Article:

Sakai, K, Liu, R orcid.org/0000-0003-0627-3184, Kusakabe, T et al. (1 more author) (2017) Pareto-Improving Social Optimal Pricing Schemes Based on Bottleneck Permits for Managing Congestion at a Merging Section. International Journal of Sustainable Transportation, 11 (10). pp. 737-748. ISSN 1556-8318

https://doi.org/10.1080/15568318.2017.1312646

Please cite this article as:
Sakay, K., Liu, R., Kusakabe, T. and Asakura, Y. (2017) Pareto-improving social optimal pricing schemes based on tradable bottleneck permits for managing congestion at a merging section. **International Journal of Sustainable Transportation,**
http://dx.doi.org/10.1080/15568318.2017.1312646

Pareto-Improving Social Optimal Pricing Schemes Based on Tradable Bottleneck Permits for Managing Congestion at a Merging Section

Katsuya Sakai[a*], Ronghui Liu[b], Takahiko Kusakabe[c], Yasuo Asakura[a]
[a] Transport Studies Unit, Tokyo Institute of Technology, Japan
phone/fax: +81-3-5734-2575
[b] Institute for Transport Studies, University of Leeds, UK
[c] Center for Spatial Information Science, The University of Tokyo, Japan

Abstract

Tradable bottleneck permits scheme proposed by Akamatsu et al. (2006) is one of the first-best pricing schemes and has been shown to be able to minimize social cost. Under the scheme, road administrators issue permits which allow permit holders to pass a bottleneck at specified times and create a market where drivers can freely trade the permits. However, the scheme is not always Pareto-improving such that it may harm some drivers. The objective of this study is to design Pareto-improving tradable bottleneck permits schemes for a V-shaped two-to-one merge bottleneck. Firstly, the paper formulates the morning commute model in the network and describes the arrival time choice equilibrium in the network with merging. Secondly, we show that the first-best pricing scheme under tradable bottleneck permits for this V-shaped network does not always achieve a Pareto improvement, with the cost of one group of drivers is increased by the permit pricing, a phenomena akin to the bottleneck paradox of Arnott et al. (1993). We propose therefore three implementations of tradable bottleneck permits for Pareto-improving: (i) merging priority rule is included in the tradable bottleneck permits scheme by creating different market for each origin; (ii) the permit revenues are refunded as monetary compensation to drivers whose cost is increased; and (iii) the permit revenues are used to expand bottleneck capacity. For each implementation, we derive their equilibrium solutions and demonstrate that a Pareto improvement is achieved and social cost is decreased by using the permit revenues for expanding the bottleneck capacity.

Keywords: Social optimal pricing**;** Tradable Bottleneck Permits; Pareto improvement; Merging

1. Introduction

Morning commute through traffic congestion to a central business district (CBD) is a common feature in cities around the world. The congestion is not caused so much by high travel demand but by a concentration of demand during a short (peak) period. Therefore, the congestion can be eliminated by an appropriate management to break up the demand among the longer time interval.

The departure time choice equilibrium is considered to deal with such problems about the morning commute. (e.g. Vickley, 1969) Under the equilibrium, no drivers can improve their travel cost by unilaterally changing their behavior. There have been many studies about the departure time choice equilibrium with consideration of the drivers' heterogeneity (e.g. Arnott et al., 1988, 1992, 1994) and network shape (e.g. Kuwahara, 1990, Arnott et al. 1993).

Congestion pricing is one of the management methods to control the traffic congestion. The congestion can be completely eliminated with an appropriate congestion charge, which is the same price as drivers would suffer waiting delay cost without pricing (Arnott et al., 1990a). One of the implementation methods to realize the social optimum is Tradable bottleneck permits (TBP) scheme (Akamatsu et al. 2006).

On the other hand, there is a problem about equity, which is that some drivers get a benefit but others suffer a loss. One of the criterions for the equity is a Pareto improvement. The Pareto improvement is a property that a policy harms no one and helps at least one person. Therefore, no one is supposed to say any complains about a policy if it is Pareto improving.

This study focuses on a merging bottleneck. The merging ratio from the upper reaches changes depending on the state at the merging point. When the merging point is saturated and queue occurs, the bottleneck capacity is allocated to the upper reaches. Before the pricing is implemented, the capacity allocation depends on the merging rule at the merging point; number of lanes, merging priority, and so on. After the pricing is implemented, on the other hand, the capacity allocation is determined by the toll, which eliminates the property of origins about merging priority. In this manner, the capacity allocation rule is changed by implementing the congestion pricing, and it may not be Pareto improving.

The objectives of this study are to reveal that the optimal dynamic congestion pricing with tradable bottleneck permits (TBP) scheme at a merging section is not always Pareto improving and to propose TBP implementation for Pareto improving. We consider V-shaped merging network, which is the minimum unit network with a merging bottleneck.

In Section 2, previous studies are reviewed. In Section 3, we describe the problem and provide the basic assumptions of the study. In Section 4, we formulate the equilibrium condition and derive the solution without pricing. In Section 5, we formulate the equilibrium condition and derive the solution with the first best-pricing by using TBP. We prove that the first-best pricing is not always Pareto improving in the V-shaped network. In Section 6, we propose three schemes to achieve a Pareto improvement. Finally, conclusions are drawn in Section 7.

2. Literature reviews

2.1 Departure time choice

Vickrey (1969) firstly modelled the peak hour behavior as a departure time choice with a trade-off between schedule cost and congestion delay cost. Small (1982) applied the model to real data and estimated the values of delay and schedule costs. As the developed studies considering the more complicated network, the departure time choice equilibrium on a two-tandem bottleneck network was formulated by Kuwahara (1990). The study found that a queue may form at the upstream bottleneck as well as at the downstream bottleneck even though the demand of upstream group is lower than the upstream bottleneck capacity. The phenomenon is caused because the upstream drivers were forced to depart earlier than necessary once a queue forms at the downstream bottleneck. Arnott et al. (1993) formulated the equilibrium on the Y-shaped network and concluded that the social trip cost may increase even if the upper bottleneck capacity is increased, leading to a capacity paradox. The paradox means that decreasing inflows, for example by ramp metering, can be efficient to decreasing social cost by departure time choice as well as increase of bottleneck capacity at the merge point. Daniel et al. (2009) conducted a laboratory experiment and replicated the bottleneck paradox, considering with schedule delay as well as schedule early. Xiao et al. (2014) studied the Y-shaped network under a stochastic setting. They formulate the equilibrium departure-time patterns in a Y-shaped network with stochastic bottleneck capacity, and show that the uncertainty in bottleneck capacity increases commuters' mean trip cost and lengthens the peak period, and the cost is sensitive to upstream traffic control mechanism.

2.2 Congestion pricing

In order to control the morning traffic congestion, road pricing has been considered as one of the efficient methods. The pricing scheme is based on the economic theory of marginal cost and is a mechanism to improve social benefit (Pigou 1920). An appropriately designed pricing scheme can increase social welfare. However, if the toll is inappropriately determined, road capacity may not be fully utilized, a queue at bottleneck may form and social welfare may decrease. Arnott et al. (1990a) derived a time-varying pricing scheme which can eliminate a queue at a bottleneck and realize social optimal state. However, it is difficult to apply this so-called "first-best" scheme because it requires perfect knowledge on drivers' preferences (to the desired arrival time, schedule cost, delay cost and so on) in order to determine the appropriate time-varying tolls. As alternatives, second-best pricing schemes such as step toll have been proposed (e.g. Laih 1994, 2004; Lindsey et al. 2012). In order to determine the social optimal toll in the step toll schemes, information about the precise peak time period during which congestion occurs without pricing, is still required. Recently, new pricing schemes based on market mechanism have emerged (Verhoef et al., 1997) which, in a road transportation system, propose to use tradable permits to pay for road transport externality. Tradable credit (TC) and tradable bottleneck permits (TBP) are two such pricing schemes; they were first proposed by Yang and Wang (2011) and Akamatsu et al. (2006) respectively.

The study about TC was originally focused on mobility management. Xiao et al. (2013) extended it to departure time choice problem. See a review on the studies about TC in Grant-Muller and Xu (2014). The study of TBP was originally based on managing congestion including both route choice and departure time choice problems (Akamatsu et al. 2006). Both TC and TBP use market mechanism to determine price of credit or permit. A distinctive difference between TPB and TC is that in TBP schemes, the social optimal state can be achieved if only road administrator knows each bottleneck capacity. In TC schemes, market mechanism decides only the price of the credit and road administrator has to decide time-varying charging rate. In TBP schemes, on the other hand, bottleneck permits are issued for each pre-specifies time period, which means that demands for bottlenecks are quantitatively controlled, and drivers trade the permits in respective markets. It is proved that the charging price rates determined by market mechanism are the same as the social optimal charge rates (Akamatsu et al., 2006).

Take a simple one-to-one network (i.e. one origin to one destination) with a bottleneck for example. With TC, a road administrator issues certain number of credits, determines charging rate of the credits to pass the bottleneck in a specified time period, and creates a market where drivers can freely trade the credits. At equilibrium, the price of the credit and the flow rate at the bottleneck can be determined. If the number of credits issued and credit charge rate are properly set, the first-best time-varying pricing can be realized (Xiao et al. 2013). However, it is not easy to set them properly because information about drivers' preferences, which is not easily observed, is required to do so.

With TBP, on the other hand, the road administrator issues bottleneck permits and creates markets whereby drivers can freely trade the permits. The permits are required for drivers to pass through the bottleneck at a specified time. The number of permits issued is equal to the bottleneck capacity. As a result, the price of the permit for each time is determined by market mechanism, which is equivalent to the toll of the first best pricing. Akamatsu (2007) proved that the equilibrium solution under the TBP scheme is the same as the solution of the social optimum traffic assignment. Therefore, the social cost can be minimized in a general network by using the TBP scheme. This study focuses on the equilibrium state realized by applying TBP, not on the process of trading the permits. One example of the process is discussed by Wada and Akamatsu (2013).

2.3 Pareto improvement

Even if the social optimal state can be realized by the first-best dynamic congestion pricing, it is not always Pareto improving. It is Pareto improving if the scheme harms no one and helps at least one person. In economic theory, this property is called a Pareto improvement and is a necessary condition for a scheme to be adopted as a general policy.

The Pareto improving schemes are categorized in two groups. The first one is using the revenues. Nie and Liu (2010) derived a condition that a Pareto improvement is achieved by self-financing under a situation in which drivers can choose to travel by either a car or public transport. As a result, the study proved that self-financing and a Pareto-improving toll scheme always exist when the value of time distribution function is concave. Yodoshi and Akamatsu (2008) studied the efficiency of the TBP in a network with two-tandem bottlenecks. They studied the usage of revenues for increasing bottleneck capacity and revealed the condition in which the TBP scheme is Pareto improving. The self-financing has been widely studied in the field of congestion pricing (e.g. Arnott et al., 1990b, Xiao et al., 2012).

The second one is not using the revenues. Daganzo and Garcia (2000) showed Pareto improving pricing strategy by classifying drivers as either “free” or “paying” group. Song et al. (2014) applies road space rationing in addition to congestion pricing to achieve a Pareto improvement. Hall (2015) classified bottleneck capacity into “tolled” and “free” and showed the Pareto improvement can be practically achieved. The all of the studies take strategy of classification to achieve a Pareto improvement without using revenue.

2.4 Positioning of this study

In this study, we study congestion pricing in a two-to-one merge bottleneck from the aspect of Pareto improving. The two-to-one network has two aspects (shown in Fig. 1). One is a Y-shaped network (Fig. 1a) with two bottlenecks in tandem: one upstream of the merging point and one downstream of the merge. The other is a V-shaped network (Fig. 1b) with a single bottleneck at the point of merge.

Lago and Daganzo (2007) studied the departure time equilibrium in a Y-shaped network, considering a physical queue with the effects of merge and spillover (Fig. 1a). This study focuses on the application of TPB in the V-shaped two-to-one network, and develops Pareto improving pricing schemes to manage bottleneck congestion at the merge point. In this study, we reveal the effect of the TBP at the merge bottleneck by showing firstly that the first-best pricing by TBP scheme is not always Pareto improving in the V-shaped merging network. We derive drivers’ trip cost before and after applying TBP market selling scheme. Secondly, we propose three TBP schemes to achieve a Pareto improvement.

3. Problem statement

3.1 The merge network and model of merging behavior

We consider morning commute in a V-shaped network shown in Fig. 1b. The network has two origins (A and B in Fig. 1b), one destination at CBD, and one bottleneck caused by merging at point M. The network is the simplest one which includes a merging bottleneck. The capacity of the bottleneck is μ. The number of drivers from origin A and B is N_A and N_B respectively.

In the same way as Lago and Daganzo (2007), this study employs the capacity allocation rules proposed by Daganzo (1996), which are shown in Fig. 2. The rule derives a relationship between discharge flows from both upstream links AM and BM. The discharge flows from each approach, $\{x_A, x_B\}$, satisfy the following conditions: (i) the sum of the discharge flows should never exceed the available merge capacity, i.e. $x_A + x_B \le \mu$; and (ii) when a queue exists on approach r, the share of capacity used by this approach should be

$$\frac{x_r}{x_A + x_B} \ge \alpha_r, \quad r = A \text{ or } B, \tag{1}$$

where α_r is a "merge-priority" constant. These constants satisfy:

$$\alpha_A + \alpha_B = 1, \tag{2}$$

$$0 \le \alpha_A \le 1, \tag{3}$$

$$0 \le \alpha_B \le 1. \tag{4}$$

It follows from (2) that when both approaches are queued, we have the following relationship:

$$\frac{x_A}{x_B} = \frac{\alpha_A}{\alpha_B}. \tag{5}$$

3.2 Drivers' arrival-time behavior

Based on the most popular bottleneck model formulated by Vickrey (1969), we assume that each driver chooses arrival time t so as to minimize their trip cost. In this study, for convenience, the free flow travel times are assumed to be zero. The trip cost is then defined as the summation of the following three component costs: i) price of a permit (toll); ii) cost of schedule early/late, defined as the difference between desired arrival time and the actual arrival time; and iii) queueing delay cost.

Therefore, drivers solve the following optimal problem:

$$\min_t \ TC(t) = f_w\{w(t)\} + f_s\{s(t)\} + p(t), \tag{6}$$

where t is the actual arrival time to CBD, TC(t) is the trip cost to a driver who arrives at time t, p(t) is the price of a permit to arrive at time t, $f_s\{.\}$ is the schedule cost function, s(t) is the schedule delay for a driver who arrives at time t; $s(t)=t_0-t$, t_0 is the desired arrival time, $f_w\{.\}$ is the queuing delay cost function, and w(t) is the queueing delay for a driver who arrives at time t. In this study, the time departing the bottleneck is regard as arrival time t.

We assume that all drivers have the same desired arrival time, schedule cost function, and queueing delay cost function. Schedule cost function and queueing delay cost function are defined as linear functions as follows:

$$f_s\{s\} = \begin{cases} c_1 s & s \geq 0 \\ -c_2 s & s < 0 \end{cases}, \tag{7}$$

$$f_w\{w\} = bw, \tag{8}$$

where c_1 is the marginal cost of schedule early, c_2 is the marginal cost of schedule delay, and b is the marginal cost of queueing delay time.

We assume that the number of drivers to merging priority ratio from A is larger than that from B, i.e.

$$\frac{N_A}{\alpha_A} > \frac{N_B}{\alpha_B}. \tag{9}$$

This means that group B drivers have an advantage and can cut into the queues formed by group A drivers. This assumption is to simplify the problem. If we considered the opposite condition of Eq. (9), the relation between A and B can be simply swapped due to the symmetrical nature of the network considered.

4. Equilibrium condition and solution without pricing

In this section, we formulate the equilibrium conditions and derive the solutions for the basic problem described in the previous section, without any pricing intervention.

4.1 Equilibrium conditions

We consider four conditions here.

(1) Equilibrium conditions for arrival time choice:

At equilibrium, no one can improve his or her own generalized trip cost by changing the bottleneck departing time unilaterally. Therefore, if someone chooses an arrival time t, the generalized trip cost to that driver is equal to the equilibrium cost. On the contrary, if no one chooses arrival time t, the cost at that time is larger than the cost at equilibrium. The equilibrium condition for the driver's arrival time choice can be expressed as:

$$\begin{cases} TC_A(t) = TC_A^* & \text{if} \quad x_A(t) > 0 \\ TC_A(t) \geq TC_A^* & \text{if} \quad x_A(t) = 0 \end{cases} \quad \forall t, \tag{10}$$

$$\begin{cases} TC_B(t) = TC_B^* & \text{if} \quad x_B(t) > 0 \\ TC_B(t) \geq TC_B^* & \text{if} \quad x_B(t) = 0 \end{cases} \quad \forall t, \tag{11}$$

where $TC_A(t)$ and $TC_B(t)$ is the generalized trip cost to a group A and group B driver respectively who arrives at destination at time t, TC_A^* and TC_B^* is the generalized trip cost to group A and group B respectively at equilibrium, $x_A(t)$ and $x_B(t)$ is the flow rate of group A and group B respectively exiting from bottleneck at time t (i.e. the outflow rate of group A and B arriving at CBD).

(2) Demand-performance equilibrium at bottleneck:

At equilibrium, if a queue forms at the bottleneck (i.e. the waiting time at the bottleneck is positive), then the bottleneck is considered to be saturated and the exit flow rate from bottleneck is equal to the bottleneck capacity. On the contrary, if the waiting time is zero, the arrival flow rate is less than the capacity:

$$\begin{cases} x_A(t)+x_B(t)=\mu & \text{if} \quad w_A(t)>0 \quad \text{or} \quad w_B(t)>0 \\ x_A(t)+x_B(t)\le\mu & \text{if} \quad w_A(t)=0 \quad \text{and} \quad w_B(t)=0 \end{cases} \quad \forall t. \tag{12}$$

(3) Flow conservation for OD flow rates and OD travel demand:

OD travel demand for each origin has to be assigned in any time intervals.

$$\int_{t_A^s}^{t_A^f} x_A(t)dt = N_A, \tag{13}$$

$$\int_{t_B^s}^{t_B^f} x_B(t)dt = N_B. \tag{14}$$

where t_A^s and t_A^f denote the start and the finish time of the departure-time period for group A commuters, and t_B^s and t_B^f those of the group B commuters.

(4) Merging constraint:

If queues form in both links AM and BM, the merging proportion is defined by the priority rules at the merging point. Therefore, the merging constraint is derived as follows:

$$x_A(t)=\alpha_A\mu \quad \text{and} \quad x_B(t)=\alpha_B\mu \quad \text{if} \quad w_A(t)>0, w_B(t)>0 \quad \forall t. \tag{15}$$

4.2 Equilibrium solution

Let I(t) and O(t) represent the accumulative number of drivers arriving at the bottleneck merge and those exited the merge respectively at time t. It is noted that the slope of arrival curves, I_A and I_B, is $1/\alpha_A$, $1/\alpha_B$ times the actual entering flow rate when both groups are flowing. At equilibrium, queue occurring intervals are different between two groups. The details of the equilibrium solution are following (see Lago and Daganzo 2007):

$$t_A^s = t_0 - \frac{N_A+N_B}{\mu}\frac{c_2}{c_1+c_2}, \tag{16}$$

$$t_B^s = t_0 - \frac{N_B}{\alpha_B\mu}\frac{c_2}{c_1+c_2}, \tag{17}$$

$$t_B^f = t_0 + \frac{N_B}{\alpha_B \mu} \frac{c_1}{c_1 + c_2}, \tag{18}$$

$$t_A^f = t_0 + \frac{N_A + N_B}{\mu} \frac{c_1}{c_1 + c_2}, \tag{19}$$

$$\frac{dI_A(t)}{dt} = \begin{cases} \mu\left(1 - \frac{c_1}{b}\right)^{-1} & \text{for} \quad t_A^s \le t < t_0 - \frac{N_A + N_B}{b\mu} \frac{c_1 c_2}{c_1 + c_2} \\ \mu\left(1 + \frac{c_2}{b}\right)^{-1} & \text{for} \quad t_0 - \frac{N_A + N_B}{b\mu} \frac{c_1 c_2}{c_1 + c_2} \le t < t_A^f \\ 0 & \text{otherwise} \end{cases}, \tag{20}$$

$$\frac{dI_B(t)}{dt} = \begin{cases} \mu\left(1 - \frac{c_1}{b}\right)^{-1} & \text{for} \quad t_B^s \le t < t_0 - \frac{N_B}{b\alpha_B \mu} \frac{c_1 c_2}{c_1 + c_2} \\ \mu\left(1 + \frac{c_2}{b}\right)^{-1} & \text{for} \quad t_0 - \frac{N_B}{b\alpha_B \mu} \frac{c_1 c_2}{c_1 + c_2} \le t < t_B^f \\ 0 & \text{otherwise} \end{cases}, \tag{21}$$

$$\frac{dO_{A+B}(t)}{dt} = \begin{cases} \mu & \text{for} \quad t_A^s \le t < t_A^f \\ 0 & \text{otherwise} \end{cases}, \tag{22}$$

And trip cost to each group is derived as follows:

$$TC_A{}^* = \frac{c_1 c_2}{c_1 + c_2} \frac{N_A + N_B}{\mu}, \tag{23}$$

$$TC_B{}^* = \frac{c_1 c_2}{c_1 + c_2} \frac{N_B}{\alpha_B \mu}. \tag{24}$$

The trip cost to group B drivers is less than that to group A drivers because group B drivers have the merging priority and can cut into a queue formed by group A drivers at equilibrium.

5. Equilibrium under TBP scheme

In this section, we formulate the equilibrium conditions and derive the solution with TBP scheme. Under the TBP scheme, the number of issued bottleneck permits is equal to the bottleneck capacity, and the permits are freely traded by drivers. As a result, the price of the permits is determined equal to the first best time-varying toll. The following shows that the pricing is not always Pareto improving.

5.1 Equilibrium conditions

As described in Section 4.1, without TBP scheme, four conditions were valid. Under TBP scheme, on the other hand, there are only three valid conditions. The first one is the equilibrium condition for permit choice, which is corresponding to the equilibrium condition for arrival time choice. The second one is the demand-supply equilibrium conditions in permit markets, which is corresponding to the demand-supply equilibrium condition at bottleneck. The third one is flow conservation between flow rate and demand, which is the same as that without any pricing. The merging constraint is not valid under TBP scheme because no queues form at the merge bottleneck and the merging interaction, which is valid when the merging point is saturated, is eliminated.

(1) Equilibrium condition for permit choice:

At equilibrium state, no one can improve his or her own generalized trip cost by unilaterally changing their permit, which is equivalent to changing their arrival time. Therefore, the equilibrium condition is the same as that without pricing; hence Eqs. (10) and (11) also apply here.

(2) Demand-supply equilibrium (market clearing) conditions in a permit market:

At equilibrium, if the price of a certain time of permit is positive, the quantities supplied and the quantities demanded for the permit are equal; for the permit whose supply quantity exceeds the quantity demanded, the price is zero:

$$\begin{cases} x_A(t) + x_B(t) = \mu & \text{if} \quad p(t) > 0 \\ x_A(t) + x_B(t) \le \mu & \text{if} \quad p(t) = 0 \end{cases} \quad \forall t \,. \tag{25}$$

(3) Flow conservation for flow rate and travel demand:

The flow conservation for each driver group is the same as that without pricing; hence conditions (13) and (14) also apply here.

5.2 Equilibrium solution procedure

The following shows the procedure to get the equilibrium solution. From Eqs. (10) and (11), the necessary condition where a drivers choose time t as an arrival time can be derived that the derivative of generalized trip cost is equal to zero:

$$\frac{\partial TC_r}{\partial t} = 0, \quad r = A \text{ or } B \tag{26}$$

Substitute Equations (6), (7), and (8) for Eq. (26),

$$\begin{cases} b\dfrac{dw}{dt} + c_1\dfrac{ds}{dt} + \dfrac{dp}{dt} = 0 & \text{when} \quad s \ge 0 \\ b\dfrac{dw}{dt} - c_2\dfrac{ds}{dt} + \dfrac{dp}{dt} = 0 & \text{when} \quad s < 0 \end{cases} \tag{27}$$

Under the TBP scheme where the number of bottleneck permits is equal to bottleneck capacity, queue does not occur. This means:

$$\frac{\partial w}{\partial t} = 0. \tag{28}$$

Therefore, the condition where a driver chooses time t as arrival time can be derived as:

$$\frac{dp}{dt} = \begin{cases} c_1 & \text{when} \quad s \geq 0 \\ -c_2 & \text{when} \quad s < 0 \end{cases} \tag{29}$$

This means that the gradient of bottleneck permits price is derived as marginal cost of schedule delay.

On the other hand, the start and finish time of arriving is calculated by the following. From Expressions (13), (14), and (25), required time for all drivers' demand to pass through bottleneck is:

$$t^s - t^f = \frac{N}{\mu} \tag{30}$$

At equilibrium, all drivers minimize their trip cost, and generalized trip cost is always constant if at least one driver chooses time t as arrival time. This means:

$$TC_r\left(t_r^s\right) = TC_r\left(t_r^f\right) \tag{31}$$

Under the TBP scheme, queuing delay time is always zero. In addition that, the price of bottleneck permits is zero at the start and finish time. Therefore, it should be satisfied that:

$$f_s\left(t^s\right) = f_s\left(t^f\right) \tag{32}$$

From Equations (31) and (32), we can obtain t^s and t^f.

5.3 Equilibrium solution

This section describes equilibrium solution under the conditions in Section 5.1. The proof of uniqueness of the solution is provided in Appendix A. Under TBP scheme, inflow rate becomes the same as the bottleneck capacity and queue does not form. The solution is shown in Fig. 4. The inflow starting time of both group drivers becomes same, so we derive it simply as follows:

$$t_A^s = t_B^s \equiv t^s = t_0 - \frac{N_A + N_B}{\mu} \frac{c_2}{c_1 + c_2}, \tag{33}$$

$$t_A^f = t_B^f \equiv t^f = t_0 + \frac{N_A + N_B}{\mu} \frac{c_1}{c_1 + c_2}. \tag{34}$$

$$\frac{dO_{A+B}(t)}{dt} = \frac{dI_{A+B}(t)}{dt} = \begin{cases} \mu & \text{for} \quad t^s \leq t < t^f \\ 0 & \text{otherwise} \end{cases} \tag{35}$$

The RHSs of Eqs. (33) and (34) are the same as those of Eqs. (16) and (19), which means that application of TBP does not change the time period when the bottleneck capacity is fully used. The result shows that the bottleneck capacity is shared by group A and B fairly and trip costs to both groups become same.

$$TC_A{}^* = \frac{c_1 c_2}{c_1 + c_2} \frac{N_A + N_B}{\mu} \tag{36}$$

$$TC_B{}^* = \frac{c_1 c_2}{c_1 + c_2} \frac{N_A + N_B}{\mu} \tag{37}$$

This is the same as the solution for a one-to-one commuter problem with one-bottleneck, and N_A+N_B drivers.

The revenues of TBP, assumed to be represented by R, can be calculated as:

$$R = \int_{t_A^s}^{t_A^f} p(t) x_A(t) dt + \int_{t_B^s}^{t_B^f} p(t) x_B(t) dt \tag{38}$$

At the equilibrium, the revenues are

$$R = \frac{1}{2} \frac{c_1 c_2}{c_1 + c_2} \frac{(N_A + N_B)^2}{\mu} \tag{39}$$

5.4 Comparison between before and after the TBP pricing applied

Comparing the results without pricing and those with TBP schemes, we can derive the changes in trip cost to the two groups of drivers and revenues as follows:

$$\Delta TC_A{}^* = 0, \tag{40}$$

$$\Delta TC_B{}^* = \frac{c_1 c_2}{c_1 + c_2} \frac{1}{\mu} \alpha_A \left(\frac{N_A}{\alpha_A} - \frac{N_B}{\alpha_B} \right), \tag{41}$$

$$\Delta R = \frac{1}{2} \frac{c_1 c_2}{c_1 + c_2} \frac{(N_A + N_B)^2}{\mu}. \tag{42}$$

There is no change in trip cost to group A drivers because they pay in TBP the equivalent of the queueing delay cost without pricing. On the other hand, the trip cost to group B is increased with the introduction of TBP scheme. The reason is that the arrival time interval of group B drivers under TBP scheme becomes wider, and as such, they now have to pay in TBP more than the queueing delay cost without pricing. This phenomenon is based on the bottleneck paradox by Arnott et al. (1993). Without pricing, group B can decrease their trip cost by cutting into the queue formed by group A and not increasing trip cost to group A. However, after the congestion is eliminated by the pricing, this effect is also eliminated. Akamatsu et al. (2006) state that the price of the permits corresponds to queueing delay cost. However, the above results show that this relation is not always satisfied in a merging network.

6. TBP implementations for a Pareto improvement

In the previous section, we show that the application of TBP to the merge bottleneck is not always Pareto improving because the cost to group B drivers is increased by the permit pricing. For any pricing scheme to be politically acceptable, it is reasonable to expect that the scheme achieves a Pareto improvement, i.e. it harms no one and helps at least one person.

In this section, we propose three schemes to be Pareto improving with TBP at the merging section. The three schemes are that: (1) differentiated permits are issued for each driver group; (2) revenues are rebated in cash as compensation; and (3) revenues are used to finance expansion of the bottleneck. Scheme 1 does not use the TBP revenues. Schemes 2 and 3 use the revenues.

6.1 Differentiated permit (Scheme 1)

In this scheme, TBP revenues are not rebated to drivers and remain with the road administrator. In order to achieve a Pareto improvement in this way, the arrival time period of each group must not be changed between with and without pricing so as not to increase the trip cost to any drivers. The road administrator issues bottleneck permits for each group without changing the arrival time range of the groups. The scheme is equivalent to placing the pricing point to the upstream of the merge.
Under the previous section's assumptions, the solution of the permit allocation is shown in Table 1. The allocation rate can be determined by the discharge flow of each group at equilibrium without pricing.

6.1.1 Equilibrium conditions and solution

In the scheme, the equilibrium conditions about demand-performance in permit markets are derived as follows:

$$\begin{cases} x_A(t) + x_B(t) = \mu & \text{if} \quad p(t) > 0 \\ x_A(t) + x_B(t) \le \mu & \text{if} \quad p(t) = 0 \end{cases} \quad \text{when} \quad t < t_A^s,\ t_B^f < t, \tag{43}$$

$$\begin{cases} x_A(t) = \mu & \text{if} \quad p_A(t) > 0 \\ x_A(t) \le \mu & \text{if} \quad p_A(t) = 0 \\ \qquad x_B(t) = 0 \end{cases} \quad \text{when} \quad t_A^s < t < t_B^s,\ t_B^f < t < t_A^f, \tag{44}$$

$$\begin{aligned} &\begin{cases} x_A(t) = \alpha_A \mu & \text{if} \quad p_A(t) > 0 \\ x_A(t) \le \alpha_A \mu & \text{if} \quad p_A(t) = 0 \end{cases} \\ &\begin{cases} x_B(t) = \alpha_B \mu & \text{if} \quad p_B(t) > 0 \\ x_B(t) \le \alpha_B \mu & \text{if} \quad p_B(t) = 0 \end{cases} \end{aligned} \quad \text{when} \quad t_B^s < t < t_B^f, \tag{45}$$

Under these conditions, the discharge flow rate of each group becomes the same as shown in Fig.

3(a), where the congestion is eliminated by TBP. That means each driver pays in TBP equivalent to the queueing delay cost without pricing. This is because permit markets are severally created for group A and group B, which eliminates the interaction of two groups in the markets.

The revenues in this case are equal to the total delay cost without pricing because drivers pay in TBP as equivalent to their delay cost.

6.1.2 Pareto improvement

Comparing the drivers' costs and the TBP revenues between with and without pricing, we show how the scheme is Pareto improving. Table 2 shows the cost to each group and revenues of the road administrator under the different schemes. Without pricing, the costs are given in equations (23) and (24), and no revenue is generated. With scheme 1, the costs to the drivers are not changed by the pricing (e.g. equations (23)) because the drivers pay in TBP equivalent to the queueing delay cost without pricing. On the other hand, revenues remain in road administrators, which would be spent as queueing delay cost if the pricing were not applied. Therefore, the scheme is Pareto improving.

6.2 Monetary compensation (Scheme 2)

In this scheme, revenues are spent as compensation to drivers whose trip cost is increased by the effect of the pricing. In Scheme 1, TBP price was controlled by creating TBP market for individual group and the price of TBP become equivalent to the queueing delay cost without pricing. In Scheme 2, on the other hand, derivers will be paid back after first paying to get TBP whose price may be more expensive than the queueing delay cost without pricing.

In order to achieve a Pareto improvement under the scheme, the road administrator should pay back as much as the increase of the cost to the cost-increased drivers. Under the conditions in Section 3, the compensation amount for a driver in each group is determined equivalent to the increase of cost shown in Eqs. (40) and (41):

Compensation for group A: 0

Compensation for group B: $\dfrac{c_1 c_2}{c_1 + c_2} \dfrac{1}{\mu} \alpha_A \left(\dfrac{N_A}{\alpha_A} - \dfrac{N_B}{\alpha_B} \right)$

6.2.1 Equilibrium conditions and solution

The equilibrium conditions are the same as formulated in Section 5.1 because the compensations are paid ex post facto. Therefore, the equilibrium solutions are also same as derived in Section 5.2. However, group B drivers can get the compensation that is equivalent to the cost added by the application of TBP.

6.2.2 Pareto improvement

The costs and revenues are the same as those under Scheme 1 because the increased cost to group B is rebated from the revenues. With Scheme 2, road administrators pay back to group B drivers less than they pay in TBP. Therefore, this implementation is also Pareto improving.

6.3 Finance expansion of bottleneck (Scheme 3)

In this scheme, the road administrators spend revenue generated by the bottleneck permits on expanding bottleneck capacity. Yodoshi and Akamatsu (2008) studied the scheme for a straight network with two-tandem bottlenecks. Capacity expansion subject to minimizing social cost to all drivers is an optimization problem. Below, we firstly formulate the optimal problem as a minimization problem of social cost. Secondly, we explain the equilibrium solution, and finally, we show the requirement for Pareto improving.

We decide how much bottleneck capacity should be increased to minimize the social cost. The best capacity improvement $\mathit{\Delta}\mu^*$ is given by minimizing social cost which consists of the cost for improving the bottleneck capacity and the sum of drivers' schedule cost. Assuming that the bottleneck capacity is improved when starting the pricing with TPB, we derive the present social cost with the discount rate r. The minimization problem is derived as:

$$\min \quad \mathrm{SC}(\mathit{\Delta}\mu) = \frac{\mathrm{TS}(\mu + \mathit{\Delta}\mu)}{r} + \mathrm{K}(\mathit{\Delta}\mu), \tag{46}$$

$$\text{s.t.} \quad \mathit{\Delta}\mu \geq 0,$$

where SC is the social cost, $\mathit{\Delta}\mu$ is the capacity increase, TS is the total schedule cost, and K is the capacity increasing cost.

K is assumed to be homogeneous of degree 1 with respect to $\mathit{\Delta}\mu$:

$$\mathrm{K}(\mathit{\Delta}\mu) = m \cdot \mathit{\Delta}\mu, \tag{47}$$

where m is the marginal cost of increasing bottleneck capacity.

From the expressions (46) and (47), the optimality conditions can be derived as:

$$\begin{cases} \dfrac{1}{r}\dfrac{d\mathrm{TS}}{d(\mathit{\Delta}\mu)} + m = 0 & \text{if} \quad \mathit{\Delta}\mu > 0 \\ \dfrac{1}{r}\dfrac{d\mathrm{TS}}{d(\mathit{\Delta}\mu)} + m > 0 & \text{if} \quad \mathit{\Delta}\mu = 0 \end{cases}. \tag{48}$$

Total schedule cost (TS) is calculated by multiplying the triangles area formed by f_s in Fig. 3 and the departure rates:

$$\mathrm{TS} = \frac{1}{2}\frac{c_1 c_2}{c_1 + c_2}\frac{(N_A + N_B)^2}{\mu + \mathit{\Delta}\mu}. \tag{49}$$

From the expression (48) and (49), the optimal amount of capacity increase is derived as:

$$\Delta\mu^* = \sqrt{\frac{c_1 c_2}{2mr(c_1 + c_2)}} \cdot (N_A + N_B) - \mu . \tag{50}$$

This means that the optimal bottleneck capacity always exists and that it relates to the number of commuters, marginal costs of schedule delays, marginal cost for capacity-increase, and the discount rate. The optimal bottleneck capacity μ^* is therefore:

$$\mu^* = \sqrt{\frac{c_1 c_2}{2mr(c_1 + c_2)}} \cdot (N_A + N_B) . \tag{51}$$

In order to achieve the minimum social cost, the road expansion only needs to increase the difference between the optimal capacity and the present capacity. If the present bottleneck capacity is larger than the optimal capacity, the road expansion is not needed. The optimal road expansion can always be realized within the revenues of the bottleneck permits. (See appendix A for the proof)

6.3.1 Equilibrium

In scheme 3, the bottleneck capacity is expanded to the optimal capacity μ^*. The cost at equilibrium is derived by replacing the bottleneck capacity in Eqs. (36) and (37) for the optimal capacity.

6.3.2 Pareto improvement

The changes in revenues without pricing and with Scheme 3 are shown in Table 2. The scheme decreases the cost to group A drivers and increases the revenues of the bottleneck permits, but the cost to group B drivers is not always decreased.

A Pareto improvement can be achieved if the change in cost for group B drivers is not positive. The condition for the Pareto improvement is derived as:

$$\frac{c_1 c_2}{c_1 + c_2} \frac{N_A + N_B}{\mu^*} \le \frac{c_1 c_2}{c_1 + c_2} \frac{N_B}{\alpha_B \mu} . \tag{52}$$

The LHS of (52) is the trip cost to group B in the scheme. The RHS is the trip cost to group B without any pricing. Substituting eq. (51) for expression (52) and solving the expression for m, we get the following condition:

$$m \le \frac{1}{2r} \frac{c_1 c_2}{c_1 + c_2} \left(\frac{N_B}{\alpha_B \mu} \right)^2 . \tag{53}$$

Therefore, a Pareto improvement can be achieved if the marginal capacity-expansion cost m satisfies (53), i.e. is less than a value determined by the number of group B commuters, present bottleneck capacity, marginal costs of schedule delay, and the discount rate.

6.4 Comparison and discussion

In this section, we compare the schemes studied in the paper and state the features of each. Table 3 presents a summary of the proposed schemes. It can be seen that, from the perspective of social cost minimization, scheme 3 is the most effective because the scheme increases bottleneck capacity and total schedule cost of drivers is reduced. However, the scheme is not always Pareto improving. On the other hand, Schemes 1 and 2 can always achieve a Pareto improvement. Scheme 1 controls the bottleneck permits allocation for each group separately without changing the discharge flow rates of respective group between before and after applying TBP. The scheme has an advantage that it is available only by observing the discharge flow rates of respective groups. Scheme 2 is reasonable to conduct because the procedure is only compensation, which can be realized for example by discounting the charge rate of bottleneck permits.

7. Conclusion

This study focuses a merging section and reveals the effect of the social optimal pricing scheme that is implemented with TBP. The feature of a merge bottleneck is that the capacity allocation for each upstream link depends on the congestion state at the merging point. If the both upstream links are congested and queues form, the capacity at merging is determined by the merging priority rules.

The paper shows that the TBP pricing is not always Pareto improving at the merge bottleneck because the pricing eliminates the priority rules at the merging. The mechanism is that the pricing by using TBP eliminates a physical queue at the bottleneck but the prices of the permits are determined by a length of imaginary point queue at the bottleneck, which is actually not formed at bottleneck. This fact relates to the bottleneck paradox identified by Kuwahara (1990) and Arnott et al. (1990a).

The paper proposes three TBP schemes to be Pareto improving. The first one is that differentiated permits are issued for each driver group so as that TBP revenues are not returned in cash or services to drivers. In this scheme, each driver pays in TBP equivalent to his/her queueing delay cost which was generated without the TBP pricing. As a result, TBP revenues remain in the road administrator and that is Pareto improving. The second one is that revenues are rebated in cash as compensation. The scheme resulted in that the cost and revenues were the same as those under the first scheme. The last one is that revenues are used to finance expansion of the bottleneck. The good point of the third scheme is that the social cost decreases more than only applying TBP at the bottleneck. However, a Pareto improvement is sometimes not achieved, for example, in case that construction cost is high, social discount rate is high, and bottleneck capacity is already enough.

References

Akamatsu, T., 2007. Tradable Network Permits: A New Scheme for the Most Efficient Use of Network Capacity. Working Paper, Tohoku University.

<http://www.plan.civil.tohoku.ac.jp/~akamatsu/Publications/PDF/TS-TNP-P1(070311).pdf>.

Akamatsu, T., Sato, S., Nguyen, X.L., 2006. Tradable time-of-day bottleneck permits for morning commuters (in Japanese). JSCE Journal of Infrastructure Planning 62(4), 605-620.

Arnott, R., de Palma, A., Lindsey, R., 1988. Schedule Delay and Departure Time Decisions with Heterogeneous Commuters. Transportation Research Record (1197), 56-67

Arnott, R., de Palma, A., Lindsey, R., 1990a. Departure time and route choice for the morning commute. Transportation Research Part B 24(3), 209-228.

Arnott, R., de Palma, A., Lindsey, R., 1990b. Economics of a bottleneck. Journal of Urban Economics 27(1), 111-130.

Arnott, R., de Palma, A., Lindsey, R., 1992. Route choice with heterogeneous drivers and group-specific congestion costs. Regional Science and Urban Economics 22(1), 71-102.

Arnott, R., de Palma, A., Lindsey, R., 1993. Properties of dynamic traffic equilibrium involving bottlenecks, including a paradox and metering. Transportation Science 27(2), 148-160.

Daganzo, C.F., 1996. The nature of freeway gridlock and how to prevent it. In: Lesort, J.B. (Ed.), Transportation and Traffic Theory (Proceedings of the 13th International Symposium). Pergamon Press, Amsterdam, 629-646.

Daganzo, C.F., Garcia, R. C., 2000. A Pareto Improving Strategy for the Time-Dependent Morning Commute Problem. Transportation Science 34(3), 303-311.

Daniel, T.E., Gisches, E.J., Rapoport, A., 2009. Departure times in Y-shaped traffic networks with multiple bottlenecks. American Economic Review 99(5), 2149-2176.

Grant-Muller, S., Xu, M., 2014. The role of tradable credit schemes in road traffic congestion management. Transport Reviews 34(2), 128-149.

Hall, J. D., 2015. Pareto Improvements from Lexus Lanes: the Effect of Pricing a Portion of the Lanes on Congested Highways. Working paper in University of Toronto <http://individual.utoronto.ca/jhall/documents/PI_from_LL.pdf>.

Iryo, T., Yoshii, T., Asakura, Y., 2005. Mathematical analysis of equilibrium in departure time choice problems (in Japanese), JSCE Journal of Infrastructure Planning and Management 779(VI-66), 105–118.

Iryo, T., Yoshii, T., 2007. Equivalent optimization problem for finding equilibrium in the bottleneck model with departure time choices, Mathematics in transport: Selected Proceedings of the 4th IMA International Conference on Mathematics in Transport, 231–244.Kuwahara, M., 1990. Equilibrium queueing patterns at a two-tandem bottleneck during the morning peak. Transportation Science 24(3), 217-229.

Kuwahara, M., 1990. Equilibrium queueing patterns at a two-tandem bottleneck during the morning peak. Transportation Science, 24(3), 217-229.

Lago, A., Daganzo, C.F., 2007. Spillovers, merging traffic and the morning commute. Transportation

Research Part B 41(6), 670-683.

Laih, C.-H., 1994 Queueing at a bottleneck with single- and multi-step tolls. Transportation Research Part B 28(3), 197-208.

Laih, C.-H., 2004. Effects of the optimal step toll scheme on equilibrium commuter behaviour. Applied Economics 36, 59-81.

Lindsey, C.R., van den Berg, V. A., Verhoef, E.T., 2012. Step tolling with bottleneck queuing congestion. Journal of Urban Economics 72(1), 46-59.

Nie, Y.M., Liu, Y., 2010. Existence of self-financing and Pareto-improving congestion pricing: Impact of value of time distribution. Transportation Research Part A: Policy and Practice 44(1), 39-51.

Pigou, A., 1920. The economics of welfare. London, Macmillan.

Small, K.A., 1982. The Scheduling of Consumer Activities: Work Trips. The American Economic Review 72(3), 467-479.

Song, Z., Yin, Y., Lawphongpanich, S., Yang, H., 2014. A Pareto-improving hybrid policy for transportation networks. Journal of Advanced Transportation 48(3), 272-286.

Verhoef, E., Nijkamp, P., Rietveld, P., 1997. Tradeable permits: their potential in the regulation of road transport externalities. Environment and Planning B 24(4), 527-548.

Vickrey, W.S., 1969. Congestion theory and transport investment. The American Economic Review 59 (2), 251-260.

Wada, K., Akamatsu, T., 2013. A hybrid implementation mechanism of tradable network permits system which obviates path enumeration: An auction mechanism with day-to-day capacity control. Transportation Research Part E 60, 94-112.

Xiao, F., Qian, Z.S., Zhang, H.M., 2013. Managing bottleneck congestion with tradable credits. Transportation Research Part B 56, 1-14.

Xiao, F., Shen, W., Zhang, H.M., 2012. The morning commute under flat toll and tactical waiting. Transportation Research Part B: Methodological 46(10), 1346-1359.

Xiao, L.-L., Liu, R., Huang, H.-J., 2014. Stochastic bottleneck capacity, merging traffic and morning commute. Transportation Research Part E. Vol 64, 48-70.

Yang, H., Wang, X., 2011. Managing network mobility with tradable credits. Transportation Research Part B 45(3), 580-594.

Yodoshi, M., Akamatsu, T., 2008. Pareto improvement properties of tradable permit systems for a tandem bottleneck network (in Japanese). Infrastructure Planning Review 25(4), 897-908.

Appendix A: Proving uniqueness of equilibrium solution under TBP scheme

Appendix A proves the following proposition.

Proposition: On the assumption that all drivers have the same linear generalized trip cost function, if the equilibrium solution exists under the TBP scheme, it is unique.

Proof: According to Iryo et al. (2005) and Iryo and Yoshii (2007), there is an optimization problem whose solution is the equilibrium solution. The optimization problem is to minimize the all drivers' total schedule cost of delay time unit. Under the equilibrium conditions derived in Section 5.1, the equivalent optimization problem is derived as:

$$\min_{\{x_r(t)\}} \int_t \frac{f_s\{s(t)\}}{p(t)} \left\{ \sum_{r=A,B} x_r(t) \right\} dt \tag{A1}$$

$$\begin{aligned} \text{s.t.} \quad & \sum_{r=A,B} x_r(t) \leq \mu \quad && \forall t \\ & \int_t x_r(t) dt = N_r && \forall r = \{A, B\} \\ & x_r(t) \geq 0 && \forall r = \{A, B\}, \forall t \end{aligned} \tag{A2}$$

With the assumptions that schedule cost function is linear, the form of optimization problem (A1) with constraints (A2) is completely linear. Therefore, if the problem has a solution, it is unique.

Appendix B: Condition where the optimal road expansion can be self-financed (financed only by the revenues from the bottleneck permits)

The condition that the optimal road expansion cost is covered by the revenue of TBP is derived as:

$$\frac{R(\varDelta\mu^*)}{r} \geq K(\varDelta\mu^*), \tag{B1}$$

where R is the revenues of TBP in a day, and r is social discount rate. The revenue can be calculated by eq. (31) and derived as:

$$R(\varDelta\mu^*) = \frac{1}{2} \frac{c_1 c_2}{c_1 + c_2} \frac{(N_A + N_B)^2}{\mu + \varDelta\mu^*}. \tag{B2}$$

The RHS of equation (B2) is the same as that of equation (42) because of the linear assumption on the schedule cost and queueing delay cost. Substitute Eqs. (40) and (B2) for condition (B1),

$$\frac{1}{2r} \frac{c_1 c_2}{c_1 + c_2} \frac{(N_A + N_B)^2}{\mu + \varDelta\mu^*} \geq m \cdot \varDelta\mu^*. \tag{B3}$$

Substitute Eq. (43) for condition (B3),

$$\frac{1}{2r} \frac{c_1 c_2}{c_1 + c_2} \frac{(N_A + N_B)^2}{\sqrt{\frac{c_1 c_2}{2mr(c_1 + c_2)}} \cdot (N_A + N_B)} \geq m \cdot \left(\sqrt{\frac{c_1 c_2}{2mr(c_1 + c_2)}} \cdot (N_A + N_B) - \mu \right) \tag{B4}$$

Divide both sides of condition (B4) by m (>0) and we can get

$$\sqrt{\frac{c_1 c_2}{2mr(c_1 + c_2)}} (N_A + N_B) \geq \sqrt{\frac{c_1 c_2}{2mr(c_1 + c_2)}} (N_A + N_B) - \mu \tag{B5}$$

Therefore, the condition where the road expansion cost can be covered by TBP is derived as:

$$\mu \geq 0, \tag{B6}$$

which means that the optimal road expansion can always be realized within the TBP revenues.