WORKING PAPER NO. 11

PRICING BOTTLENECK CONGESTION AND MODAL SPLIT

by

Takatoshi Tabuchi

Faculty of Economics
Kyoto University

Faculty of Economics,
Kyoto University,
Kyoto, 606 JAPAN

WORKING PAPER NO. 11

PRICING BOTTLENECK CONGESTION AND MODAL SPLIT

by

Takatoshi Tabuchi

Faculty of Economics
Kyoto University

March 1991

PRICING BOTTLENECK CONGESTION AND MODAL SPLIT

by

Takatoshi Tabuchi

Faculty of Economics, Kyoto University

Sakyo-ku, Kyoto 606, Japan

March 1991

ABSTRACT

This paper analyzes commuting congestion when there is a mass transit and a road with a bottleneck between a residential area and a workplace. We investigate the optimality and efficiency of several road pricing régimes and obtain simple and practical rules to attain the social optimum. We then make a welfare comparison between these road toll régimes, and show that the road tolls are effective especially in case of heavy bottleneck congestion.

1. Introduction

It is often experienced in large cities that driving into a CBD by car takes much more time than driving out of a CBD during the rush hour even when the number of auto drivers into a CBD is less than that out of a CBD. This is due to the radial structure of the road network, i.e., the number of roads and lanes decreases as one approaches to a CBD. Roughly speaking, such a road network is compared to a road with a bottleneck like a sandglass, and cars are compared to the sands.

In a sandglass example, the speed of the falling sands is much faster than that of the accumulated sands although each speed is fixed and independent of the amount of the sands. In case of urban traffic, the speed on an uncongested road is much faster than that at the congested bottleneck though each speed is constant and almost independent of the number of cars. Consequently, the commuting time is determined solely by the length of the queue, and hence the number of cars during the rush hour. If a commuter wants to avoid such a long queue to reduce the time to commute, she has to arrive at the CBD much earlier or later than the start time of work. This is also considered to be a cost incurred by the commuter.

Vickrey (1969) first developed a model of such an endogenous departure time with bottleneck congestion, which is fully investigated by Arnott, de Palma and Lindsey, hereafter ADL, (1990) and Braid (1989), among others. This paper extends the ADL model of bottleneck congestion by introducing an alternative commuting mode of a mass transit which has no congestion but a higher fixed cost of commuting. In brief, we consider the situation that commuters can not only choose the departure time from home, but also select the transport mode between the road with a bottleneck and the mass transit. Using a road toll, we derive a demand function for the road service. We then investigate the optimality and efficiency of several road pricing régimes and obtain simple rules to attain the social optimum that minimizes the total commuting cost, which consists of the travel time cost, the fixed travel cost, and the schedule delay cost.

In section 2, after depicting the general setting of the model, we examine three kinds of equilibria in each subsection: the no-toll equilibrium in subsection 2.1, the uniform toll equilibrium in subsection 2.2, and the fine toll equilibrium in subsection 2.3. We make a welfare comparison in section 3. Section 4 concludes the paper.

2. The Model

2.1 No-Toll Equilibrium

Between a residential area and a central business district, there are two commuting routes: (a) a road with a bottleneck and (b) a railway. Automobiles are used in the road while a mass transit is utilized on the railway. The number of users is denoted by N_a and N_b respectively. Since the mass transit arrives on time (t^*), its users do not have to pay the schedule delay costs while automobile commuters if congested have to incur the costs. Due to the physical constraint of the bottleneck capacity, some of automobile commuters necessarily arrive earlier or later than t^* unless N_a is less than the capacity.

When the road is free from congestion (i.e., the rate of arrival to the bottleneck is less than its capacity), the commuting time is assumed to be a constant value T_o. If the arrival rate exceeds the capacity, a queue develops. The commuting time is then written by

$$T(t) = T_o + Q(t)/s, \tag{1}$$

where t is the departure time, Q(t) is the queue length measured by the number of automobiles, and s is the bottleneck capacity. Let r(t) be the rate of departure from home, then the queue length is given by

$$Q(t) = \int_{t_o}^{t} r(u)du - s(t-t_o) \qquad \text{for} \quad Q(t) > 0, \tag{2}$$

where t_o is the most recent time at which there was no queue.

In addition to the travel time costs, the auto commuters have to incur the costs of schedule delay as mentioned above. Following Vickrey (1969) and ADL (1990a), we assume that the cost is proportional to the travel time and to the schedule delay. The total cost of auto commute is then expressed as

$$C_a = \alpha(\text{travel time}) + \beta \max(\text{time early},0) + \gamma \max(\text{time late},0) + \text{toll},$$

where the Greek letters are shadow prices of time, which are considered to take different values.[1] Specifically, the total cost is given by

$$\begin{aligned} C_a &= \alpha T(t) + \beta[t^*-t-T(t)] && \text{for} \quad t\in[t_o,\tilde{t}), \\ &= \alpha T(t) + \gamma[t-t^*+T(t)] && \text{for} \quad t\in(\tilde{t},t_1], \end{aligned} \tag{3}$$

where t_1 is the time at which the queue ends, and $\tilde{t}$ is the departure time at which an individual arrives at work on time t^*, i.e., $\tilde{t} + T(\tilde{t}) = t^*$.

In equilibrium, every auto commuter is unable to find a departure time which reduces her total cost. In other words, C_a in (3) should be constant for all t. Hence, using (1) and (3), r(t) in (2) is solved as:

$$\begin{aligned} r(t) &= \frac{\alpha}{\alpha-\beta}\, s && \text{for} \quad t\in[t_o,\tilde{t}), \\ &= \frac{\alpha}{\alpha+\gamma}\, s && \text{for} \quad t\in(\tilde{t},t_1], \end{aligned} \tag{4}$$

Substituting (4) into (2), the queue length is given by

$$\begin{aligned} Q(t) &= \frac{\beta s}{\alpha-\beta}(t-t_o) && \text{for} \quad t\in[t_o,\tilde{t}), \\ &= \frac{\beta s}{\alpha-\beta}(\tilde{t}-t_o) - \frac{\gamma s}{\alpha+\gamma}(t-\tilde{t}) && \text{for} \quad t\in(\tilde{t},t_1], \end{aligned} \tag{5}$$

Equation (4) indicates that the arrival rate is piecewise constant and equation (5) shows that a queue develops linearly and dissipates linearly.

Now, since

$$\tilde{t} + T(\tilde{t}) = t^*,$$

$$Q(t_1) = 0, \qquad \text{and}$$

$$\int_{t_0}^{t_1} r(u)du = N_a,$$

we can determine the three unknowns as follows:

$$t_0 = t^* - T_0 - \frac{\delta N_a}{\beta s}, \quad t_1 = t^* - T_0 + \frac{\delta N_a}{\gamma s}, \quad \tilde{t} = t^* - T_0 - \frac{\delta N_a}{\alpha s}, \tag{6}$$

where $\delta \equiv \beta\gamma/(\beta+\gamma)$. Thus, the total cost per capita is obtained as

$$C_a = \frac{\delta N_a}{s} + \alpha T_0. \tag{7}$$

The constancy of C_a implies that no individual is able to decrease her total cost by changing her departure time t.

Let us next consider the mass transit commute. For the sake of simplicity, we assume that the total cost by mass transit C_b is constant against the number of its users N_b. That is, we are assuming that increased passengers of the mass transit would be accommodated by increasing the number of vehicles without further cost.[2]

Since no individual can decrease her total cost by altering her trip mode,

$$C_a \leq C_b, \tag{8}$$

should always hold in equilibrium. From (7) and (8), it is straightforward that the equilibrium numbers of users are:

$$\begin{aligned} (N_a^e, N_b^e) &= (N, 0) && \text{for} \quad N \leq \bar{N}_a, \\ &= (\bar{N}_a, N - \bar{N}_a) && \text{for} \quad N > \bar{N}_a, \end{aligned} \tag{9}$$

where $N[\equiv N_a + N_b]$ is the total number of commuters in the city, and the critical value $\bar{N}_a[\equiv s(C_b - \alpha T_0)/\delta]$ is the value that makes the equality in (8) hold. For future reference, we define that the city size is:

small *if* $N \leq \bar{N}_a/2$;

medium *if* $N \in (\bar{N}_a/2, \bar{N}_a)$;

large *if* $N \geq \bar{N}_a$.

Without little loss of generality, we assume $\bar{N}_a > 0$ or equivalently $C_b > \alpha T_o$ so as to guarantee the need for the road. Otherwise, everyone uses the mass transit under any nonnegative road toll. Finally, we define the total social cost in this city under the no-toll equilibrium as

$$TC^e \equiv C_a N_a + C_b N_b. \tag{10}$$

In the next two subsections, we will introduce two kinds of road toll régimes: a uniform toll and a time-varying fine toll. Feasibility of these tolls rests on social consensus in addition to their technical conditions.

2.2 Second Best by the Uniform Toll

Suppose the public authority is able to levy a uniform toll τ^u from road users. This régime would be very common to most countries under the current level of technology. The total cost per road user (7) is then modified to

$$C_a = \frac{\delta N_a}{s} + \alpha T_o + \tau^u \tag{11}$$

while that of a mass transit user C_b remains invariable with respect to its number of users N_b. Equations (8) and (11) establish an equilibrium given a level of the uniform toll.

Since commuting is a must, the trip for commute itself is perfectly inelastic. However, the introduction of the mass transit lets the road commute elastic with respect to the road toll τ^u. It should be noticed that whereas ADL (1987) and Braid (1989) assume an elastic demand *a priori*, the demand function in this model is endogenously determined because a shift to the mass transit commute takes place here by the toll τ^u.

Using (8) and (11), the demand for the road service N_a is expressed as a function of its price τ^u:

$$N_a(\tau^u) = 0 \qquad \text{for} \quad \tau^u \geq C_b,$$

$$= \max\{N, \bar{N}_a(1-\tau^u/C_b)\} \qquad \text{for} \quad \tau^u < C_b.$$

Obviously, the demand function is piecewise linear and weakly monotone decreasing in the road toll τ^u.

Suppose the toll revenue is assumed to be equally redistributed to both users, then the total social cost in this city is redefined by

$$TC^u \equiv C_a N_a + C_b N_b - \tau^u N_a. \tag{12}$$

The public authority minimizes (12) with respect to τ^u subject to (8) and (11), which also determines the optimum number of users of the road and the mass transit. After some simple calculations, we obtain the optimal pricing of the uniform toll as follows:

$$\tau^u = 0 \qquad \text{for} \quad N \leq \bar{N}_a/2,$$

$$= \frac{C_b - \alpha T_o}{2} \qquad \text{for} \quad N > \bar{N}_a/2. \tag{13}$$

Put it in another way,

Proposition 1

In small size of cities, no toll should be levied. In medium or large cities, a fixed amount of the uniform toll (13) should be imposed.

From the uniform toll pricing (13), the corresponding distribution of users of the road and mass transit is shown to be:

$$(N_a^u, N_b^u) = (N, 0) \qquad \text{for} \quad N \leq \bar{N}_a/2,$$

$$= (\bar{N}_a/2, N-\bar{N}_a/2) \qquad \text{for} \quad N > \bar{N}_a/2. \tag{14}$$

Comparing the second-best distribution (14) with the no-toll equilibrium distribution (9), we can draw the following two propositions.

Proposition 2

In medium size of cities, the public authority should construct the mass transit if the cost of construction is low enough.

This corresponds to the case of $N \in (\bar{N}_a/2, \bar{N}_a)$, where $N_b^e=0$ (i.e., the mass transit is not demanded under no-toll equilibrium), but $N_b^u>0$ (i.e., the mass transit is needed under the uniform toll pricing). It should be noted that without exercising the road toll, the mass transit is not voluntarily built by a private sector since no one uses the mass transit when N is less than $\bar{N}_a$. Needless to say, such a government intervention is justified because the bottleneck congestion creates the negative externality.

Suppose the city is in the growing stage, then it is optimal to construct a mass transit when N becomes $\bar{N}_a/2$ and impose the uniform toll. This transit may be run by a private sector only after N exceeds $\bar{N}_a$.

The next proposition deals with the case of large city size, where $N_a^u=N_a^e/2$ always holds.

Proposition 3

If there are some mass transit users under no-toll equilibrium, then a uniform toll should be levied such that the number of road users is reduced to half.

Since mass transits are utilized in almost every large cities, the road toll (or a subsidy for mass transits) is necessary in those cities. The level of the toll should be adjusted so as to halve the number of road users. We would like to emphasize that the rule of Proposition 3 is so simple that it can be easily applied without estimating the set of parameters, α, β, γ, s, T_o and C_b when the uniform toll is the only possible policy means.

In conclusion, by employing the above mentioned three propositions, we can sum up the uniform toll pricing scheme in the following way:

1. *Suppose there exists no-toll pricing in every city. Observe first that the number of automobile commuters* $\bar{N}_a$ *in large cities* ($N>\bar{N}_a$), *where the mass transit pays.*
2. *Determine the level of the uniform toll* τ_o *which reduces* $\bar{N}_a$ *by half in the large cities by trial and error.*
3. *Using the value of* $\bar{N}_a$, *identify medium cities* $N\in(\bar{N}_a/2,\bar{N}_a)$. *Construct a mass transit there if its cost is small enough.*
4. *Apply the same level of the uniform toll* τ_o *to all of medium and large cities* ($N>\bar{N}_a/2$).

It seems obvious that this pricing scheme is not a mere theory. It is a feasible and practical scheme under limited data on urban transportation. Notice that the scheme is applicable to every city.

2.3 Social Optimum by the Fine Toll

So far, the road toll has been assumed to be constant. However, if pricing technology allows the public authority to exercise a time-dependent fine toll, then the total social cost will be reduced further. Specifically, a time-varying road toll can eliminate any queue although it does not reduce the schedule delay costs. Employing the result by ADL (1990), the optimal fine toll for the road with a bottleneck is expressed as:

$$
\begin{aligned}
\tau(x,t) &= 0 && \text{for } t \le t_o, \\
&= x - \beta(t^* - t - T_o) && \text{for } t \in (t_o, \tilde{t}), \\
&= x - \gamma(t + T_o - t^*) && \text{for } t \in [\tilde{t}, t_1), \\
&= 0 && \text{for } t \ge t_1,
\end{aligned} \tag{15}
$$

where $x \le \delta N_a/s$. This condition assures that commuting by automobile occurs only at the period $[t_o, t_1]$.

Equation (15) implies that the fine toll is piecewise linear, and should be collected in proportion to the queue length (5) of no-toll equilibrium. If the set of the parameters is difficult to estimate in practice, the fine toll should be adjusted such that the arrival rate $r(t)$ equals the capacity s, where the queue is about to vanish.

Similar to the previous subsections, the total social cost under the fine toll is rewritten by

$$
TC^f = C_b N - \frac{xN_a}{2} = C_b N - \frac{sx^2}{2\delta}, \tag{16}
$$

which should be minimized with respect to x by the public authority. Note that the second equality is due to $\tau(x, t_o) = 0$, or $x \le \delta N_a/s$. The binding constraint is (8), which is $C_a \le C_b$, or $x \le \delta \bar{N}_a/s$.

Clearly, the minimizer of (16) is obtained as $x = \max\{\delta \bar{N}_a/s, \delta N/s\}$. As a result, the optimal distribution of users of the road and mass transit is

$$
\begin{aligned}
(N_a^f, N_b^f) &= (N, 0) && \text{for } N < \bar{N}_a, \\
&= (\bar{N}_a, N - \bar{N}_a) && \text{for } N \ge \bar{N}_a.
\end{aligned} \tag{17}
$$

Note that in (17), $N < \bar{N}_a$ corresponds to $x = \delta N/s$ and $C_a < C_b$ (only the road is used), and $N \ge \bar{N}_a$ corresponds to $x = \delta \bar{N}_a/s$ and $C_a = C_b$ (both modes are used).

Comparing (17) with (9), we immediately witness that this social optimum distribution of the road and mass transit users coincides with the equilibrium distribution. In other words, *Proposition 4*

If a fine toll (15) is feasible, then it should be levied in proportion to the queue length, but it should not change the number of road users.

The distributions of between the road and mass transit users under the above three kinds of toll régimes are summarized in Table 1.

Pricing Régime	Small City $N \le \bar{N}_a/2$	Medium City $\bar{N}_a/2 < N \le \bar{N}_a$	Large City $\bar{N}_a < N$
No Toll	$(N,0)$	$(N,0)$	$(\bar{N}_a, N-\bar{N}_a)$
Uniform Toll	$(N,0)$	$(\bar{N}_a/2, N-\bar{N}_a/2)$	$(\bar{N}_a/2, N-\bar{N}_a/2)$
Fine Toll	$(N,0)$	$(N,0)$	$(\bar{N}_a, N-\bar{N}_a)$

Table 1 The Number of Road Users and The Number of Mass Transit Users Under Three Pricing Régimes [$\bar{N}_a \equiv (C_b - \alpha T_o)s/\delta$]

Road tolls are often criticized in that the right to use roads by everyone is infringed. If the above fine toll is technologically feasible, such criticism is misdirected. The fine toll does replace the commuting time cost at a queue by the toll which is to be equally redistributed to all commuters.

Notice that whereas the uniform toll eliminates a part of the queue by inducing some road users to convert to the mass transit, the fine toll can get rid of the whole queue by compelling road users to wait at home. As is demonstrated in the next section, the latter toll is, of course, superior.

3. Welfare Comparison

The total social costs per capita for three toll régimes are listed in Table 2 for comparison. It is observed that the total social cost under the fine toll is half the cost under the no-toll or under the uniform toll in small cities $N(\leq\bar{N}_a/2)$, but that the differences in the total social costs gradually diminish as city size N gets larger. It is also observed that in small cities the uniform toll is immaterial whereas the fine toll is not. Remember that the mass transit is not utilized in small cities.

Pricing Régime	Small City $N\leq\bar{N}_a/2$	Medium City $\bar{N}_a/2<N\leq\bar{N}_a$	Large City $\bar{N}_a<N$
No Toll	$C_b - \frac{\delta}{s}(\bar{N}_a-N)$	$C_b - \frac{\delta}{s}(\bar{N}_a-N)$	C_b
Uniform Toll	$C_b - \frac{\delta}{s}(\bar{N}_a-N)$	$C_b - \frac{\delta\bar{N}_a^2}{4sN}$	$C_b - \frac{\delta\bar{N}_a^2}{4sN}$
Fine Toll	$C_b - \frac{\delta}{s}(\bar{N}_a-\frac{N}{2})$	$C_b - \frac{\delta}{s}(\bar{N}_a-\frac{N}{2})$	$C_b - \frac{\delta\bar{N}_a^2}{2sN}$

Table 2 Total Social Costs Per Capita Under Three Pricing Régimes

In general, both the uniform and fine toll régimes are effective when commuters want to avoid the schedule delays (β, γ), the road capacity (s) is small, and/or the city size (N) is large. On the other hand, the régimes are useless when commuters do not care the schedule delays, there is enough number of road lanes, and/or the city size is small. Briefly speaking, the toll régimes are indispensable in highly congested cities, which is in accord with our intuition.

Following ADL (1990), let us next evaluate the relative efficiency of the uniform toll by the following formula:

$$\text{eff}^{u}. \equiv \frac{TC^{e}-TC^{u}}{TC^{e}-TC^{f}} .$$

Substituting the values in Table 2, we have

$$\begin{aligned} \text{eff}^{u}. &= 0 && \text{for } N \leq \bar{N}_a/2, \\ &= 2\left[1 - \frac{\bar{N}_a^2}{2N}\right]^2 && \text{for } N \in (\bar{N}_a/2, \bar{N}_a), \\ &= \frac{1}{2} && \text{for } N \geq \bar{N}_a. \end{aligned}$$

It follows from this that the relative efficiency of the uniform toll increases monotonically from 0 to 1/2 until the city size becomes $\bar{N}_a$, and becomes constant after that. It is also found that the upperbound of the relative efficiency of the uniform toll is 50%. It is hoped that the fine toll régime becomes technologically feasible.

4. Concluding Remarks

ADL's (1990) model provides a fundamental and operational framework in analyzing bottleneck congestion which is widespread in big cities. This paper extended the ADL's model by introducing another commuting mode. That is, we assumed there is a mass transit in addition to a road with a bottleneck both connecting a residential area and a workplace.

Reformulating ADL's model, we derived equilibria under the no-toll, the uniform toll, and the fine toll régimes respectively, and obtained several results. First, a governmental supply of transportation facilities is justified in order to reduce the negative externality of bottleneck congestion. It is socially desirable to construct a mass transit and levy a uniform road toll by a public authority in some cases even if there were no demand for the transit (Proposition 2).

Second, suppose both modes are in use and a bottleneck queue is generated in no-toll equilibrium. If the uniform toll is the only feasible policy instrument, then it should be imposed such that the number of road users is reduced by half (Proposition 3). It should be stressed that the road pricing policies derived in Proposition 3 can be conducted without estimating the parameters of α, β, γ, s, T_o and C_b. In practice, it can be attained simply by adjusting N_a by means of the road tolls through trial and error.

On the other hand, if the time-varying fine toll is politically and technologically possible, then it should be charged in proportion to the queue length, but the number of road users should remain unchanged (Proposition 4).

Finally, we made a welfare comparison between these road toll régimes. It was shown that the road tolls are effective in reducing the social cost particularly in case of heavy bottleneck congestion. The ability to lessen the social cost of the uniform toll increases as the city size gets large although it is at most 50% of the ability of the fine toll.

Footnote

1 According to Small's (1982) estimates, they are $\beta<\alpha<\gamma$.

2 Even if C_b were associated with N_b, most of the results obtained in this paper remain unchanged.

References

Arnott, R., A. de Palma and R. Lindsey (1987) "Bottleneck congestion with elastic demand," Discussion Paper #690, Queen's University.

Arnott, R., A. de Palma and R. Lindsey (1990) "Economics of a bottleneck," *Journal of Urban Economics* 27, 111-30.

Braid, R.M. (1989) "Uniform versus peak-load pricing of a bottleneck with elastic demand," *Journal of Urban Economics* 26, 320-27.

Small, K.A. (1982) "The scheduling of consumer activities: Work trips," *American Economic Review* 72, 467-79.

Vickrey, W.S. (1969) "Congestion theory and transport investment," *American Economic Review* 59,251-61.