Bottleneck Congestion with Traffic Jam:

A reformulation and correction of earlier results

Se-il Mun

Graduate School of Economics, Kyoto University

Yoshida Honmachi, Sakyo-ku, Kyoto 606-8501, Japan

Fax: +81-75-753-3492,

E-mail: mun@econ.kyoto-u.ac.jp

September 2006

Abstract

This paper presents a dynamic model of road congestion with traffic jam, based on shockwave theory. Our model includes both the traditional flow congestion model and the Vickrey type bottleneck model as special cases. The model is applied to the peak-load problem for the morning rush hour. We derive no-toll equilibrium with commuters' departure choice and optimal road use minimizing the total cost for commuting. It is shown that, as the peak period congestion is heavier, introducing a peak-load toll attains larger gain in social welfare, with relatively smaller loss of private welfare.

JEL classification: R41; R48; D62

Keywords: Congestion; Bottleneck; Peak-load; Departure choice

1. Introduction

In the literature on transport economics, there are two major approaches to modeling traffic congestion on a roadway: the flow model and the bottleneck model. The flow model is based on the relationship between flow rate (= number of vehicles entering the road within a certain period of time) and speed. In this model, travel time cost is a function of the traffic flow rate. Models of this type have been widely applied to static analysis, for which Walters (1961) is the seminal work. On the other hand, the bottleneck model, first developed by Vickrey (1969), describes the congestion mechanism as queuing behind a bottleneck. Travel time is waiting time within the queue. Many economists seem to believe that these two approaches are independent from each other, coming from different backgrounds. In fact, however, they are abstractions of components from a more general framework of traffic flow theory (Lighthill and Whitham (1955)).

Mun (1994) attempted to reunify two approaches by introducing shockwave theory into the flow model. Shockwave theory is useful for describing the development of queues in a manner consistent with the flow model. Mun (1999) applied the model to the peak-load problem for the morning commute. It was shown that, in some cases, private welfare is improved by introducing a peak-load toll. This result is contrary to intuition. The present paper reformulates the previous model by relaxing the assumption concerning the propagation of queues. This reformulation yields a clearer relationship between our model and earlier models: our model includes both the traditional flow congestion model and the Vickrey type bottleneck model as special cases. Furthermore, it turns out that the result obtained in Mun (1999) is not valid under new model: private welfare is always decreased by introducing a peak-load toll.

The paper is organized as follows. Section 2 briefly reviews two major approaches to modeling traffic congestion, namely the flow model and the bottleneck model. Section 3 presents a dynamic model of traffic jam development based on shockwave theory, and compares it with earlier models described in Section 2. In Section 4, we apply the model to the peak-load problem for the morning commute: we formulate no-toll equilibrium and optimal road use, and derive peak-load toll to decentralize the optimum. Moreover we present a simulation analysis of equilibrium and optimal solutions to examine the effects of peak-load toll on private and social welfare. Section 5 concludes the paper.

2. Flow congestion and bottleneck model

2-1 Flow congestion

States of traffic flow are characterized by three variables: traffic volume (or flow rate), q; traffic density, K; and speed (velocity), V. The relations among these variables are given by the following equations.

$$V = f(K), \qquad f' < 0 \tag{1}$$

$$q = KV \tag{2}$$

Based on the above equations, the relation between traffic volume and speed is depicted in Figure 1. Two curves, $V_N(q)$ and $V_J(q)$ are drawn, so two values of speed exist for a given value of traffic volume. The $V_J(q)$ curve, for values of V less than V_M defines the relation between traffic volume and speed in a traffic jam. Hereafter, subscripts "N" and "J" of variables represent normal flow (no jam), and jammed flow, respectively.

Figure 1

Since the opportunity cost of travel time is a major part of the trip cost, it is assumed that the trip cost is proportional to the travel time. Following Walters (1961), most economic analyses of traffic congestion discuss the problem of traffic congestion supposing a uniform roadway. In this case, the travel time is defined as $T = L/V$, where L is the length of the roadway. Applying the relation in Fig. 1 to the above definition of travel time, the trip cost curve has a backward-bending portion. Economists refer to the backward-bending part of the cost curve "hyper-congestion". Some economists attempted to deal with hyper-congestion in a static framework supposing a uniform roadway (Else (1981), Nash (1982)), which were not successful and faced with difficulty (see, e. g., Small and Chu (2003)). As discussed in Mun (1994, 1999), Small and Chu (2003), Verhoef (2003), hyper-congestion, or traffic jam, is a dynamic phenomenon, which is not compatible with a static framework. Mun (1994, 1999) argues that traffic jams never occur in a uniform roadway, and the cost curve does not bend backward even if a traffic jam occurs.

Henderson (1985) analyzes the peak-load problem based on the model of flow congestion. He assumes that the speed of travelers entering the road is solely a function of the number of travelers entering with them, and the rise in traffic flow rate increases travel time. This is equivalent to assuming that only the upper segment of the curve in Fig. 1 applies. Let us rewrite the Henderson's model along the context of the traffic flow model described by (1) (2). Travel time for a driver entering the road at time *t* is obtained by the following formula

$$T(t) = L \big/ V_N(Q(t)) \tag{3}$$

where $Q(t)$ is the flow rate at time t[1]. Since $\frac{dV_N}{dQ} < 0$ from Figure 1, $\frac{dT}{dQ} > 0$ always holds. In the analysis of peak-load problem, flow rate $Q(t)$ varies over time and so does the travel time $T(t)$ accordingly. Variation of travel time is described as follows,

$$\dot{T}(t) = -\frac{L}{\left(V_N(Q(t))\right)^2}\frac{dV_N}{dQ(t)}\dot{Q}(t) \tag{4}$$

where · denotes a time derivative. The above discussion is an interpretation of Henderson's model in the context of traffic flow theory. This formulation, however, has a limitation as a tool for dynamic analysis of the peak-load problem, since travel time is bounded below a certain level, in other words it must not exceed L/V_M. The dynamic equilibrium of peak hours requires that travel time should increase over time for a substantially long period. Therefore equilibrium may not exist when the peak period is sufficiently long (i.e., the total number of commuters is large).

2-2 Bottleneck model

Suppose a single road connecting two nodes, say, the origin and the destination of trips, which has a bottleneck just before the destination. It is assumed that vehicles drive at a constant speed from the origin to the bottleneck point: the travel time for this portion of the trip is constant, and represented as T_f. A queue develops when the

[1] Henderson (1985) defines $Q(t)$ as the arrival rate at the destination (flow rate at the exit of the roadway). On the other hand, this paper defines $Q(t)$ as the departure rate (flow rate at the entrance of the roadway). This difference has no problem in the setting of Henderson (1985), since it is assumed that the speed for a traveler is constant throughout the journey. In other words, a flow rate at the entrance at time *t* is equal to the flow rate at the exit at time $t + T(t)$.

traffic flow rate (= departure rate at the origin) exceeds the bottleneck capacity, C_b. The travel time for a vehicle departing at t is represented as follows

$$T(t) = T_f + \frac{M(t)}{C_b} \tag{5}$$

where $T(t)$ is travel time, $M(t)$ is length of queue. The first term in the RHS of (5) is the time to travel from home to the bottleneck point, assuming that vehicles move at free speeds. The second term represents the waiting time within the queue behind the bottleneck. Note that $M(t)$ is measured as the number of vehicles within the queue. The queue length that a trip maker departing at t encounters is calculated as follows

$$M(t) = \int_{t_a}^{t} [Q(u) - C_b]du \tag{6}$$

where t_a is the time at which a queue starts to develop, and $Q(t)$ is the departure rate at t.

Differentiating (5) and (6) with respect to time, we have an equation of travel time variation as follows

$$\dot{T}(t) = \frac{\dot{M}(t)}{C_b} = \frac{Q(t) - s}{C_b} \tag{7}$$

Although this type of bottleneck model presents a useful framework for analyzing the dynamic process of bottleneck congestion in a tractable form, it introduces simplifying assumptions: it neglects flow congestion in which the speed of traffic depends on the flow rate; it ignores the space occupied by the queue. The above formulation implies that the physical queue length is zero: there is a large parking space at the entrance of the bottleneck so that all vehicles stop and wait their turns. Although this seems unrealistic, the analysis in the next section shows that the simple bottleneck model is consistent with traffic flow theory.

3. Dynamic model of traffic congestion

The model presented in this section is a modified version of that in Mun (1999).

Suppose a road connects two nodes A and B as illustrated in Figure 2, where trips originating at A are destined for B. The physical condition is uniform throughout the road with the exception of the bottleneck located just before the end point B. The capacity of the bottleneck and the rest of the road are C_b and C_a, respectively. All vehicles using the road are homogeneous.

Figure 2

As long as traffic volume, $Q(t)$, entering the road via point A at time *t* is less than C_b, an equal volume of vehicles will flow out of point B after a while[2]. In this case, the speed and density of the traffic are $V_N(Q(t))$ and $K_N(Q(t))$, respectively. On the other hand, when traffic volume exceeds C_b, overflowing vehicles form a queue behind the bottleneck. Within the queue, traffic is jammed and the speed and density are determined by $V_J(C_b)$ and $K_J(C_b)$, respectively (see Fig.1). Traffic flow is discontinuous at the tail end of the queue. The movement of the queue's tail end is called the shockwave (Lighthill and Whitham (1955)). The propagation speed of the shockwave, $G(t)$, is calculated by the following formula (see Appendix A for derivation):

$$G(t) = \frac{C_b - Q(t)}{K_J(C_b) - K_N(Q(t))} \tag{8}$$

[2] In this paper, *Q(t)* is also called “input flow”.

It is apparent that when $Q(t) > C_b$, the sign of $G(t)$ is negative, and the shockwave moves upstream, opposite the direction of traffic flow. Put simply, for this case the length of the queue is growing. By the same logic, the shockwave moves downstream when $Q(t) < C_b$.

The length of the queue at t, given by $H(t)$, is calculated by

$$H(t) = \int_{t_a}^{t} -G(u)du \,, \tag{9}$$

where t_a is the time at which the traffic volume upstream exceeds C_b and the queue is initiated.

The dynamics of $H(t)$ is described by a differential equation,

$$\dot{H}(t) = -G(t), \tag{10a}$$

The above equation does not always apply. Since the length of the queue must be a non-negative number, it follows that

$$\dot{H}(t) = 0 \text{, if } \quad H(t) = 0 \text{ and } Q(t) < C_b \tag{10b}$$

Suppose that at time *t* vehicles enter the road with a flow rate equal to $Q(t)$, and they encounter the queue behind the bottleneck, the length of which is $J(t)$. Then travel time $T(t)$ is defined as follows

$$T(t) = \frac{J(t)}{V_J(C_b)} + \frac{L - J(t)}{V_N(Q(t))} \tag{11}$$

where L is the length of the roadway. The first term represents the time required to pass through the queue and the second term is the time to travel the remainder. The traffic speed within the queue is determined by $V_J(C_b)$, since traffic is jammed and volume is constrained by the bottleneck capacity, C_b (see Figure 2).

It should be noted that $J(t)$ shown in (11) is slightly different from $H(t)$ in (9).

The former is the queue length that the driver entering the road at *t* encounters, while the latter is the queue length at time *t*. There is a time lag between the time of entry and the time that the driver encounters the queue. The driver entering at time *t* reaches the queue's tail end at $t+\frac{L-J(t)}{V_N(Q(t))}$. Thus the following relation holds

$$J(t)=H\left(t+\frac{L-J(t)}{V_N(Q(t))}\right) \tag{12}$$

And the dynamics of queue length with respect to departure time is derived as follows[3]

$$\dot{J}(t)=\frac{C_b-Q(t)}{K_J(C_b)-C_b/V_N(Q(t))} \tag{13}$$

This treatment is the modification made to Mun (1999)[4].

Travel time variation is obtained by differentiating (11), and reduced to the following form.

---

[3] Differentiating (12) with respect to *t* and using the relation of (10a), we have

$$\dot{J}(t)=\dot{H}(t_N)\left\{1-\frac{\dot{J}(t)}{V_N(Q(t))}\right\}$$

$$=\frac{-G(t_N)}{1-\frac{G(t_N)}{V_N(Q(t))}}$$

where $t_N=t+\frac{L-J(t)}{V_N(Q(t))}$. Inserting (8) into the above yields (13).

[4] Mun (1999) does not distinguish between $H(t)$ and $J(t)$, but describes the dynamics of queue length as $\dot{J}(t)=-G(t)$.

$$\dot{T}(t) = \frac{\partial T}{\partial Q}\dot{Q}(t) + \frac{\partial T}{\partial J}\dot{J}(t)$$

$$= -\frac{L - J(t)}{\{V_N(Q(t))\}^2}\frac{dV_N}{dQ}\dot{Q}(t) + \left(\frac{1}{V_J(C_b)} - \frac{1}{V_N(Q(t))}\right)\dot{J}(t)$$

$$= -\frac{L - J(t)}{\{V_N(Q(t))\}^2}\frac{dV_N}{dQ}\dot{Q}(t) + \frac{Q(t) - C_b}{C_b} \tag{14}$$

The last expression is obtained by substituting (13) for $\dot{J}(t)$ in the second term of the RHS. When there is no queue, the second term on the RHS disappears and (14) coincides with (4). In this case, the model is equivalent to Henderson's model. On the other hand, when there is a queue, the second term on the RHS of (14) is exactly same as (7). In other words, evolution of queuing time by the Vickrey-type bottleneck model is consistent with that by the model based on the physics of traffic flow with shockwave. Furthermore, if the speed in the normal flow is independent of the flow rate, i.e., $\frac{dV_N}{dQ} = 0$, the first term disappears and the model coincides with the pure bottleneck model. Therefore the present model includes Henderson's model and the Vickrey type bottleneck model as special cases.

4. Application to Peak-load Problem for the Morning Commute

The peak-load problem has been analyzed, *inter alia*, by Vickrey (1969), Henderson (1985) and Arnott, de Palma and Lindsey (1990). Henderson (1985) analyzes the peak-load problem based on the static model of traffic flow as in Section 2-1. His model predicts that road users are worse off by introducing a peak-load toll. On the other hand, Arnott et al. (1990) apply the Vickrey type bottleneck model as in Section 2-2, and show that the optimal peak-load toll improves social welfare without affecting

the level of private welfare: road users are well off with and without the toll. Mun (1999) uses the model that unifies two types of formulations, and shows that road users may be either better off or worse off by tolling. This section uses the model presented in the Section 3, and repeats the analysis in Mun (1999), to examine whether the reformulation affects the above result.

4-1 No-toll equilibrium

Suppose that every morning a fixed number, *N*, of identical individuals commute from home to the office, driving along a single lane road as in Figure 2. Nodes A and B in the figure are, respectively, the residential area and CBD. Each commuter chooses the departure time so as to minimize commuting cost, which is the sum of travel time cost and scheduling cost. The scheduling cost is incurred by arriving at the office earlier or later than the specified work start time; in the early arrival stage, it is the opportunity cost of waiting time before work, while in the late arrival stage it is the penalty of late arrival. It is assumed that all commuters have the same work start time, $\hat{t}$, as in Arnott, et al. (1990). Then the commuting cost for an individual departing home at *t* is specified as

$$C^1(t) = \alpha T(t) + \beta(\hat{t} - t - T(t)), \qquad if \;\; t + T(t) \le \hat{t} \tag{15a}$$

$$C^2(t) = \alpha T(t) + \gamma(t + T(t) - \hat{t}), \qquad if \;\; t + T(t) > \hat{t} \tag{15b}$$

where $C^1(t)$ and $C^2(t)$ are the commuting costs for individuals arriving earlier and later than the work start time, respectively. α is the monetary value of travel time, β and γ are respectively the values of time early and late. Empirical studies suggest $\beta < \alpha < \gamma$ (see, e.g., Small (1982)). The first term in the RHS of each equation above is the travel time cost and the second term is the scheduling cost.

Equilibrium obtains when no commuter has an incentive to change his departure time. Since commuters are identical, commuting cost must be the same at all times that departures occur. Hence the equilibrium conditions can be stated as follows

$$C^1(t) = C^*, \qquad if \quad t_1 \le t \le \hat{t} - T(\hat{t}) \tag{16a}$$

$$C^2(t) = C^*, \qquad if \quad \hat{t} - T(\hat{t}) < t \le t_2 \tag{16b}$$

$$\int_{t_1}^{t_2} Q(t)dt = N \tag{16c}$$

where t_1 and t_2 are the times at which the first and last commuters depart home, respectively. In other words, the rush hour starts at t_1 and ends at t_2. C^* is the equilibrium commuting cost that is constant during the rush hour. Note that *Q(t)* is the departure rate from the residential area at time *t*.

The following equations are obtained by differentiating (16a)(16b).

$$\dot{T}(t) = \begin{cases} \dfrac{\beta}{\alpha - \beta}, & if \quad t_1 \le t \le \hat{t} - T(\hat{t}) \\ \dfrac{\gamma}{\alpha + \gamma}, & if \quad \hat{t} - T(\hat{t}) < t \le t_2 \end{cases} \tag{17}$$

where $\dot{T}(t)$ is given by (14). The equilibrium pattern of departure *Q(t)* is obtained by solving the above differential equation.

4-2 Optimal Road Use and Peak-Load Toll

Socially optimal road use is achieved when the total commuting cost is minimized. The objective function to be minimized is defined as follows

$$W = \int_{t_1}^{\tilde{t}} C^1(t)Q(t)dt + \int_{\tilde{t}}^{t_2} C^2(t)Q(t)dt \tag{18}$$

where $C^1(t), C^2(t)$ are defined by Eq.(15a), (15b), respectively. $\tilde{t}$ is the departure time of a commuter who incurs no scheduling cost by arriving at $\hat{t}$.

The problem is formulated as a two-stage optimal control problem (Tomiyama (1985)) in which early ($t_1 \le t \le \tilde{t}$) and late ($\tilde{t} < t \le t_2$) arrival stages are integrated. The control variable is the departure rate, $Q(t)$; state variables are queue length, $J(t)$. $\tilde{t}$ is called the switching time, a control parameter in the problem.

The optimality conditions are the same as those presented in Mun (1999), except that the state equation of $J(t)$ is described by (13). Therefore we omit the explanation of the optimality conditions.

The peak-load toll to decentralize the optimal road use is derived. For the toll to be compatible with the decentralized departure choice of commuters, the following equations must hold at each time:

$$C^1(t) + \tau(t) = C^O \qquad for \;\; t_1 \le t \le \tilde{t} \tag{19a}$$

$$C^2(t) + \tau(t) = C^O \qquad for \;\; \tilde{t} < t \le t_2 \tag{19b}$$

where C^o is the private commuting cost (inclusive of toll) in the equilibrium with optimal peak-load toll. It is postulated that tolls should take non-negative values.

By comparing the optimal condition with (19), we obtain the optimal toll schedule as follows:

$$\tau(t) = \begin{cases} 0, & for\ t \le t_1 \\ (\alpha - \beta)E(Q(t)), & for\ t_1 < t \le t_1^{'} \\ (\alpha - \beta)E(C_b) + \beta(t - t_1^{'}), & for\ t_1^{'} \le t \le \tilde{t} \\ (\alpha + \gamma)E(C_b) + \gamma(t_2^{'} - t), & for\ \tilde{t} \le t \le t_2^{'} \\ (\alpha + \gamma)E(Q(t)), & for\ t_2^{'} \le t < t_2 \\ 0, & for\ t_2 \le t \end{cases} \tag{20}$$

where $E(Q(t)) = Q(t)\dfrac{\partial T(t)}{\partial V_N}\dfrac{dV_N}{dQ(t)}$, that is, the externality effect among vehicles in the normal (unjammed) flow when traffic volume is equal to *Q(t)*, which is called *static externality* hereafter. For the time period, $t_1 < t \le t_1'$ or $t_2' \le t < t_2$, the departure rate

is smaller than the bottleneck capacity, and the toll is equal to the value of static externality. This rule is essentially equivalent to that derived in Henderson (1985). On the other hand, for $t_1' \le t \le t_2'$, the value of $E(C_b)$ is constant, consequently the toll profile is linear. This property is similar to that derived by Arnott, de Palma and Lindsey (1990) using a Vickrey-type bottleneck model. The optimal peak-load toll derived in the present paper includes both Henderson-type and Vickrey-type as special cases. The Henderson-type toll internalizes static externality, while the Vickrey type toll eliminates queue. Note that $\tau(t)$ have non-negative values, and no commuter has an incentive to change departure time to $t < t_1$ or $t > t_2$.

4-3 Simulations

Specifications of the functional form and parameter values are the same as in Mun (1999), which are presented in Appendix B.

Table 1 summarizes equilibrium and optimal solutions for two cases, Case I ($N = 1500$) and Case II ($N = 4000$). In the table, time variables are normalized by setting the work start time to zero: for example, rush hour starting time in Case I, -56.16, means that the first commuter departs 56.16 minutes earlier than the work start time. In both cases, the rush hours at optimum start earlier and end later than those at equilibrium. In other words, rush hours at optimum are longer than at equilibrium. As discussed in Mun (1999), the length of the rush hour has important implications for commuter's private welfare: private welfare is higher as the rush hour is shorter. Private welfare is represented by private costs in the table, which correspond to C^* in (16) for equilibrium and C^O in (19) for optimum. The table shows that, for both Cases I and II, private commuting costs at optimum are larger than at equilibrium. The above results differ from those obtained by Mun (1999), which shows that, for the same parameter

values as in Case II, private commuting cost under tolling is smaller than that under no-toll equilibrium.

Figure 3 plots private welfare and social welfare for various values of N. For a fixed value of C_b, the levels of congestion become heavier for larger N. Recall that private welfare is represented by C^* and C^O. Social welfare is represented by the total commuting cost, defined as W in Eq. (18a). W^O and W^* are the total commuting cost at optimum and equilibrium, respectively. The figure shows that W^O/W^* decreases with the total number of commuters. This implies that social welfare gains from introducing a peak-load toll are larger when the traffic congestion in no-toll equilibrium is heavier.

Table 1

The figure shows that C^O/C^* always exceeds unity, suggesting that private welfare is always decreased by imposing a toll[5]. Although I examined various combinations of parameters to check whether different results are obtained, there was no case in which private welfare is improved by the introduction of a toll. Thus the result of Mun (1999) regarding the welfare effect of tolling is not valid under the present model. Difference between the present model and the previous one is the formulation of state equation (13). It is observed that the difference is quantitatively subtle, but qualitatively significant.

[5] Verhoef (2002) reports similar results based on a car-following model in the same setting as in the present paper, which incorporate bottleneck and flow congestion. Note that the car-following model is a micro foundation of a traffic flow model, since Eqs. (1), (2) are derived as the stationary state of the car-following process.

Recall that the present model combines Henderson-type and Vickrey-type models, and in the Henderson-type model private welfare is always lowered by a toll while in the Vickrey-type it is unchanged. It is reasonable that combining the results of two types yields the result that private welfare is decreased. It is also observed that C^O/C^* is increasing when N is smaller than 1000, and begins to decrease for larger N. For a small value of N, congestion within normal flow (static externality) is dominant, since the input flow rate hardly exceeds the bottleneck capacity. In this case, the role of the toll is largely to internalize the static externality as in Henderson (1985), and imposing this type of toll tends to decrease private welfare. On the other hand, when N is large (traffic jam is severe in the no-toll case), the loss of private welfare from tolling is relatively small. This is due to the dominance of the effect of the Vickrey-type toll that eliminates the queue without changing the level of private welfare. This result also suggests that acceptability barriers to road pricing may be easier to surmount in areas where congestion is either light or heavy compared to areas in which it is of medium intensity.

Figure 3

5. Conclusion

This paper presents a dynamic model of traffic congestion based on shockwave theory, and applies it to the peak-load problem for the morning commute.

The model in Mun (1999) is reformulated by relaxing the assumption regarding the propagation of a queue. This reformulation yields a clear relationship between the present model and earlier models of traffic congestion. In particular, the dynamic

equation describing travel time evolution is derived as the combination of Henderson-type flow model and Vickrey type bottleneck model.

Numerical simulations show that the private welfare is always reduced by peak-load toll. This result is different from that in Mun (1999), but plausible under the assumptions of the present model. It is also shown that introducing a peak-load toll attains a larger gain in social welfare with a smaller loss of private welfare, as peak period congestion is heavier.

The above result should not be considered as the final answer. It might be the case in a more realistic model that tolling improves private welfare. It is commonly observed that the flow rate at the bottleneck is not constant but decreasing with queue length. If the traffic volume passing through the bottleneck without a queue is larger than that with a queue, tolling to eliminate the queue may improve private welfare. Yang and Huang (1997) attempt to deal with this problem by assuming that the bottleneck capacity is a decreasing function of the queue length. Rather than assuming the capacity as an exogenously given function, we should develop a model that describes endogenously the process of decreasing the flow rate through interaction between traffic flows in the queue and downstream bottleneck. The car following model (e.g., Verhoef (2002)) may be a promising tool to deal with this problem.

Appendix A Derivation of (8)

Suppose that, at time *t*, the tail end of the queue is located at x_g and moving with the speed equal to G, as illustrated in Figure 2. Traffic volume and density in the upstream of x_g are Q and K_N, respectively. And traffic volume and density in the

downstream of x_g are C_b and K_J, respectively. While the tail end moves from x_g to $x_g + G$ between t and t+1, the number of vehicles in the road section $[x_g, x_g + G]$ changes from $K_J G$ to $K_N G$. For the same time period, C_b vehicles flow out and Q vehicles flow into that section. Thus $Q - C_b = K_N G - K_J G$ holds, then we have (8).

Appendix B Specification of functional form and parameter values for simulations

The density-speed relationship, $f(K)$ in Eq.(1), is specified by the following Greenshield formula:

$$f(K) = V_f \left\{ 1 - \left(K / k_u \right) \right\} \quad .$$

where V_f is the "free speed" and k_u is the maximum density at which the speed equals zero, namely the traffic comes to a standstill.

For the simulations, the road conditions are given as follows:

Length of roadway:	L = 20 (km)
Capacity of the main part of the road :	C_a = 3600 (Vehicles/hour) = 60 (Vehicles/min)
Capacity of the bottleneck:	C_b = 2400 (Vehicles/hour) = 40 (Vehicles/min)
Free speed:	V_f = 60 (km/hour)
Maximum density:	k_u =180 (Vehicles/km).

These values are derived from the empirical study of a Japanese Expressway by Makigami et al. (1984). The value of travel time, α, is assumed to be 2000 Yen/hour, which also comes from another Japanese empirical study. Regarding the value of scheduling cost, in view of the estimates by Small (1982), we assume that

$\beta / \alpha = 0.4, \gamma / \alpha = 2.1$.[6]

References

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

Chu, X., 1993, Comment on "Dynamic user equilibrium departure times and route choice on idealized traffic arterials", Mimeo, University of California, Irvine.

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

Else, P.K., 1981, A reformulation of the theory of optimal congestion taxes, *Journal of Transport Economics and policy*, XV, 217-232 (1981).

Evans, A.W., 1992, Road congestion: The diagrammatic analysis, *Journal of Political Economy*, 100, 211-217 .

Henderson, J.V., 1981, The economics of staggered work hours, *Journal of Urban Economics*, 9, 349-364.

Lighthill, M.H. and G.B. Whitham, 1955, On kinematic waves, II: a theory of traffic flow on long crowded roads, *Proceedings of the Royal Society* (London), A229, 317-345.

Makigami, Y., T. Nakanishi and S. Kim, 1984, Simulation model applied to Japanese expressway, *ASCE Journal of Transportation Engineering*, 110, 94-111.

Matsumoto, S., 1988, Travel time variability and target time of arrival on commuter journeys, working paper, Transport Studies Unit, Oxford University.

[6] Empirical studies on scheduling cost for Japan are quite rare. Matsumoto (1988) is an exception: he estimated the disutility of scheduling delay based on the departure choice model using Japanese data. However, the monetary value of the disutility was not obtained in his study.

Mun, S., 1994, Traffic jams and the congestion toll, *Transportation Research B*, Vol.28B, pp.365-375.

Mun, S., 1999, Peak-load pricing of a bottleneck with traffic jam, *Journal of Urban Economics*, 46, 323-349.

Small, K., 1982, The scheduling of consumer activities: Work trips, *American Economic Review*, 72, 467-479.

Small, K. and X. Chu, 2003, Hypercongestion, *Journal of Transport Economics and Policy*, 37, 319-52.

Verhoef, E., 2001, An integrated dynamic model of road traffic congestion based on simple car-following theory, *Journal of Urban Economics*, 49, 505-542.

Verhoef, E., 2002, Inside the Queue: Hypercongestion and Road Pricing in a Continuous Time –Continuous Place Model of Traffic Congestion, *Journal of Urban Economics*, 54, 531-65.

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

Walters, A.A., 1961, The theory and measurement of private and social cost of highway congestion, *Econometrica*, Vol.29, 676-699.

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

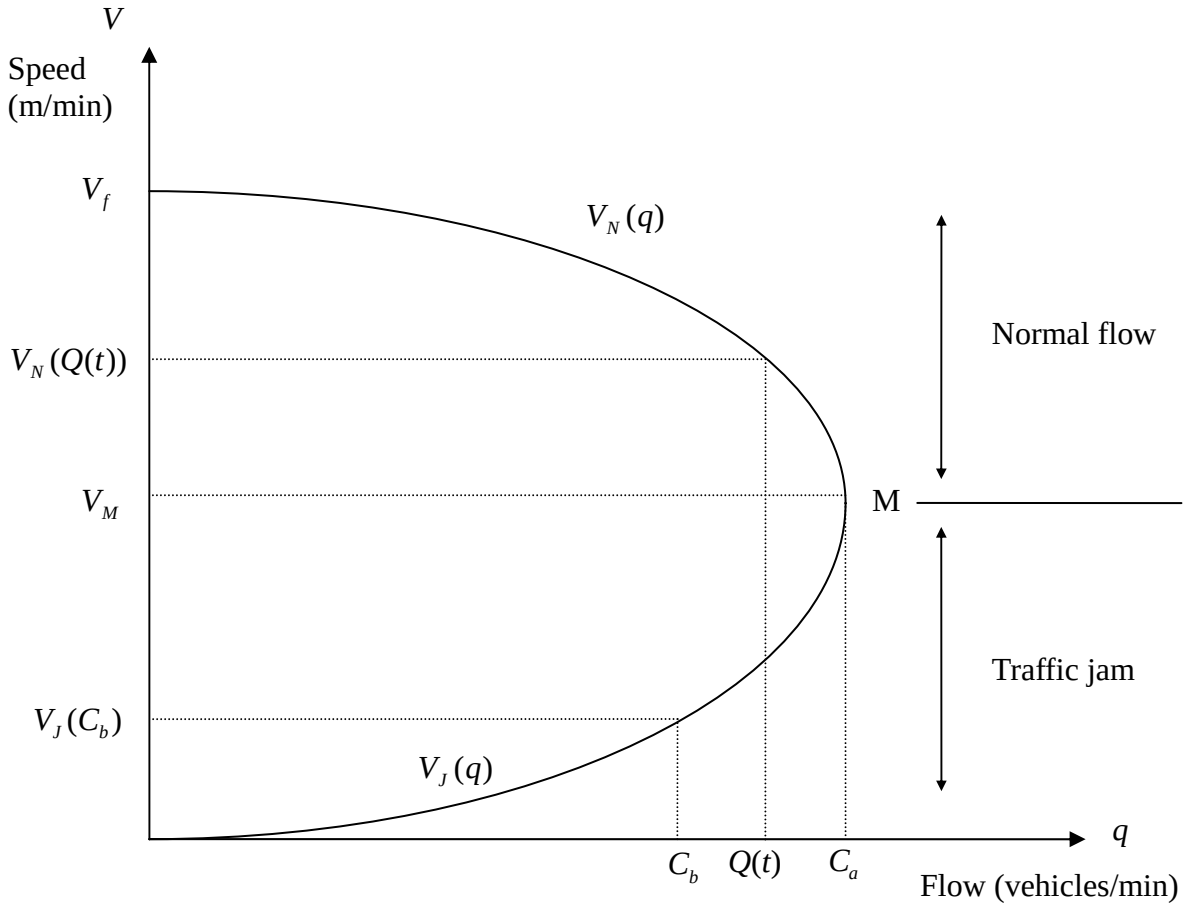

Figure 1 Relation between flow and speed

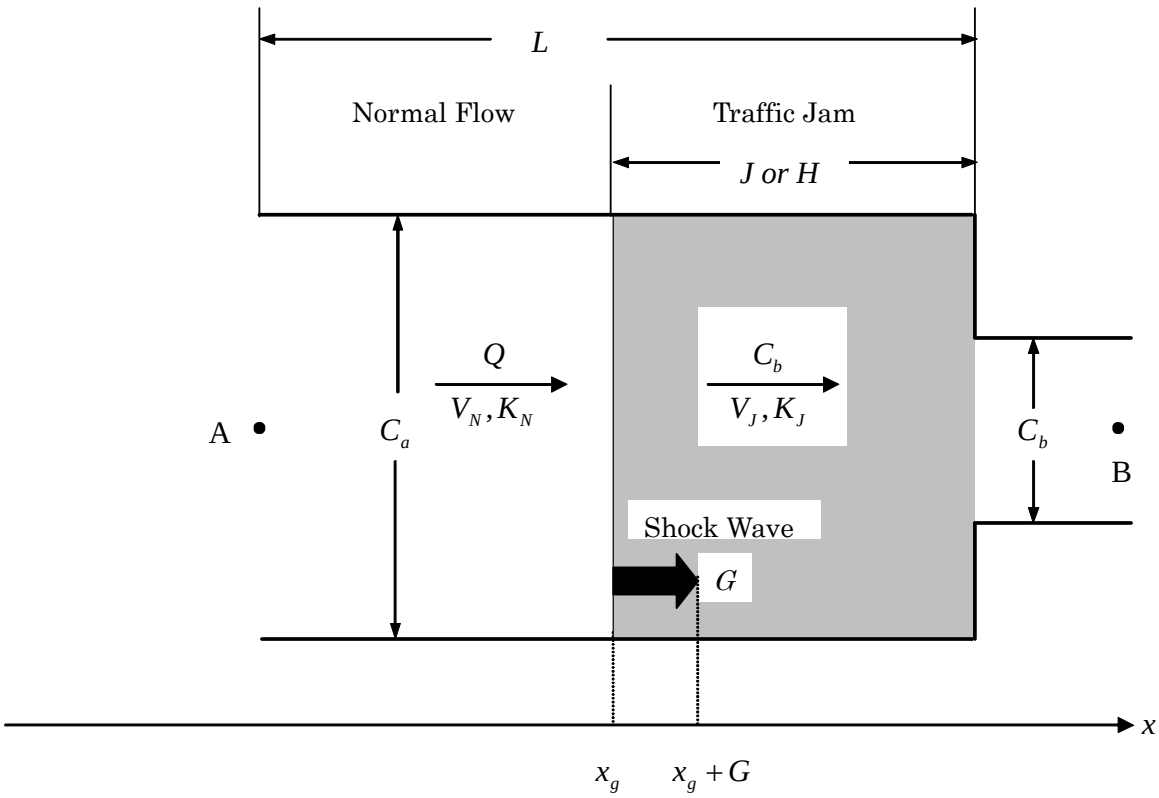

Figure 2 State of traffic on the roadway

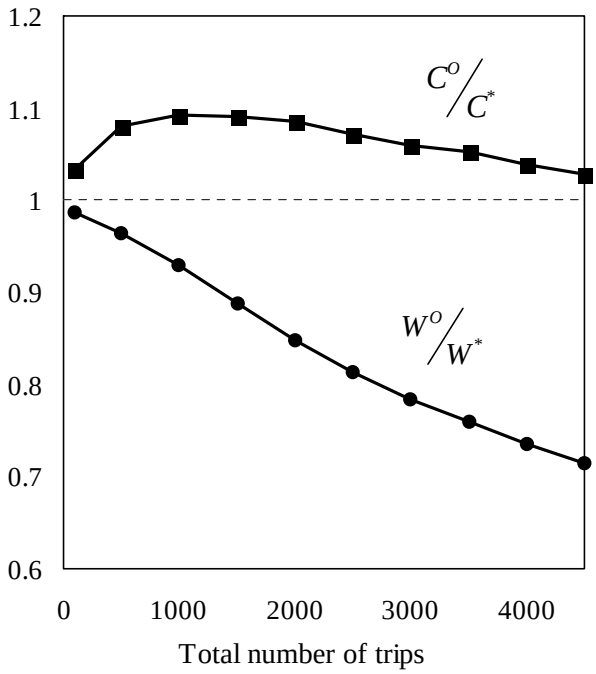

Figure 3 Social welfare and private welfare for various levels of congestion

Table 1 Comparison of equilibrium and optimal solutions

	Case I		Case II	
	Equilibrium	Optimum	Equilibrium	Optimum
Rush hour start time (min)	-56.16	-65.37	-112.13	-117.37
Rush hour ending time (min)	-13.11	-11.36	-2.45	-1.45
Switching time (min)	-34.46	-25.36	-56.85	-25.36
Peak queue length (m)	2062.6	0.0	8110.3	0.0
Peak toll (Yen)	0.0	426.3	0.0	1119.6
Average toll (Yen)	0.0	250.8	0.0	574.5
Private cost (Yen)	1148.8	1271.6	1895.1	1964.9
(Optimum/Equilibrium)	1.107		1.037	
Total commuting cost (Yen)	1723129.6	1531203	7580290.0	5561524.0
(Optimum/Equilibrium)	0.889		0.734	
Total scheduling cost (Yen)	278006.7	312958.2	2022648.4	2244351.5
Total travel time cost (Yen)	1445123	1218244.7	5557641.5	3317172.5