \tag{21}$$

When the capacity of the bottleneck at the second stage is $\overline{s}$, $\overline{SDE}(t_4) = 0$. Thus we have

$$R\left(t_4\right) = N_1 + \overline{s}\left(t^* - t_2\right). \tag{22}$$

Situation IV. Commuters always arrive late and incur a queuing delay in $(t_4, t_5]$.

Using the equilibrium condition, we can obtain the departure rate within this interval is:

$$r_4\left(t\right) = \frac{\alpha\overline{s}\left(1-\theta\right)}{\left(\alpha + \gamma\right)\ln\theta^{-1}}, \quad t_4 < t \le t_5. \tag{23}$$

The boundary condition for this situation is

$$R\left(t_5\right) = N_1 + \overline{s}\left(t_5 - t_2\right). \tag{24}$$

That means the queuing length at time instant t_5 is 0, if the capacity at the second stage is $\overline{s}$.

Situation V. Commuters always arrive late and, depending on capacity, may or may not incur a queuing delay in $(t_5, t_e]$.

For any time instant t in $(t_5, t_e]$, there is a watershed capacity of the bottleneck such that the queuing length equals

zeros; i.e., $R(t) - N_1 = s(t - t_2)$; the corresponding watershed capacity is

$$s = \frac{R(t) - N_1}{t - t_2}. \tag{25}$$

We then get the departure rate in this interval:

$$r_5(t) = \frac{(\alpha+\gamma)\left(R(t) - N_1\right)/(t - t_2) - (\alpha\theta + \gamma)\overline{s}}{(\alpha+\gamma)\left(\ln\left(R(t) - N_1\right) - \ln\left(\theta\overline{s}(t - t_2)\right)\right)}, \quad t_5 < t \le t_e. \tag{26}$$

The boundary condition for this situation is $r_5(t_e) = 0$, and we could also have

$$R(t_e) = N_1 + s^*(t_e - t_2), \quad \text{where } s^* = \frac{(\alpha\theta + \gamma)\overline{s}}{\alpha + \gamma}. \tag{27}$$

Since t_e are the ending times of the rush hour period, there is no more departure after that; we have $R(t_e) = N$, and From (13) and (27) we have

$$t_e = t_0 + \frac{N_1}{\widehat{s}} + \frac{N - N_1}{s^*}. \tag{28}$$

The mean trip cost is identical for all commuters, so $E[C(t_0)] = E[C(t_e)] = -t_0\beta$. Thus, we have

$$t_0 = \frac{N_1}{\widehat{s}}\left(-\frac{\gamma}{\beta+\gamma}\right) + \frac{N - N_1}{s^*}\left(\frac{1}{k_0 - 1}\right), \tag{29}$$

and

$$t_e = \frac{N_1}{\widehat{s}}\left(\frac{\beta}{\beta+\gamma}\right) + \frac{N - N_1}{s^*}\left(\frac{k_0}{k_0 - 1}\right). \tag{30}$$

Here

$$k_0 = 1 - \frac{(1-\theta)(\beta+\gamma)\overline{s}}{(\alpha+\gamma)s^*(\ln s^* - \ln\theta\overline{s})}. \tag{31}$$

Using the boundary conditions and (13), we can obtain

$$t_1 = \frac{N_1}{\widehat{s}}\left(\frac{\alpha-\beta}{\alpha} - \frac{\gamma}{\beta+\gamma}\right) + \frac{N - N_1}{s^*}\left(\frac{1}{k_0 - 1}\right),$$
$$t_2 = \frac{N_1}{\widehat{s}}\left(\frac{\beta}{\beta+\gamma}\right) + \frac{N - N_1}{s^*}\left(\frac{1}{k_0 - 1}\right), \tag{32}$$

$$t_3 = \frac{N_1}{\widehat{s}}\left[\frac{\alpha-\beta}{\alpha}\left(1 + \frac{\theta\ln\theta}{1-\theta}\right) - \frac{\gamma}{\beta+\gamma}k_1\right] + \frac{N - N_1}{s^*}\left(\frac{k_1}{k_0 - 1}\right), \tag{33}$$

$$t_4 = \frac{N_1}{\widehat{s}}\left[\frac{\alpha+\gamma}{\alpha}\left(1 + \frac{\ln\theta}{1-\theta}\right) - \frac{\gamma}{\beta+\gamma}k_2\right] + \frac{N - N_1}{s^*}\left(\frac{k_2}{k_0 - 1}\right), \tag{34}$$

$$t_5 = \frac{N_1}{\widehat{s}}\left[1 + \frac{(1-\theta)\gamma}{(1-\theta)\alpha + (\alpha+\gamma)\ln\theta} - \frac{\gamma}{\beta+\gamma}k_3\right] + \frac{N - N_1}{s^*}\left(\frac{k_3}{k_0 - 1}\right), \tag{35}$$

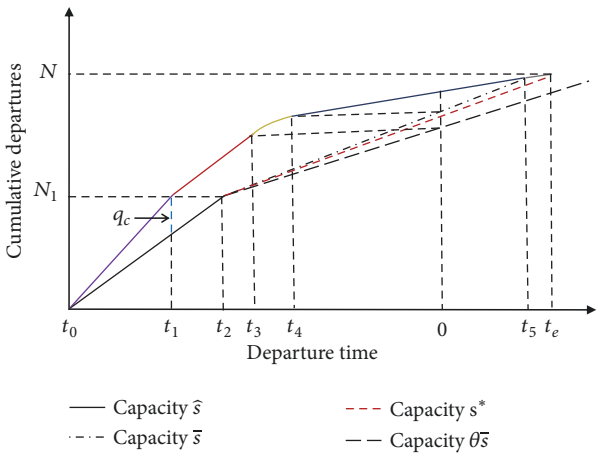


Figure 1: Equilibrium departures in a bottleneck.

where

$$k_1 = 1 - \frac{\alpha-\beta}{\alpha}\frac{\theta\ln\theta^{-1}}{1-\theta},$$
$$k_2 = \frac{\alpha+\beta+\gamma}{\alpha} + \frac{(\alpha+\gamma)\ln\theta}{\alpha(1-\theta)} \tag{36}$$
$$\text{and } k_3 = 1 + \frac{(1-\theta)(\beta+\gamma)}{\alpha(1-\theta) + (\alpha+\gamma)\ln\theta}.$$

The equilibrium cumulative departure and arrival curves are shown in Figure 1. From the figure, one can confirm that at the beginning, the departure rate is constant and then exceeds the capacity so that a queue forms; the last commuter who departs at time point t_1 will be the last one who passes the bottleneck before capacity drops, and commuters departing after that one will pass the bottleneck after the capacity drops. Under the influence of capacity drop, the departure rate after t_1 has a decrease but remains constant until t_3. When the capacity equals $\theta\overline{s}$, commuters departing at time instant t_3 will arrive at CBD on time. Then the departure rate gradually decreases until t_4. For commuters departing at time pint t_4, the capacity which equals $\overline{s}$ will make the commuters arrive on time. Thereafter, commuters arrive late, the departure rate will remain constant until t_5, and the departure rate continues to decrease until 0 at the last departure interval $(t_5, t_e]$.

2.4. Properties of the Watershed Bottleneck Model. To reveal interesting properties of the proposed bottleneck model and understand the impact of the stochastic capacity drop, here we present the following theorems.

Theorem 1. *At equilibrium, the expected trip cost is a strictly monotonically increasing function of the travel demand.*

Proof. At equilibrium, the expected trip cost is identical for all commuters. Since $E[C(t_0)] = -t_0\beta$ and $t_0 = (N_1/\widehat{s})(-\gamma/(\beta+\gamma)) + ((N - N_1)/s^*)(1/(k_0 - 1))$, we have

$$\frac{\partial E[C(t_0)]}{\partial N} = \frac{-\beta}{s^*(k_0 - 1)}. \tag{37}$$

Because $0 < \theta < 1$, it is easy to conclude $k_0 - 1 < 0$; then $\partial E[C(t_0)]/\partial N > 0$. Therefore, the expected trip cost is a strictly monotonically increasing function of the travel demand. □

For the departure rate and cumulative departures, the following two theorems could be presented.

Theorem 2. *At equilibrium, the departure rate is a monotonically decreasing function of the departure time t, $t \in [t_0, t_e]$.*

Proof. The theorem coincides with Theorem 2 in Xiao et al. [33] or Proposition 3 in Lindsey [36]. □

Theorem 3. *The following inequality holds:*

$$s^*(t - t_2) \le R(t) - N_1 \le \overline{s}(t - t_2), \quad t \in [t_5, t_e]. \tag{38}$$

Proof. Since the queue may exist during $[t_5, t_e]$, we have

$$R(t) \ge N_1 + \theta\overline{s}(t - t_2). \tag{39}$$

Additionally, the departure rate $r_5(t)$ is nonnegative, so we have

$$R(t) \ge N_1 + s^*(t - t_2) \tag{40}$$

and, on the other hand, since the queue may not exist during $[t_5, t_e]$,

$$R(t) - N_1 \le \overline{s}(t - t_2) \tag{41}$$

holds. □

For the impact of the stochasticity of capacity drop, the following theorem presents some analytical results.

Theorem 4. *When the value of the parameter θ is increasing, the length of peak period will decrease.*

Proof. From (30), the length of peak period is

$$t_e - t_0 = \frac{N_1}{\widehat{s}} + \frac{N - N_1}{s^*}. \tag{42}$$

Since s^* is a monotonically increasing function of θ and s^* as denominator, $t_e - t_0$ is a monotonically decreasing function of θ. □

Next, we discuss the effect of the capacity drop on equilibrium departures.

Theorem 5. *At equilibrium, the departure rates $r_3(t)$ and $r_5(t)$ are dependent on the critical queue length q_c.*

Proof. Because $r_3(t)$ is the solution to the nonlinear equation (20) and according to (16), the critical queue length q_c is increasing with respect to N_1; hence $r_3(t)$ is dependent on the critical queue length q_c.

Similarly, based on (26), the departure rate $r_5(t)$ is dependent on the critical queue length q_c. □

The departure rates $r_1(t)$, $r_2(t)$, and $r_4(t)$ are constant, and they are independent of the critical queue length q_c.

Theorem 6. *At equilibrium, the expected trip cost is a strictly monotonically decreasing function of the critical queue length q_c.*

Proof. The first order derivative of the expected trip cost with respect to q_c is

$$\begin{aligned}\frac{\partial E[C(t_0)]}{\partial q_c} &= \alpha\left(\frac{\gamma}{\widehat{s}(\beta+\gamma)} + \frac{1}{s^*(k_0 - 1)}\right) \\ &= \alpha\left(\frac{(1-\theta)\overline{s}\gamma - \widehat{s}(\alpha+\gamma)(\ln s^* - \ln\theta\overline{s})}{\widehat{s}\,\overline{s}(1-\theta)(\beta+\gamma)}\right).\end{aligned} \tag{43}$$

To prove $\partial E[C(t_0)]/\partial q_c < 0$, we need only to prove

$$(1-\theta)\overline{s}\gamma - \widehat{s}(\alpha+\gamma)(\ln s^* - \ln\theta\overline{s}) < 0. \tag{44}$$

The first order derivative of the left term of (44) with respect to θ is

$$-\overline{s}\gamma + \widehat{s}(\alpha+\gamma)\frac{\gamma}{\theta(\alpha\theta+\gamma)} > -\overline{s}\gamma + \widehat{s}\gamma > 0, \tag{45}$$

Then at point $\theta = 1$, term $(1-\theta)\overline{s}\gamma - \widehat{s}(\alpha+\gamma)(\ln s^* - \ln\theta\overline{s})$ reaches its maximum, and the maximum is 0, so $\partial E[C(t_0)]/\partial q_c < 0$ holds. □

Theorem 7. *With a fixed number of commuters and assuming that capacity drop would happen, the length of the morning peak period is shorter when the critical queue length q_c is larger, and vice versa.*

Proof. According to (28), the length of the morning peak period is as follows:

$$\begin{aligned}t_e - t_0 &= \frac{N_1}{\widehat{s}} + \frac{N - N_1}{s^*} = N_1\left(\frac{1}{\widehat{s}} - \frac{1}{s^*}\right) + \frac{N}{s^*} \\ &= \frac{\alpha q_c}{\beta}\left(\frac{1}{\widehat{s}} - \frac{1}{s^*}\right) + \frac{N}{s^*}.\end{aligned} \tag{46}$$

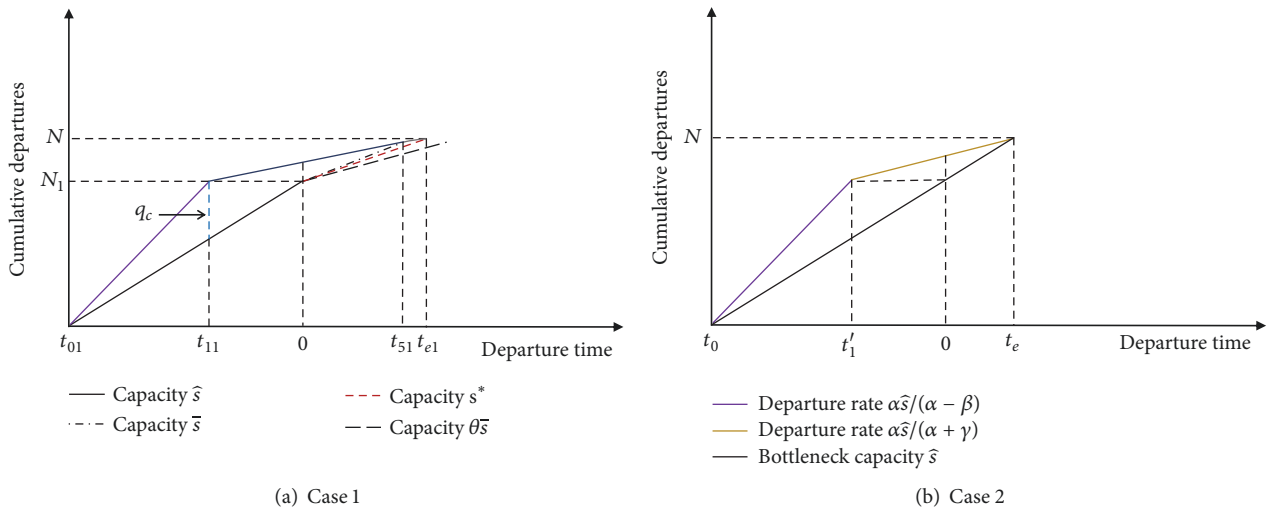
Since $0 < \theta < 1$, then $s^* = (\alpha\theta+\gamma)\overline{s}/(\alpha+\gamma) < (\alpha+\gamma)\overline{s}/(\alpha+\gamma) = \overline{s} < \widehat{s}$; we have $1/\widehat{s} - 1/s^* < 0$. Therefore, the length of peak period $t_e - t_0$ is monotonically decreasing with respect to q_c. □

The above three theorems confirm the influence of capacity drop. Briefly, the length of the morning peak period is longer and the expected trip cost of all travelers is larger if the critical queue length q_c is shorter. In fact, the length of first stage is as follows:

$$t_2 - t_0 = \frac{N_1}{\widehat{s}} = \frac{q_c}{\widehat{s}}\frac{\alpha}{\beta}, \tag{47}$$

so decreasing the critical queue length q_c is equivalent to the capacity drop taking place earlier.

2.5. Possible Equilibrium Traffic Flow Patterns. As shown in Figure 2, the other two possible departure patterns can happen, which depend on the critical queue length q_c. The



(a) Case 1 (b) Case 2

Figure 2: Possible equilibrium departure flow patterns.

details of the two possible departure patterns are given as follows:

Case 1 in Figure 2(a). Three situations (i.e., Situations I, IV, and V) happen in this case, and the corresponding departure rates of the three situations are $r_1(t)$, $r_4(t)$, and $r_5(t)$, respectively. We can observe from Figure 2(a) that commuters departing before t_{11} experience schedule delay early and always experience queuing; they will pass the bottleneck before the bottleneck capacity drops. Thereafter, the queue length increased to q_c when $t = t_{11}$, and capacity drop happens when $t = 0$. Commuters during (t_{11}, t_{51}) experience schedule delay late and always experience queuing, and commuters departing after t_{51} experience schedule delay late and possibly experience queuing.

Case 2 in Figure 2(b). Two situations (i.e., Situations I and IV) happen in this case, and the corresponding departure rates of the three situations are $\alpha s/(\alpha-\beta)$ and $\alpha s/(\alpha+\gamma)$. We can observe from Figure 2(b) that commuters departing before t_1' experience schedule delay early and always experience queuing; the queue length initially increases when $t < t_1'$, but the queue length $Q(t_1') < q_c$; then the queue length decreases when $t > t_1'$, so the capacity drop will not be triggered, and the bottleneck model follows the deterministic model. Commuters during (t_1', t_e) experience schedule delay late and always experience queuing.

Meanwhile in case 1, by definition, we have

$$t_2 = t^* = 0, \tag{48}$$

and

$$N_1 = \hat{s}(-t_{01}). \tag{49}$$

Since the departure rates in the case have not changed, according the derivation process of t_{01}, (29) still holds; substituting (48) and (49) into (29), we have

$$\bar{N}_1 = \frac{(N/s^*)(1/k_0 - 1)}{(1/s^*(k_0 - 1) - \beta/\hat{s}(\beta+\gamma))},$$

$$\bar{q}_c = \bar{N}_1 * \frac{\beta}{\alpha}. \tag{50}$$

If $q_c < \bar{q}_c$, the commuting behaviors could be investigated analytically in the five situations; if $q_c \geq \bar{q}_c$, no commuters experience schedule delay early and pass the bottleneck after capacity drops.

The critical time points in Case 1 are given as follows:

$$t_{01} = \frac{N/\hat{s}s^*(1-k_0)}{(1/s^*(k_0-1) - \beta/\hat{s}(\beta+\gamma))}, \tag{51}$$

$$t_{e1} = \frac{-N\beta/\hat{s}s^*(\beta+\gamma)}{(1/s^*(k_0-1) - \beta/\hat{s}(\beta+\gamma))}, \tag{52}$$

$$t_{11} = \frac{N\beta/\hat{s}s^*\alpha(1-k_0)}{(1/s^*(k_0-1) - \beta/\hat{s}(\beta+\gamma))}, \tag{53}$$

$$t_{51} = \frac{(N\beta/\hat{s}s^*\alpha(1-k_0))(\alpha(1-\theta)/((1-\theta)\alpha + (\alpha+\gamma)\ln\theta))}{(1/s^*(k_0-1) - \beta/\hat{s}(\beta+\gamma))}. \tag{54}$$

Case 2 and the basic deterministic bottleneck model have the same set of formulae for time instants t_0 and t_e. Let t_1' in Case 2 be the departure time for which an individual arrives at work on time. Then

$$t_1' = t^* - \frac{\beta\gamma N}{\alpha(\beta+\gamma)s}. \tag{55}$$

Theorem 8. *The maximum critical queue length q_c which can trigger capacity drop is bigger than the longest queue in the basic deterministic bottleneck model; i.e., $\bar{q}_c > q_0$.*

Proof. According to (8) and (50), we have the following.

$$\bar{q}_c - q_0 = \frac{(N/s^*)(1/(k_0-1))}{(1/s^*(k_0-1) - \beta/\hat{s}(\beta+\gamma))(\beta/\alpha)} - \frac{N}{\alpha}$$

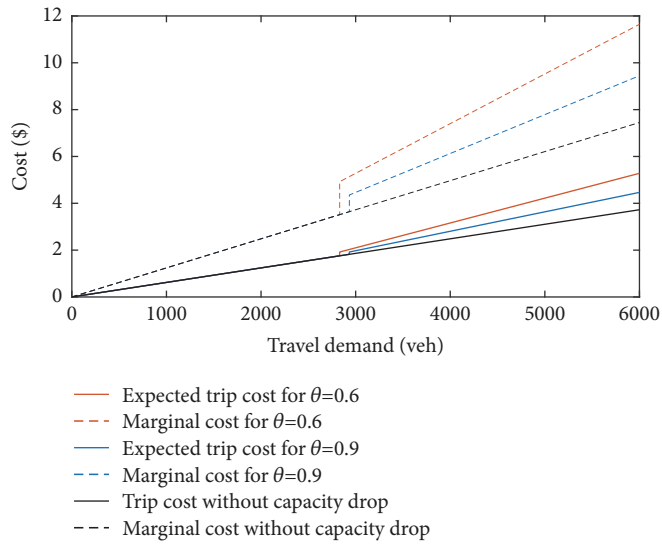


Figure 3: Expected trip cost and marginal cost functions.

$$\cdot \frac{\beta\gamma}{\beta+\gamma} = \frac{N\beta}{\alpha}\left(\frac{1/s^*(k_0-1)}{(1/s^*(k_0-1)-\beta/\hat{s}(\beta+\gamma))} - \frac{\gamma}{\beta+\gamma}\right) = \frac{N\beta}{\alpha(\beta+\gamma)} \cdot \frac{\beta}{(\alpha+\gamma)\hat{s}(\ln s^* - \ln\theta\overline{s}) + \beta(1-\theta)\overline{s}}(\hat{s}(\alpha+\gamma) \cdot (\ln s^* - \ln\theta\overline{s}) - (1-\theta)\overline{s}\gamma) \tag{56}$$

According to (44), we have $\hat{s}(\alpha+\gamma)(\ln s^* - \ln\theta\overline{s}) - (1-\theta)\overline{s}\gamma > 0$. Because $0 < \theta < 1$, it is easy to conclude $\ln s^* - \ln\theta\overline{s} > 0$. Therefore, the right-hand side of (56) is more than 0, so, $\overline{q}_c > q_0$ holds. □

Theorem 8 presents an interesting result. That is, if commuters believe the capacity drop will happen in the bottleneck, they will leave home earlier to avoid potential losses due to capacity drop, even if one may face a queue longer than the longest queue with the capacity drop. And in fact the capacity drop with critical queue length $\overline{q}_c$ should never happen.

Next, we study the expected trip cost varying with the traffic demand N over three distinct possible equilibrium patterns.

According to (50), for a given critical queue length q_c, the capacity drop is activated over some time interval when $N = N_{case1}$, where

$$N_{case1} = q_c \frac{\alpha s^*(k_0-1)}{\beta}\left(\frac{1}{s^*(k_0-1)} - \frac{\beta}{\hat{s}(\beta+\gamma)}\right). \tag{57}$$

The three distinct possible equilibrium patterns, i.e., Case 2, Case 1, and the basic Case depicted in Figure 1, arise from successively larger values of N. Thus the demand-varying expected trip cost can be formulated as a piecewise function.

$$ec(N) = E[C(t_0)] = \begin{cases} \frac{N}{\hat{s}}\frac{\beta\gamma}{\beta+\gamma}, & N < N_{case1} \\ \frac{\alpha}{\hat{s}}q_c, & N = N_{case1} \\ q_c\alpha\left(\frac{\gamma}{\hat{s}(\beta+\gamma)} + \frac{1}{s^*(k_0-1)}\right) - \frac{\beta N}{s^*(k_0-1)}, & N > N_{case1} \end{cases} \tag{58}$$

It is straightforward to verify that the expected trip cost function is upper semicontinuous at N_{case1}.

Similar with the work by Fosgerau and Small [20], we present the marginal cost to understand the relationship between system performance and congestion.

The marginal cost is the derivative of total cost with respect to N. Thus, we have the following.

$$mc(N) = \begin{cases} \frac{2N}{\hat{s}}\frac{\beta\gamma}{\beta+\gamma}, & N < N_{case1} \\ q_c\alpha\left(\frac{\gamma}{\hat{s}(\beta+\gamma)} + \frac{1}{s^*(k_0-1)}\right) - \frac{2\beta N}{s^*(k_0-1)}, & N \geq N_{case1} \end{cases} \tag{59}$$

Using the parameters presented in Table 1, the expected trip cost and marginal cost are given in Figure 3. The figure shows that the expected trip cost and marginal cost increase nonlinearly; both are discontinuous at N_{case1}. Hence, when heavy congestion exists, the congestion is more sensitive to travel demand. In addition, the larger capacity uncertainty (i.e., lessening the value of the parameter θ) will further increase the sensitivity.

2.6. Comparison with Related Works. As Vickrey [7] developed a bottleneck model to capture the dynamics of traffic congestion, at equilibrium state, all commuters experience the same travel cost no matter when they leave home. Fosgerau and Small [20] analyzed an apparently complex system

Table 1: Base parameters for a simple example.

Parameters	Values
α	6.4
β	3.9
γ	15.21
$\widehat{s}$	5000
$\overline{s}$	4000
θ	0.9
q_c	1500

similarly to a bottleneck whose capacity declines when the queue is long enough. Xiao et al. [6] assumed the bottleneck capacity follows a uniform distribution and investigated the equilibrium associated with the stochastic bottleneck. The stochasticity of capacity variation combined is investigated in the watershed bottleneck model proposed in this paper, there are two stages with different capacity. The first stage follows the schedule delay early part of the deterministic bottleneck model, but the bottleneck capacity will drop in the second state into a stochastic amount when the queue at the bottleneck reaches a certain level. Therefore, travelers passing through the bottleneck with a high capacity need to leave home earlier, and under the influence of capacity drop and stochastic capacity, travelers departing at the same time of a day may arrive at their destination either early or late, but the commuters' mean trip cost is constant with respect to a time instant.

From the theoretical aspects, Theorems 1, 2, and 4 in this paper are coincident with Theorems 1 and 2 and Proposition 3 in Xiao et al. [6]. Different from Xiao et al. [6], the departure rates in Situations III and V are dependent on the critical queue length q_c and the occurrence time of capacity drop. In addition, the critical queue length q_c will affect the equilibrium traffic flow pattern; this paper distinguishes two cases of equilibria associated with the capacity drop and derived the boundary conditions for all cases. We found that shrinking the length of the first stage will result in an increase in the length of peak period; when the critical queue length is set as $\overline{q}_c$, only Situations I, IV, and V will happen in the bottleneck, and $\overline{q}_c > q_0$, which means the capacity drop that should not happen also happened.

The three distinct possible equilibrium patterns could emerge with successively larger travel demand. Under the hypercongestion circumstances, the capacity drops and the expected trip cost and marginal cost increase with travel demand. The conclusion coincides with the results for a bottleneck with variable capacity investigated by Fosgerau and Small [20]. Moreover, it is found that the stochasticity of capacity variation will lead to the discontinuity of expected trip cost. In addition, larger capacity stochasticity could further increase the marginal cost with a higher speed with the travel demand, which means higher stochasticity could make the marginal cost more sensitive with travel demand variation.

From the above theorems, one could expect shorter congestion time period if the dispersion of capacity drop is relatively low. Therefore, it is reasonable to consider that control strategies aimed at traffic flow regularity could enhance the efficiency of the commuting systems from the congestion duration time point of view.

3. Time-Varying Toll and Step Toll Schemes

The capacity drop extends the morning peak and leads to loss of efficiency; hence it is meaningful to find approaches to alleviate the negative effects of capacity drop. From the theorems given in the previous section on bottleneck capacity drop, approaches to shortening the queue length could be expected to avoid the capacity drop. As the time-varying toll scheme could completely eliminate queue at bottleneck and therefore could avoid the capacity drop, in this section we design a time-varying toll scheme to enhance the efficiency of bottleneck. In addition, because of the difficulty for using the time-varying toll scheme in real world commuting system, we also propose a step toll scheme to achieve higher performance more practically.

3.1. Time-Varying Toll Scheme. The time-varying toll was designed as the first-best toll by Arnott et al. [29] in the general bottleneck model; they showed that the toll scheme could eliminate queue completely without raising schedule delay. Similarly, in this paper, we present the toll pricing which is designed to match the marginal travel time cost in the no-toll equilibrium:

$$f(t) = \begin{cases} 0, & t < t_0 \\ \overline{C} + t\beta, & t \in [t_0, 0) \\ \overline{C} - t\gamma, & t \in [0, t_e] \\ 0, & t > t_e \end{cases} \tag{60}$$

where t_0 and t_e are given by (6) and $\overline{C}$ is given by (7).

Since the queue could be eliminated completely under the time-varying toll scheme, the total cost of a trip and the peak period are the same as in the no-toll equilibrium. Noticing that the capacity drop studied in this paper associated with the bottleneck will not be triggered under the time-varying toll scheme, then the equilibrium flow pattern is the same as the case if capacity drop phenomenon is not considered.

Although the resulting departure and arrival patterns are identical for all sizes of travel demand, the performance for deploying the policy for the different travel demand is different. Firstly, relative to the congestion equilibrium with capacity drop, the bottleneck capacity could keep at the highest level under this toll scheme, which could decrease the commuters' mean travel cost and shorten the peak period. Secondly, all queuing costs are eliminated under the toll scheme, and the travel cost for each commuter is $N\beta\gamma/2\widehat{s}(\beta + \gamma)$. Combining both of the two cost saving parts, the total saving cost could be formulated as the difference between expected trip cost (ec) with and without time-varying toll as follows.

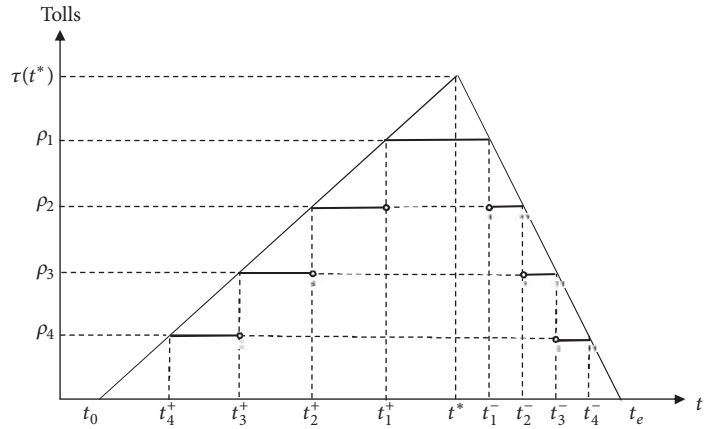


Figure 4: The 4-step toll inscribed in the optimal time-varying toll.

$$ec - ec\left(time\ varying\ tolled\right) = \begin{cases} \dfrac{N}{2\hat{s}}\dfrac{\beta\gamma}{\beta+\gamma}, & N < N_{case1} \\ q_c\alpha\left(\dfrac{\gamma}{\hat{s}(\beta+\gamma)} + \dfrac{1}{s^*(k_0-1)}\right) - \dfrac{\beta N}{s^*(k_0-1)} - \dfrac{N}{2\hat{s}}\dfrac{\beta\gamma}{\beta+\gamma}, & N \geq N_{case1} \end{cases} \tag{61}$$

It is not difficult to verify that the total cost saving increases with travel demand N. Essentially, under condition $N < N_{case1}$, the commuters' difference between expected cost with and without time-varying toll is zero because the cost saving parts could be covered by the toll.

3.2. Step Toll Scheme. Since the time-varying toll is difficult to implement and hard to be accepted for government and commuters, now we try to find a more practical toll scheme, which is step toll scheme. This type of toll scheme could limit the queue length in a range to avoid capacity drop with finite number of steps.

As reviewed above, there are mainly three types of step toll scheme including ADL model, Laih model, and braking model [30]. In this paper, we design a step toll scheme based on Laih model to avoid the capacity drop, and we also try to find least number of steps that should be used to achieve optimal efficiency.

Firstly, we introduce some notations about the step toll scheme used in Laih model. It is assumed that there are m levels of toll pricing in the schedule; the tolling period i spans a time interval $[t_i^+, t_i^-]$ with a toll level ρ_i, $i = 2, \ldots, m$. The tolling periods are indexed and period 1 is the central period containing t^*. For clarity, a diagrammatic sketch for a 4-step toll scheme is illustrated in Figure 4. As is presented, the lower-charge interval $[t_{i+1}^+, t_{i+1}^-]$ is disconnected during the higher-charge interval $[t_i^+, t_i^-]$; optimal toll levels for the Laih model are

$$\begin{aligned} \rho_1 &= \frac{m}{m+1}\frac{N}{s}\frac{\beta\gamma}{\beta+\gamma}, \\ \rho_i &= \frac{m+1-i}{m}\rho_1, \quad i = 2, \ldots, m. \end{aligned} \tag{62}$$

The optimal tolling periods for the Laih model are

$$\begin{aligned} t_i^+ &= t_0 + \frac{\rho_i}{\beta}, \\ t_i^- &= t_e - \frac{\rho_i}{\gamma}, \\ & i = 1, \ldots, m. \end{aligned} \tag{63}$$

To have a deeper insight of the equilibrium under the proposed step toll scheme, we discuss the queue length evolution process in equilibrium.

Case 1 (before tolling period 1). Under the proposed multistep toll scheme, the length of queue is 0 at the moment when the toll ρ_i is applied and also immediately before it is lifted. For commuters arriving early within $[t_i^+, t_{i-1}^+]$, the one arriving just before t_{i-1}^+ must experience a queuing time $(\rho_{i-1} - \rho_i)/\alpha$, which is longer than the queuing time the commuter arriving after t_{i-1}^+ may face and pay the toll ρ_{i-1}. Therefore, the longest queue length during time interval $[t_i^+, t_{i-1}^+]$ is $(\rho_{i-1} - \rho_i)s/\alpha$, which could be simplified as

$$q_i^+ = \frac{(\rho_{i-1} - \rho_i)\, s}{\alpha} = \frac{1}{m+1} q_0, \quad i = 2, \ldots, m+1. \tag{64}$$

where q_{m+1}^+ denotes the longest queue length during time interval $[t_0, t_m^+]$ and q_0 is the longest queue under the equilibrium in the original setting without capacity drop.

Case 2 (during tolling period 1). The departure rate at equilibrium is $\alpha s/(\alpha - \beta)$ for commuters who arrive early and is $\alpha s/(\alpha + \gamma)$ for commuters who arrive late, which is also the same as the rates in the no-toll equilibrium. The optimal multistep toll could be set so that the queue reaches 0 at t_1^-, the

total number of departures in $[t_1^+, t_1^-]$ is $s(t_1^- - t_1^+) = N/(m+1)$, and the longest queue length in $[t_1^+, t_1^-]$ is, therefore,

$$q_1 = \frac{1}{m+1} q_0. \tag{65}$$

Case 3 (after tolling period 1). As in the Laih model, it is assumed that commuters who arrive just after t_{i-1}^- may wait on a set of secondary lanes without impeding other commuters who arrive before t_{i-1}^- and use the main set of lanes; there are separate queues which will dissipate at t_i^-. To achieve the equilibrium, the commuter arriving just after t_{i-1}^- must experience a waiting time that equals $(\rho_{i-1} - \rho_i)/\alpha$, which is equivalent to the case in which the commuter faces a queue with length $(\rho_{i-1} - \rho_i)s/\alpha$. However, as illustrated in Figure 5, the longest queue at the secondary lanes is

$$\begin{aligned} q_i^- &= \frac{1}{\alpha+\gamma} \frac{N}{m+1} \frac{\beta\gamma}{\beta+\gamma} < \frac{1}{\alpha} \frac{N}{m+1} \frac{\beta\gamma}{\beta+\gamma} \\ &= \frac{1}{m+1} q_0, \quad i = 2, \ldots, m+1. \end{aligned} \tag{66}$$

In summary, the longest queue length under the m-step toll scheme should be $q_0/(m+1)$. Thus, if we let

$$m^* = \left\lfloor \frac{q_0}{q_c} \right\rfloor, \tag{67}$$

where $\lfloor \bullet \rfloor$ denotes the ceil operator, then the longest queue length is

$$q^* = \frac{1}{m^*+1} q_0 = \frac{1}{\lfloor q_0/q_c \rfloor + 1} q_0 < q_c. \tag{68}$$

It is straightforward to conclude that the m^*-step toll scheme can guarantee that the commuters face a queue length less than the critical length q_c and capacity drop does not arise; on the other hand, the m^*-step toll scheme with least number of steps is convenient for practical application.

Similar to the time-varying toll, the m^*-step toll could avoid the capacity drop phenomenon and further shorten the queue length at bottleneck. The average cost with the m^*-step toll for each commuter is $(1 - m/2(m+1))(N/\hat{s})(\beta\gamma/(\beta+\gamma))$; then the total cost saving for each commuter could be formulated as follows.

$$ec - ec\,(step\ tolled) = \begin{cases} \dfrac{m}{2\,(m+1)} \dfrac{N}{\hat{s}} \dfrac{\beta\gamma}{\beta+\gamma}, & N < N_{case1} \\ q_c\alpha \left(\dfrac{\gamma}{\hat{s}\,(\beta+\gamma)} + \dfrac{1}{s^*\,(k_0-1)} \right) - \dfrac{\beta N}{s^*\,(k_0-1)} - \left(1 - \dfrac{m}{2\,(m+1)}\right) \dfrac{N}{\hat{s}} \dfrac{\beta\gamma}{\beta+\gamma}, & N \geq N_{case1} \end{cases} \tag{69}$$

From (69), it could also be concluded that the step toll asymptotically approaches the efficiency of the first-best toll as $m \longrightarrow \infty$.

4. Numerical Examples

In this section, numerical results are presented for the bottleneck model without toll, with a time-varying toll, and with a multistep toll. As in Arnott et al. (1990b) and Xiao et al. [33], we first use α = 6.4\$/hr, β = 3.9\$/hr, and γ = 15.21\$/hr and consider the situation with N = 6000 veh, $\hat{s}$ = 5000 veh/hr, $\overline{s}$ = 4000 veh/hr, and θ = 0.9, To ensure that all five situations could appear in no-toll equilibrium, the watershed number N_1 must be smaller than $\overline{N}_1 \approx 5000$ veh; thus we consider N_1 = 3000 veh, and the corresponding critical queue length that activates the capacity drop is q_c = 1828 veh.

4.1. No-Toll Equilibrium in the Bottleneck. To obtain the departure rates in Situation III and V, we solve the differential equations (20) and (26) using the Euler method with step size equal to 0.005. Figure 6 shows the cumulative departure flows in the bottleneck. The rush hour starts at -1.12 and ends at 0.26. Note that t_2 separates the two stages mentioned above, but it is irrelevant to these watershed times that separate the five situations, and it is not always before t_3.

The mean total travel cost, mean travel time cost, and the mean schedule delay and early costs are shown in Figure 7; at equilibrium, the mean total travel costs of all commuters are the same and equal to 4.35; the travel time cost reaches the highest point at time instant around -0.69, but the time point is neither at the one with the minimum of schedule delay early/late cost nor at the intersection of the schedule delay early cost curves and the schedule delay late cost curves. As expected, the queue exists and the travel time cost is nonzero at the end of the peak period. Interestingly, although the capacity of bottleneck will turn from deterministic to stochastic as it drops, the mean travel time costs are still linearly increasing or decreasing with departure time.

To understand the impact of stochasticity of the bottleneck capacity, Table 2 reports the sensitivity analysis results for different values of θ and q_c. We change the θ-value from 0.75 to 1.0 when q_c = 1828 veh, and q_c from 500 to 2500 when θ = 0.9, respectively; it is shown that both the length of peak period and the mean trip cost increase as the θ-value decreases, which is consistent with Theorem 4. Because smaller θ-value will bring more uncertainty, then commuters will depart earlier to avoid potential losses, which will also increase the total travel cost of the whole system. As many works use the concept reliability to capture the stochasticity of the transportation systems [37], higher level of stochasticity may lead to more unreliability and make the systems less efficient consequently. In addition, from the table, we can

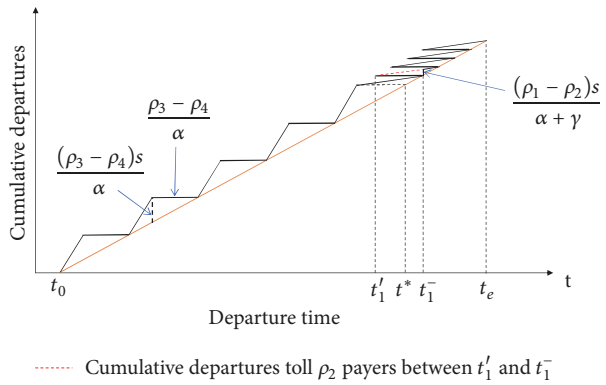


Figure 5: Equilibrium with a 4-step toll: the Laih model.

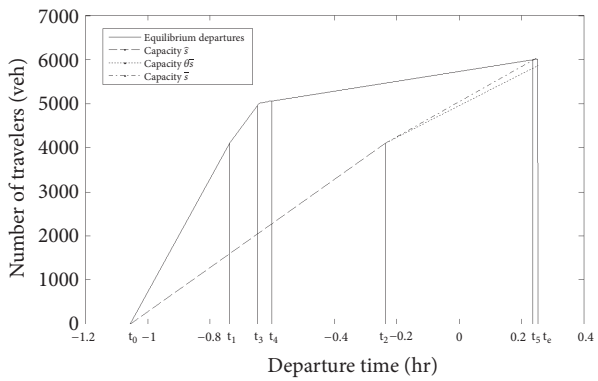


Figure 6: Equilibrium departures in the bottleneck.

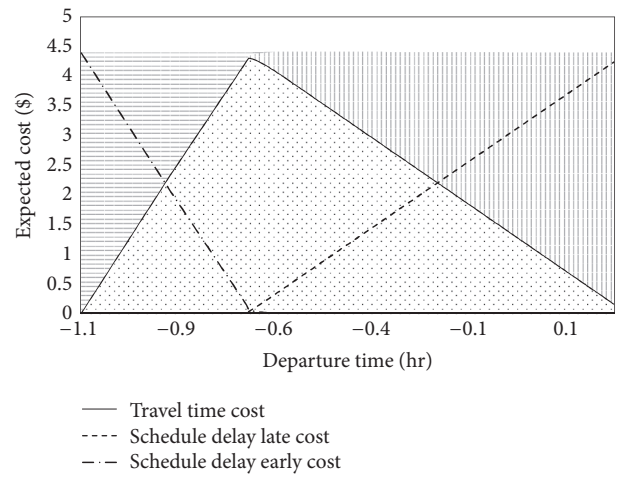


Figure 7: Mean trip cost and its components in no-toll equilibrium.

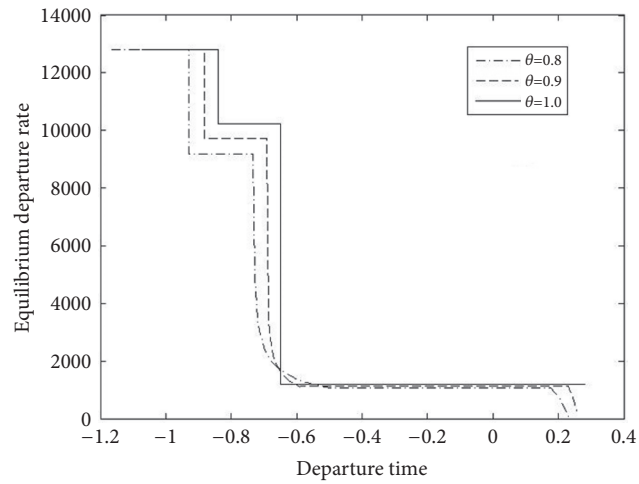


Figure 8: Influence of parameter θ on departure rate.

also find that when $\theta = 1$, time points $t_3 = t_4$ and $t_5 = t_e$. It means that the elimination of the stochasticity of bottleneck model degrade into the deterministic model. By contrast, both the length of peak period and the mean trip cost decrease as the q_c increases, which is consistent with Theorem 4. Since increasing the critical queue length is equivalent to prolonging the period of the first stage of the bottleneck, commuters departing at the same time could enjoy higher capacity than the one in the model with shorter critical queue length.

Figure 8 depicts the departure rates for different θ-values. It shows that in equilibrium the departure rate is monotonically decreasing with time, and the departure rate is around zero at the end of the peak period. Moreover, under the influence of the capacity drop of bottleneck, the departure rate suffers "a cliff-like drop" too, and when the θ-value approaches one, the stochastic bottleneck model follows the deterministic model.

4.2. Time-Varying Toll in the Bottleneck Model. For equilibrium under time-varying toll scheme, the results are given in Table 3 and are identical to the case without capacity drop as discussed in Section 2.

4.3. Step Toll in the Bottleneck Model. With the input data of $\alpha = 6.4$ \$/hr, $\beta = 3.9$ \$/hr, $\gamma = 15.21$\$/hr, $N = 6000$ veh, $s = 5000$ veh/hr, the longest queue in bottleneck model when capacity drop is not considered is $q_0 = 2910$ veh. To avoid the capacity drop, the number of steps in the toll schedule needs to satisfy $m \geq \lfloor q_0/q_c \rfloor = 1$; i.e., a single-step toll can avoid the capacity drop under this condition.

Despite the practicability which step toll schemes may have, the loss of efficiency should also be considered. In the following, we compare the performance of the bottleneck with single-step toll and the time-varying toll in Table 3. Specifically, the notations TC, TTC, and SDC denote total travel cost, total travel time cost, and total schedule delay cost, respectively.

From Table 4, one could see that in terms of the total travel time cost, the single-step toll is not so effective compared to time-varying toll. Compared with classical Vickrey model, the single-step toll could eliminate half of the total queuing time that exists in the no-toll equilibrium when capacity drop does not happen.

As multistep toll scheme can avoid capacity drop with the least number of steps, Figure 9 shows the least number of steps of multistep toll scheme for different q_c values. From the figure, one can see that q_c value could affect the least number

Table 2: Influence of parameter θ and q_c on the mean trip cost and the watershed time instants.

		$E[C]$	t_0	t_1	t_2	t_3	t_4	t_5	t_e	$t_e - t_0$
θ	1.00	4.19	-1.07	-0.84	-0.47	-0.65	-0.65	0.28	0.28	1.35
	0.95	4.27	-1.09	-0.86	-0.49	-0.67	-0.62	0.25	0.27	1.36
	0.90	4.35	-1.12	-0.88	-0.52	-0.69	-0.59	0.23	0.26	1.37
	0.85	4.44	-1.14	-0.91	-0.54	-0.71	-0.54	0.20	0.25	1.38
	0.80	4.54	-1.16	-0.93	-0.56	-0.73	-0.49	0.18	0.23	1.40
	0.75	4.65	-1.19	-0.96	-0.59	-0.76	-0.43	0.15	0.22	1.41
q_c	500	4.81	-1.23	-1.17	-1.07	-0.77	-0.56	0.22	0.27	1.50
	1000	4.64	-1.19	-1.06	-0.86	-0.74	-0.57	0.22	0.26	1.45
	1500	4.47	-1.15	-0.95	-0.65	-0.71	-0.58	0.23	0.26	1.40
	2000	4.29	-1.10	-0.84	-0.44	-0.68	-0.59	0.23	0.26	1.36
	2500	4.12	-1.06	-0.74	-0.24	-0.65	-0.60	0.24	0.25	1.31

Table 3: Peak period and individual travel cost under two schemes ($\theta = 0.9$).

	t_0	t_e	$E[C]$	$t_e - t_0$
No-toll	-1.1162	0.2567	4.3530	1.3729
Time-varying toll	-0.9551	0.2449	3.7249	1.2000

Table 4: TC, TTC, and SDC with time-varying toll and single-step toll.

	TC ($)	TTC ($)	SDC ($)	The Longest queue (veh)
Time-varying toll	22350	0	11175	0
Single-step toll	22350	5587.5	11175	1455

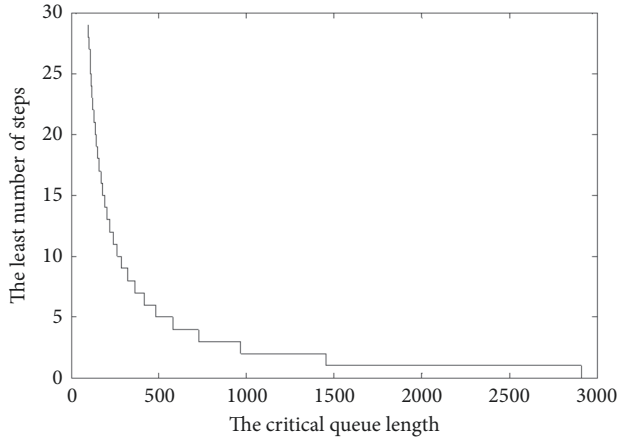


Figure 9: The least number of steps for different q_c values.

of steps. Furthermore, as the q_c value increases, the decrement speed of the least number of steps becomes slow.

5. Conclusions

In this paper, we analyze the impacts of capacity drop on commuting systems using a bottleneck model. Following the empirical findings by Lorenz and Elefteriadou [4] and Brilon et al. [5], it is assumed that the bottleneck capacity could drop stochastically because of the irregularity of traffic flow. Under the assumption that the commuters' departure time choice follows UE principle in terms of mean travel cost, analytical solutions have been derived in five situations. Several properties of the proposed bottleneck model with stochastic capacity drop are discussed, including monotonicity of travel cost and departure rate, and the relationship between dispersion degree and length of peak period. In addition, the other two possible departure patterns associated with the capacity drop are given.

Furthermore, two types of toll pricing schemes are also proposed to avoid the capacity drop, including time-varying toll scheme and step toll scheme. Specifically, the time-varying toll could eliminate the queue completely. The resulting equilibrium flow pattern is the same as that when capacity drop could not happen. To seek more practical approaches to enhance the efficiency of the commuting system, step toll pricing scheme is proposed and discussed. For any given critical queue length, multistep toll with least number of steps could be designed to achieve the optimal efficiency of the bottleneck under the step toll schemes.

In addition, numerical results are also presented to demonstrate the impact of capacity drop. It is shown that the capacity drop phenomenon would increase the mean travel cost, and it is helpful to reduce the critical queue length to enhance the efficiency of the system by using approaches for regulating the traffic flow.

Appendix

A. Notations

N: Number of commuters

s: Bottleneck capacity

t_{free}: Free flow travel time

t^*: Preferred arrival time

t_0: Earliest departure time

t_t: Departure time at which an individual arrives at the CBD on time t^*

t_e: Ending time of the rush hour

$r(\bullet)$: Departure rate

$R(\bullet)$: Cumulative departures

$Q(t)$: Queue length at time t

$T(t)$: Queuing time at time t

$C(t)$: Total cost at time t

$\overline{C}$: Travel cost at equilibrium

α: Values of travel time

β: Values of schedule delay early

γ: Values of schedule delay late

q_0: Longest queue length at equilibrium

q_c: Critical point of queue length for capacity drop

$\widehat{s}$: Full capacity of the bottleneck

$\overline{s}$: Upper bound of the dropped capacity

λ: Parameter to characterizing the capacity drop, as $\overline{s} = \lambda\widehat{s}$

θ: Parameter to characterizing the lower bound of dropped capacity

t_1: The time point at which the queue length reaches q_c

t_2: The time point at which the capacity drops

t_3: The time point at which separate commuters "always arrive early" and "can arrive early or late"

t_4: The time point at which separate commuters "arrive early or late" and "always arrive late and incur a queuing delay"

t_5: The time point at which separate commuters "always arrive late and incur a queuing delay" and "always arrive late and, depending on capacity, may or may not incur a queuing delay"

t_{01}: Earliest departure time in case 1

t_{11}: The time point in case 1 at which the queue length reaches $\overline{q}_c$

t_{51}: The time point in case 1 at which separate commuters "always arrive late and incur a queuing delay" and "always arrive late and, depending on capacity, may or may not incur a queuing delay"

t_{e1}: Ending time of the rush hour in case 1

t_1': The departure time in case 2 for which an individual arrives at work on time

N_1: Cumulative number of departures by time t_1

$f(t)$: Toll pricing matches the marginal cost

m: Numbers of toll levels

t_i^+: Start time point of the tolling period at level i

t_i^-: End time point of the tolling period at level i

ρ_i: Toll at level i

m^*: Least number of step toll levels

q^*: Longest queue length with least number of step toll levels.

Data Availability

All relevant data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work is jointly supported by the National Natural Science Foundation of China (Grant No. 71371028, 71621001) and Beijing Natural Science Foundation (Grant No. 9172013).

References

[1] M. J. Cassidy and R. L. Bertini, "Some traffic features at freeway bottlenecks," *Transportation Research Part B: Methodological*, vol. 33, no. 1, pp. 25–42, 1999.

[2] L. Zhang and D. Levinson, "Some properties of flows at freeway bottlenecks," *Transportation Research Record*, no. 1883, pp. 122–131, 2004.

[3] A. Srivastava and N. Geroliminis, "Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model," *Transportation Research Part C: Emerging Technologies*, vol. 30, pp. 161–177, 2013.

[4] M. R. Lorenz and L. Elefteriadou, "Defining freeway capacity as function of breakdown probability," *Transportation Research Record*, vol. 1776, pp. 43–51, 2001.

[5] W. Brilon, J. Geistefeldt, and M. Regler, "Reliability of freeway traffic flow: a stochastic concept of capacity," in *Proceedings of the 16th International symposium on transportation and traffic theory*, College Park Maryland, 2005.

[6] L.-L. Xiao, H.-J. Huang, and R. Liu, "Congestion behavior and tolls in a bottleneck model with stochastic capacity," *Transportation Science*, vol. 49, no. 1, pp. 47–65, 2015.

[7] W. Vickrey, "Congestion theory and transport investment," *American Economic Review*, vol. 59, pp. 251–261, 1969.

[8] R. Arnott, A. de Palma, and R. Lindsey, "A structural model of peak-period congestion: a traffic bottleneck with elastic demand," *The American Economic Review*, vol. 83, no. 1, pp. 161–179, 1993.

[9] R. Lindsey, "Existence, uniqueness, and trip cost function properties of user equilibrium in the bottleneck model with multiple user classes," *Transportation Science*, vol. 38, no. 3, pp. 293–314, 2004.

[10] G. Ramadurai, S. V. Ukkusuri, J. Zhao, and J.-S. Pang, "Linear complementarity formulation for single bottleneck model with

heterogeneous commuters," *Transportation Research Part B: Methodological*, vol. 44, no. 2, pp. 193–214, 2010.

[11] R. Lindsey, *Effects of Driver Information in the Bottleneck Model*, Department of Economics, University of Alberta, 1996.

[12] V. van den Berg and E. T. Verhoef, "Congestion tolling in the bottleneck model with heterogeneous values of time," *Transportation Research Part B: Methodological*, vol. 45, no. 1, pp. 60–78, 2011.

[13] R. Arnott, A. de Palma, and R. Lindsey, "A temporal and spatial equilibrium analysis of commuter parking," *Journal of Public Economics*, vol. 45, no. 3, pp. 301–335, 1991.

[14] Z. S. Qian, F. E. Xiao, and H. M. Zhang, "The economics of parking provision for the morning commute," *Transportation Research Part A: Policy and Practice*, vol. 45, no. 9, pp. 861–879, 2011.

[15] Z. S. Qian, F. E. Xiao, and H. M. Zhang, "Managing morning commute traffic with parking," *Transportation Research Part B: Methodological*, vol. 46, no. 7, pp. 894–916, 2012.

[16] R.-Y. Guo, H. Yang, H.-J. Huang, and X. Li, "Day-to-day departure time choice under bounded rationality in the bottleneck model," *Transportation Research Part B: Methodological*, pp. 551–570, 2017.

[17] V. A. C. van den Berg and E. T. Verhoef, "Autonomous cars and dynamic bottleneck congestion: the effects on capacity, value of time and preference heterogeneity," *Transportation Research Part B: Methodological*, vol. 94, pp. 43–60, 2016.

[18] F. Zhang, W. Liu, X. Wang, and H. Yang, "A new look at the morning commute with household shared-ride: How does school location play a role?" *Transportation Research Part E: Logistics and Transportation Review*, vol. 103, pp. 198–217, 2017.

[19] Y. Liu and Y. Li, "Pricing scheme design of ridesharing program in morning commute problem," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 156–177, 2017.

[20] M. Fosgerau and K. A. Small, "Hypercongestion in downtown metropolis," *Journal of Urban Economics*, vol. 76, no. 1, pp. 122–134, 2013.

[21] M. Fosgerau, "Congestion costs in a bottleneck model with stochastic capacity and demand," in *Munich Personal RePEc Archive paper*, University Library of Munich, Germany, 2008.

[22] M. Fosgerau, "On the relation between the mean and variance of delay in dynamic queues with random capacity and demand," *Journal of Economic Dynamics & Control*, vol. 34, no. 4, pp. 598–603, 2010.

[23] H. Li, P. H. Bovy, and M. C. Bliemer, "Departure time distribution in the stochastic bottleneck model," *International Journal of Its Research*, vol. 6, no. 2, pp. 79–86, 2008.

[24] B. W. Y. Siu and K. L. Hong, *Equilibrium Trip Scheduling in Congested Traffic under Uncertainty*, Springer US, 2009.

[25] L.-L. Xiao, H.-J. Huang, and L.-J. Tian, "Stochastic bottleneck model with heterogeneous travelers," *Journal of Transportation Systems Engineering and Information Technology*, vol. 14, no. 4, pp. 93–98, 2014.

[26] L.-L. Xiao, R. Liu, and H.-J. Huang, "Stochastic bottleneck capacity, merging traffic and morning commute," *Transportation Research Part E: Logistics and Transportation Review*, vol. 64, pp. 48–70, 2014.

[27] J. Laval, M. Cassidy, and C. Daganzo, "Impacts of lane changes at merge bottlenecks: a theory and strategies to maximize capacity," in *Traffic and Granular Flow'05*, pp. 577–586, 2007.

[28] W. Liu, Y. Yin, and H. Yang, "Effectiveness of variable speed limits considering commuters' long-term response," *Transportation Research Part B: Methodological*, vol. 81, pp. 498–519, 2015.

[29] R. Arnott, A. de Palma, and R. Lindsey, "Economics of a bottleneck," *Journal of Urban Economics*, vol. 27, no. 1, pp. 111–130, 1990.

[30] C. R. Lindsey, V. A. C. Van den Berg, and E. T. Verhoef, "Step tolling with bottleneck queuing congestion," *Journal of Urban Economics*, vol. 72, no. 1, pp. 46–59, 2012.

[31] Z.-C. Li, W. H. K. Lam, and S. C. Wong, "Step tolling in an activity-based bottleneck model," *Transportation Research Part B: Methodological*, vol. 101, pp. 306–334, 2017.

[32] C.-H. Laih, "Effects of the optimal step toll scheme on equilibrium commuter behaviour," *Applied Economics*, vol. 36, no. 1, pp. 59–81, 2004.

[33] F. Xiao, Z. Qian, and H. M. Zhang, "Managing bottleneck congestion with tradable credits," *Transportation Research Part B: Methodological*, vol. 56, pp. 1–14, 2013.

[34] Y. M. Nie and Y. Yin, "Managing rush hour travel choices with tradable credit scheme," *Transportation Research Part B: Methodological*, vol. 50, pp. 1–19, 2013.

[35] L. Tian, H. Yang, and H. Huang, "Tradable credit schemes for managing bottleneck congestion and modal split with heterogeneous users," *Transportation Research Part E: Logistics and Transportation Review*, vol. 54, pp. 1–13, 2013.

[36] R. Lindsey, "Optimal departure scheduling for the morning rush hour when capacity is uncertain," in *Proceedings of the 7th World Conference on Transport Research*, World Transport Research, 1996.

[37] B. Du and D. Z. W. Wang, "Continuum modeling of park-and-ride services considering travel time reliability and heterogeneous commuters—a linear complementarity system approach," *Transportation Research Part E: Logistics and Transportation Review*, vol. 71, pp. 58–81, 2014.