OPTIMAL CONGESTION CHARGES IN GENERAL EQUILIBRIUM

Dong-Joo MOON, Ph.D.
Research Professor
School of Civil, Urban, and Geosystem Engineering
Seoul National University
Seoul, Korea

Chang-Ho PARK, Ph.D.
Professor
School of Civil, Urban, and Geosystem Engineering
Seoul National University
Seoul, Korea

(Received December 13, 2001)

This paper deals with pricing and investment decision problems of multi-route and multi-period highway systems in which the congestion is a significant factor in the assessment of system costs. This study approaches this congestion pricing scheme with two different social welfare maximization problems, both of which search for the optimal solutions through general equilibrium analysis. These two optimization problems have an identical structure except financial constraints that reflect different decision environments.

One welfare maximization problem involves estimating the first-best social optimal solution. This problem yields the optimal solution for the implementation scheme to impose the differentiated congestion charge for each trip alternative in terms of travel route and trip period. The optimal congestion charge for this problem has the expression similar to that derived in previous studies dealing with congestion pricing.

Another maximization problem involves characterizing the second-best optimal solution. In this problem, it is assumed to impose the congestion toll only on a single highway link. This problem yields the second-best congestion toll different from the first-best one. This second-best optimal congestion toll has the structure to reflect its impact on other highway links exempt from the congestion charge program.

Key Words: Highway service process, Marginal cost, Congestion charge, Investment decision, Social optimality

1. INTRODUCTION

Highway transportation systems have a number of peculiar characteristics that seem responsible for their being dealt with as separate topics of marginal cost pricing. One is ‘congestion phenomena’ that the average user cost increases as the number of users increase. Another one is ‘peaking phenomena’ that the demand concentrates on a certain period of the day or week. The other one is ‘network situation’ that more than one alternative route is available to trip-makers.

Congestion pricing might be the most appropriate pricing scheme to accommodate these characteristics of highway systems. This economic analysis scheme is a special kind of marginal cost pricing, which deals with the optimal decision rules for the investment and pricing of excludable public goods. This pricing scheme has an advantage of reflecting the effects of congestion and peaking demand in searching for the optimal decision rules.

This aspect of congestion pricing is well examined in the pricing and investment decision problem of urban highways[1]. This example deals with the optimization problem to maximize the net benefit, NB, namely:

$$\max NB \equiv \sum_t \alpha_t \left[\int_0^{Q_t} P_t(q_t)dq_t - Q_t F_t(Q_t;y) \right] - SC(y)$$

Here, the parameter α_t represents the duration of the tth period for every $t \in \langle 1, T \rangle$, and satisfies the condition that $\sum \alpha_t = 1.0$ and $\alpha_t \rangle 0$. The benefit of highway users is expressed as the inverse demand function P_t. The cost is expressed as the sum of user cost $\sum \alpha_t Q_t F_t(Q_t;y)$ in T periods and supplier cost $SC(y)$, where F_t refers to the average user cost function, and y represents the highway capacity.

This net benefit maximization problem searches for the optimal investment and pricing rules through partial equilibrium analysis. This optimization problem reflects the congestion effect by using the average user cost function $F_t(Q_t)$, which is monotonically increasing with respect to Q_t. In addition, the optimization problem expresses the total benefit and costs as the sum of T periods, in order to accommodate the effect of peaking phenomena.

However, this approach has a limitation associated with partial equilibrium analysis, which is not suitable to assess the impact of the congestion charge on other transportation facilities or services. This approach therefore has the difficulty to address a number of relevant policy questions related to the congestion charge. One such policy question would be to characterize the optimal solution for investment and pricing decisions, which can attain the first-best social optimality in welfare economics. Another one would be to find the optimal pricing scheme for a certain highway route, which can comprehend its effect on the welfare changes of trip-makers using other competing alternative routes.

The main objective of this study is to find the optimal decision rules for the investment and pricing of multi-route and -period highway systems in the presence of congestion. One of our key concerns is to compare the decision rule for the multi-route highway system with the rule for single highway route, which has been examined in most studies for congestion pricing[1,2]. Another one is to assess a number of relevant policy questions associated with congestion pricing, including those mentioned above.

We approach these policy questions with two different social welfare maximization problems, which search for the optimal investment and pricing rules through general equilibrium analyses. One maximization problem involves estimating the optimal congestion charge corresponding to the first-best optimal solution. This maximization problem is defined to estimate the adequate amount of congestion charges for all the highway routes. Another involves estimating the second-best optimal solution for the case that the congestion charge is levied only upon a particular route.

These two welfare maximization problems have an identical structure except for the constraint about the implementation program of congestion charges. Specifically, the objective function of these two problems corresponds to the social welfare function, which was firstly applied in assessing marginal cost pricing by Mohring[3]. The structure of constraints is similar to the one used in analyzing the optimal taxation rule by Diamond and Mirrles[4]. One set of constraints to the optimization problems represents the process of transportation services in the presence of congestion. Another constraint involves the financial requirements faced by the supplier of transportation services.

The structure of the paper is as follows. In Section 2, we introduce an alternative approach to specifying the service process of highways, which is outlined above. In Section 3, we examine the welfare maximization problem, which searches for the solution to attain the first-best social optimality. Subsequently, in Section 4, we assess the amended welfare maximization problem defined to address a policy question, namely, the effectiveness of congestion charges on a single highway link.

2. HIGHWAY SERVICE PROCESS

The objective of this section is to explain our approach to specifying the constraints of the welfare maximization examined in subsequent sections. We firstly introduce an alternative specification for the highway service process equivalent to the production process of ordinary goods. Subsequently, we set up the highway cost function and the user equilibrium condition with the specification for the service process defined in the previous step.

2.1 Specification for service processes

To start, we examine the approach to specifying the production process of a public good free from congestion. Suppose that this public good g is the output of the production process G with the input of ordinary goods $\mathbf{x} \equiv (x_1, x_2, \cdots, x_J)$. Then, this public production process could be expressed as follows:

$$g - G(\mathbf{x}) \leq 0$$

where $g \geq 0$ and $\mathbf{x} \geq \mathbf{0}$.

On the other hand, the service mechanism of highway systems could be decomposed into two separate processes. The first involves the production process of highway facility with ordinary goods. The second is associated with the congestion process, which determines the average user cost of the system. These two processes are interconnected with the highway capacity corresponding to the output of the first process and also the input of the second process.

The output of the first process is assumed to be the highway capacity, which corresponds to the measure for the size of facilities. This highway capacity is also assumed to be a sole variable that determines the service characteristics. Furthermore, it is assumed that the capacity, denoted by y, is a real number, i.e., the highway is a divisible facility.

At this stage, it would be necessary to comment on the validity of the second assumption. The capacity of highways is not determined solely by number of lanes; but also by many other physical variables, such as width of lanes, access control method, grade, alignment, etc. It would therefore be acceptable that the image of the function to estimate the capacity with the inputs of these physical variables could be real numbers.

The input of the first process corresponds to ordinary goods $\mathbf{x} \equiv (x_1, x_2, \cdots, x_J)$, which represent the quantities of goods traded at the market, such as fuel, manpower, construction material, etc. These inputs are assumed to be available to the supplier of highway services, at the before-tax price $\mathbf{p} \equiv (p_1, p_2, \cdots, p_J)$. Also an identical type of good is represented as different goods, if their consumption periods are different.

The production of output y with inputs $\mathbf{x}$ involves a production process G, which is assumed to fulfill the following. First, the process G satisfies:

$$y - G(\mathbf{x}) \leq 0 \quad (1)$$

Second, the function G is concave and differentiable with respect to $\mathbf{x}$.

The second process involves the service mechanism to determine its average user cost. This user cost commonly refers to the value of the resources consumed from the viewpoint of society, and usually excludes highway tolls. This user cost is short-run in nature. Accordingly, the user cost of each period t is determined by a process F_t, which is separated from those of other periods.

The process F_t is actually the average user cost function, which is a kind of delay function of queuing systems. One relevant way of expressing the process F_t is:

$$F_t\,(g_t;y) - C_t \leq 0 \quad (2)$$

where C_t is a non-negative continuous variable, and represents the average user cost at the t th period. The inequality in (2) reflects that the average user cost $\mathbf{C} \equiv (C_1, C_2, \cdots, C_T)$ is the supplier's independent decision variable. Alternatively, the inequality implies that the supplier could choose $\mathbf{C}$ larger than the technically attainable value determined by the function F_t.

We make an additional number of assumptions about the structure of the average user cost function F_t. First, the function F_t is convex, increasing and differentiable with respect to output g_t. The delay function of queuing systems is of this form. Second, the function F_t is convex, decreasing, and differentiable with respect to capacity y. These assumptions are also satisfied by the delay functions that are commonly used in transportation studies, except for the differentiability, which is an extension of divisibility assumption mentioned earlier.

One example that fulfills these assumptions about convexity and differentiability is the average user cost function proposed in "Highway Capacity Manual[5]". This function is:

$$F_t\,(g_t;y) \equiv t_0 + t_1\,(g_t/y)^s \quad (3)$$

Here, the first term t_0 represents the user cost in the free flow condition of highways. The second term $t_1\,(g_t/y)^s$ represents the additional cost accrued by congestion, where constant s is reported to be 4.0 ~ 6.0.

2.2 Cost functions for highway services

The total cost function, by definition, estimates the minimum social costs necessary for the production of average output rates $\mathbf{g} \equiv (g_1, g_2, \cdots, g_T)$. The total cost function is composed of supplier cost function and user cost function. The supplier cost function involves the production process G to yield capacity y with input $\mathbf{x}$. The user cost function is related with the queuing process F_t to determine the average user cost C_t for the given g_t and y .

We define the total cost function as the solution to the social cost minimization problem associated with the production of output rate $\mathbf{g}$. This social cost minimization problem could be formulated in two steps, which are interrelated. The first step involves estimating the supplier cost function, which is the solution to the cost minimization problem to choose the optimal value of $\mathbf{x}$ necessary for the production of a given y. The Lagrangian for this minimization problem is:

$$SC\,(y;\mathbf{p}) \equiv \min\ \sum_j p_j x_j + \phi\big(y - G(\mathbf{x})\big)^{*} \quad (4)$$

where $\phi \geq 0$ is a Lagrangian coefficient. The second step involves searching for the optimal capacity y, which minimizes the total cost in producing a given vector of output rate $\mathbf{g} \equiv (g_1, g_2, \cdots, g_T)$. The decision variables for this minimization problem are y and $\mathbf{C}$. The Lagrangian for this problem, denoted by TC, is:

$$TC\,(\mathbf{g};\mathbf{p}) \equiv \min\ \Big\{SC\,(y;\mathbf{p}) + \sum_t \alpha_t g_t C_t\Big\} + \sum_t \lambda_t \alpha_t \big(F_t\,(g_t;y) - C_t\big) \quad (5)$$

where $\lambda_t \geq 0$ is the Lagrangian coefficient.

Substituting the supplier cost function in equation (4) into the total cost function in equation (5) yields the alternative expression that can provide a more comprehensive picture about the decision variables. This alternative expression involves estimating the optimal values of $\mathbf{x}$, y and $\mathbf{C}$ necessary for yielding the output rate $\mathbf{g}$. This problem is:

$$TC\,(\mathbf{g};\mathbf{p}) \equiv \min\ \Big\{\sum_j p_j x_j + \sum_t \alpha_t g_t C_t\Big\} + \phi\big(y - G(\mathbf{x})\big) + \sum_t \lambda_t \alpha_t \big(F_t\,(g_t;y) - C_t\big) \quad (6)$$

Note that this cost function yields the identical solution with the problem in equation (5) for any combination of non-negative g and p.

Subsequently, we estimate the marginal cost function with respect to outputs for the total cost function TC in equation (6). To this end, we first determine the unit for measuring the outputs of each period. We also find

* For simplicity of the presentation, we omit the terms representing the constraints for the boundaries of the decision variables, such as $y \geq 0$ and $\mathbf{x} \geq \mathbf{0}$. Note also that we describe the Lagrangian of the other optimization problems in this abbreviated form throughout this study.

that the quantity of output in the t th period is not g_t representing the output rate; but rather $\alpha_t\, g_t$, corresponding to the actual number of outputs produced during the time interval α_t.

We define the factor demand functions of $\mathbf{x}$, y and $\mathbf{C}$ with respect to independent variables $\mathbf{g}$ and $\mathbf{p}$, which are denoted by $\hat{\mathbf{x}} \equiv (\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_J)$, $\hat{y}$ and $\hat{\mathbf{C}} \equiv (\hat{C}_1, \hat{C}_2, \cdots, \hat{C}_T)$ respectively. Each factor demand function by definition estimates the optimal value of the corresponding decision variable, which satisfies the necessary conditions for the minimization problem in equation (6). We also define the functions $\hat{\phi}$ and $\hat{\mathrm{e}} \equiv (\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_T)$, for ϕ and $\ddot{\mathbf{e}} \equiv (\lambda_1, \lambda_2, \cdots, \lambda_T)$ respectively, in the same manner.

Replacing all the unknown decision variables in equation (6) with their corresponding functions defined above, differentiating this alternative expression with respect to $\alpha_t\, g_t$, and simplifying the resulting expression by substituting the necessary conditions of the optimization problem yields:

$$MC_t(\mathbf{g};\mathbf{p}) \equiv \frac{\partial TC(\mathbf{g};\mathbf{p})}{\partial \alpha_t g_t} = F_t\big(g;\hat{y}(\mathbf{g};\mathbf{p})\big) + g_t \left.\frac{\partial F_t(g_t;y)}{\partial g_t}\right|_{y=\hat{y}(\mathbf{g};\mathbf{p})} \quad\text{.................... (7)}$$

This long-run marginal cost function estimates the additional increase in total costs accrued by an increase in outputs from $\alpha_t\, g_t$ to $\alpha_t\, g_t + 1$.

The first term of equation (7) corresponds to the *average user cost function*, which estimates the average user cost per unit of trip at the t th period. This function is alternatively expressed as:

$$AUC_t(\mathbf{g};\mathbf{p}) \equiv F_t(g_t;\, \hat{y}(\mathbf{g};\mathbf{p})) \quad\text{.......... (8)}$$

The second term of equation (7) is called the *marginal congestion cost function*. This function estimates the additional user cost per trip caused by adding one more unit of trip to the system:

$$MCC_t(\mathbf{g};\mathbf{p}) \equiv g_t \left.\frac{\partial F_t(g_t;y)}{\partial g_t}\right|_{y=\hat{y}(\mathbf{g};\mathbf{p})} \quad\text{........................ (9)}$$

Note also that the functional structure of the long-run marginal cost function MC_t in equation (7) is identical to that of the short-run marginal cost function for the total user cost.

This long-run marginal cost function MC_t for a given output $\mathbf{g^0}$ has the same value as the short-run marginal cost function of the system with the optimal capacities for the given $\mathbf{g^0}$. To clarify this point, suppose the highway has the capacity $y^0 \equiv \hat{y}(\mathbf{g^0};\mathbf{p})$. Then, the short-run marginal cost for the given y^0, denoted by $SRMC_t$, satisfies the following:

$$SRMC_t(g_t^0;y^0) \equiv F_t(g_t^0;y^0) + g_t^0 \left.\frac{\partial F_t(g_t;y^0)}{\partial g_t}\right|_{g_t=g_t^0} = MC_t(\mathbf{g^0};\mathbf{p}) \quad\text{.. (10)}$$

where the second equality is a direct consequence of the definition of MC_t in equation (7).

Finally, we illustrate the configurations of various cost functions derived above with a simple example. To this end, assume that the supplier cost function SC_a has a linear relationship with the capacity y:

$$SC_a = a + by \quad\text{.......... (11)}$$

where $a \geq 0$ and $b \rangle 0$. Assume further that the average output rate g is constant across all periods, and that the average user cost function has the functional form in equation (3).

Then, the highway cost function in equation (5) can be simplified as:

$$TC_a(g) \equiv \min\{a + by + gC\} + \lambda(F(g;y) - C) \quad\text{.......... (12)}$$

For this cost function, the first order condition with respect to y is:

$$b = -t_1\, g \frac{\partial f_a}{\partial y} \quad\text{.......... (13)}$$

where $f_a \equiv (g/y)^s$. Solving the above equation with respect to y yields the following factor demand function:

$$\hat{y}(g) = \delta^{1/s+1} g \quad\text{.......... (14)}$$

where $\delta = t_1 s/b$. Note that $\delta \rangle 1$, since $T = 1$. Otherwise, the output rate exceeds the capacity.

Substituting equation (14) into equation (12), and differentiating the resulting expression with respect to g yields the following long-run marginal cost function:

$$MC_a(g) = AUC_a(g) + MCC_a(g) \quad\text{.......... (15)}$$

Here, AUC_a and MCC_a have the following specific expressions:

$$AUC_a(g) = t_0 + t_1 \delta^{-s/s+1} = t_0 + \frac{b}{s}\delta^{1/s+1} \quad\text{.......... (16)}$$

$$MCC_a(g) = t_1 s \delta^{-s/s+1} = b\delta^{1/s+1} \quad\text{.......... (17)}$$

In addition, the marginal cost function MCC_a has the following relationships with supplier costs:

$$MCC_a(g) = MSC_a(g) = b\frac{\partial \hat{y}(g)}{\partial g} = AVSC_a(g) = b\frac{\hat{y}(g)}{g} \quad\text{.......... (18)}$$

where MSC_a refers to the marginal supplier cost function, and $AVSC_a$ represents the average variable supplier cost function.

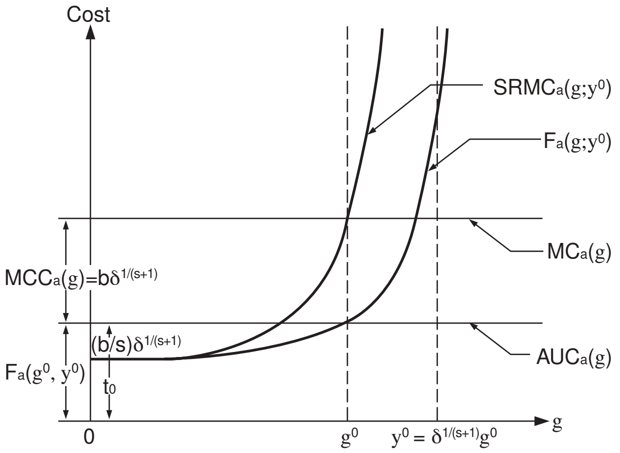

Fig. 1 Schematic illustration of highway cost function

Figure 1 schematically depicts the above results. The figure shows that the marginal cost function MC_a is equal to the sum of AUC_a and MCC_a. The figure also illustrates the relationship between MC_a and $SRMC_a$ for a given output g^0, and the relationship between MCC_a and MSC_a in equation (18).

2.3 Specification for user equilibrium

So far, we have assumed that the output rate **g** in the average user cost function F_t is the independent decision variable of suppliers. Instead, this output rate **g** is actually a dependent variable of the interaction process between a highway facility and its users. More precisely, the output corresponds to the user demand at user equilibrium, which refers to the stable state reached through the process of a trip-maker's search for optimal choices in terms of frequency of trips, time of day, route and so on.

To start, we identify a number of requirements that a specification of user equilibrium should fulfill. First, the formulation can reasonably depict user equilibrium conditions from the viewpoint of the trip-maker's behaviors. Second, it can formulate the user equilibrium of multi-route and multi-period highway systems. Third, the specification for user equilibrium should have the solution that is not necessary to be a unique one. Finally, it should be an expression that allows the analysis of congestion pricing in combination with the production process in equation (1) and the congestion process in equation (2).

We first explain our specification with a simple hypothetical market characterized as follows. Suppose only the single trip alternative serves trip-makers. Suppose also that the demand for the alternative, denoted by Q, is a function of real trip cost $C + R$, where C refers to the average user cost, and R is the congestion charge imposed by the supplier. Suppose further that the average user cost function F has a single capacity variable y.

In addition, suppose that the supplier has already made the decisions about congestion charge R and capacity y. Note that these decisions are a prerequisite to properly defining the user equilibrium. Moreover, the congestion pricing issue dealt with later involves searching for the optimal congestion charge and capacity.

Under these assumptions, one way of specifying the user equilibrium is to express it using the following two equations:

$$g = Q(C + R) \text{ and } C = F(g;y) \qquad (19)$$

The first equation corresponds to the demand function. The second one represents the average user cost function identical to equation (2). The user equilibrium refers to the state characterized by the solutions C^* and g^*, which satisfy the two equations simultaneously.

The above specification has long been used to explain the structure of the user equilibrium[6]. In addition, this specification has the advantage of satisfying the first three requirements mentioned above. Further, this specification has the structure that the existence of the user equilibrium can readily be approached, as examined later in this section. However, this specification is not an appropriate expression to fulfill the last requirement.

For this reason, we combine these two equations into one. This alternative approach yields the following specification for user equilibrium:

$$C^* = F(Q(C^* + R);y) \qquad (20)$$

Here, the left side C refers to the actual average user cost determined by the average user cost function, which is the outcome of the interaction between the highway facility and its users. The right side C in the demand function corresponds to the average user cost perceived by the trip-makers prior to their trip. Therefore, the user equilibrium defined above refers to a state in which the users' expected value of C prior to their trip equals their experienced value of C in a trip.

The proof for the existence of user equilibrium defined in equation (20) is worked out by introducing an algorithm to search C^* that satisfies the two equations (19) simultaneously. One candidate algorithm involves mapping Γ_0 to update the value of C through the iteration defined below:

$$C_{n+1} = 0.5[C_n + F(Q(C_n + R)y)] \quad (21)$$

On repeating a sufficiently large number of iterations, the value of C_n will approach a certain limit, and will satisfy the following:

$$C_{n+1} \cong C_n \cong 0.5[C_n + F(Q(C_n + R);y)] \quad (22)$$

This resulting relation is equivalent to the expression of the user equilibrium in equation (20). The search process defined in equation (21) is illustrated schematically in Figure 2.

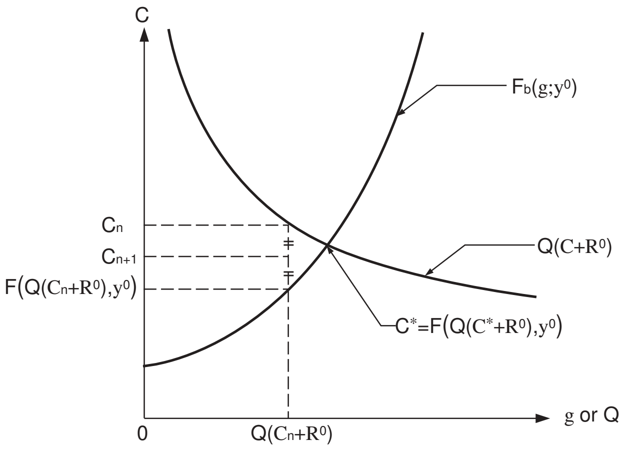

Fig. 2 Schematic illustration of user equilibrium

Subsequently, we examine the problem to specify the user equilibrium for the multiple routes connecting the two points. We approach this problem in two steps. The first step is to set up an equation, similar to equation (11), for each link and for each period respectively. The second step is to formulate the system-wide user equilibrium as a set of simultaneous equations, each of which corresponds to the one developed in the first step.

This approach can alternatively be stated as follows. Suppose that the trip demand Q_{mt}, for every $m \in \langle 1, M \rangle$ and $t \in \langle 1, T \rangle$, is a function of user's real trip costs, $\mathbf{C} + \mathbf{R} \equiv (C_{11} + R_{11}, C_{12} + R_{12}, \cdots, C_{MT} + R_{MT})$, and the after-tax prices of ordinary goods, $\mathbf{P} \equiv (P_1, P_2, \cdots, P_J)$, where the items of goods consumed $\mathbf{z} \equiv (z_1, z_2, \cdots, z_J)$ are identical to those of input $\mathbf{x}$ as defined in equation (1). Suppose also that the supplier has already set up the congestion charge $\mathbf{R_m} \equiv (R_{m1}, R_{m2}, \cdots, R_{mT})$ and capacity y_m for every m.

Then, the existence of the user equilibrium implies that there exists at least one solution $\mathbf{C}^* \equiv (C^*_{11}, C^*_{12}, \cdots, C^*_{MT})$, which satisfies the following $M \times T$ equations simultaneously:

$$C^*_{mt} = F_{mt}(Q_{mt}(\mathbf{C}^* + \mathbf{R}, \mathbf{P})/\alpha_t; y_m),$$
$$\forall m \in \langle 1, M \rangle, \forall t \in \langle 1, T \rangle \quad (23)$$

where Q_{mt}/α_t corresponds to the same quantity with g_t in equation (2).

The existence of user equilibrium for the specification (23) calls for a number of assumptions. First, every demand function Q_{mt} is single-valued and continuous in $\mathbf{C} + \mathbf{R}$. Second, the function Q_{mt} is decreasing with respect to $C_{mt} + R_{mt}$, but non-decreasing with respect to $C_{kl} + R_{kl}$, for every $k \neq m$ or $l \neq t$. Third, the value of each C_{mt} belongs to the closed interval $IC_{mt} \equiv [C^{\min}_{mt}, C^{\max}_{m}]$, where $C^{\min}_{mt} \equiv F_{mt}(0;y_m)$ is the travel cost in the free flow state that $Q_{mt} = 0$. Fourth, each R_{mt} also belongs to the closed interval $[0, R^{\max}_{m}]$, where $R^{\max}_{m}$ is a positive finite value.

We prove the existence of user equilibrium $\mathbf{C}^*$ specified in equation (23) by applying Brower's fixed point theorem. To this end, we define a mapping Γ: $\mathbf{IC} \rightarrow \mathbf{IC}$ that:

$$\Gamma(\mathbf{C}) \equiv \Gamma_{11}(\mathbf{C}) \times \Gamma_{12}(\mathbf{C}) \times \cdots \times \Gamma_{MT}(\mathbf{C}) \text{ and}$$

$$\Gamma_{mt}(\mathbf{C}) \equiv \min\{0.5[C_{mt} + F_{mt}(Q_{mt}(\mathbf{C}+\mathbf{R}, \mathbf{P})/\alpha_t; y_m)]C^{\max}_{m}\} \quad (24)$$

where $\mathbf{IC} \equiv IC_{11} \times IC_{12} \times \cdots \times IC_{MT}$. Then, the function Γ_{mt} is single-valued and continuous, since each F_m and Q_{mt} is single-valued and continuous. Accordingly, the function Γ is also single-valued and continuous. Therefore, by Brower's fixed point theorem, there is a fixed point corresponding to the solution to the simultaneous equation system (23) as well as the equation system (24).

3. THE FIRST-BEST SOCIAL OPTIMALITY

This section deals with the first-best social optimality for multi-route and multi-period highway systems in the presence of congestion. This optimality corresponds to the solution for the social welfare maximization problem faced by government, namely:

Objective : social welfare function
Constraints : production possibilities of capacity,
user equilibrium conditions, and
financial constraint (25)

The details of this maximization problem are explained below.

We define the social welfare function in an identical manner with the one introduced to assess the marginal cost pricing by Mohring[3]. This social welfare function $W(u^1, u^2, \cdots, u^I)$ is monotonically increasing and concave in the indirect utility function u^i, for every individual $i \in \langle 1, I \rangle$. Every consumer i has a neo-classical utility function, and owns an initial endowment m_i comprised of monetary income and other human resources. Each consumer i has to pay an income tax h^i imposed by the government. In addition, every individual i maximizes his utility by consuming private good $\mathbf{z^i} \equiv (z_1^i, z_2^i, \cdots, z_J^i)$ and making trips $\mathbf{q^i} \equiv (q_{11}^i, q_{12}^i, \cdots, q_{MT}^i)$ under the budget constraint of his disposable income $m^i - h^i$.

Subsequently, we make two assumptions about the supply side of ordinary goods consumed $\mathbf{z}$. First, every good $j \in \langle 1, J \rangle$ is supplied by competitive producers at the before-tax price of p_j, being equal to the marginal cost denoted by MC_j. Second, the government imposes an excise tax on every good j, which amounts to v_j. Therefore, the after-tax price of j is equal to $p_j + v_j$, and is denoted by P_j.

We set up the utility maximization problem for each consumer i. The decision variables of this consumer problem are nonnegative $\mathbf{q^i}$ and $\mathbf{z^i}$. The Lagrangian of this decision problem is:

$$U^i \equiv \max u^i\,(\mathbf{q^i}, \mathbf{z^i}) + \eta^i \Big[m^i - h^i - \sum_m \sum_t (C_{mt} + R_{mt}) - \sum_j q^i_{mt} P_j z^i_j \Big] \quad \cdots\cdots (26)$$

where η^i is the marginal utility of income. On arranging the necessary conditions for this maximization problem, we can have the following equation system:

$$\begin{bmatrix} \mathbf{s^i_{qq}} & \mathbf{s^i_{qz}} \\ \mathbf{s^i_{zq}} & \mathbf{s^i_{zz}} \end{bmatrix} \begin{bmatrix} \mathbf{C + R} \\ \mathbf{p + v} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad \cdots\cdots (27)$$

where, the left-side matrix corresponds to a Slutsky substitution matrix for a bundle of $(\mathbf{q}, \mathbf{z}) \equiv (q_{11}, q_{12}, \cdots, q_{MT}, z_1, z_2, \cdots, z_J)$, and each element of the matrix is a submatrix such that $[\mathbf{s^i_{qq}}] \equiv \{s^i_{kl,mt}\}$, $[\mathbf{s^i_{qz}}] \equiv \{s^i_{kl,j}\}$, $[\mathbf{s^i_{zq}}] \equiv \{s^i_{n,mt}\}$ and $\mathbf{s^i_{zz}} \equiv \{s^i_{n,j}\}$. For example, the (kl, mt) element of sub-matrix $[\mathbf{s^i_{qq}}]$ or $\{s^i_{kl,mt}\}$ is $s^i_{kl,mt} = \partial q^i_{mt} / \partial R_{kl} - q^i_{kl} \cdot \partial q^i_{mt} / \partial h^i$. The vector $\mathbf{C + R}$ represents the real cost of $\mathbf{q}$ including user cost $\mathbf{C}$, and $\mathbf{p + v}$ corresponds to the after-tax price of $\mathbf{z}$ (See Appendix A).

In addition, we make four sets of assumptions about the constraints of the social welfare maximization problem. First, the production function of capacities, denoted by G_m, for every m, satisfies a number of requirements specified in Section 2.1. Second, the average user cost function F_{mt} fulfills a number of assumptions identified in Section 2.1. Third, the trip demand function Q_{mt} fulfills the requirements that are necessary for the existence of a user equilibrium $\mathbf{C}$ specified in equation (23). Fourth, the decision of the government supplying all transport services fulfills the balanced budget constraint expressed by:

$$\sum_i h^i + \sum_m \sum_t R_{mt} Q_{mt} + \sum_j v_j Z_j - \sum_m \sum_j p_j x_{mj} \geq 0 \quad \cdots\cdots (28)$$

where $Z_j \equiv \sum z^i_j$. In equation (28), $\sum h^i$ represents the sum of income taxes, $\sum\sum R_{mt}\, Q_{mt}$ refers to the sum of congestion charge revenues, $\sum v_j Z_j$ corresponds to the sum of excise taxes on goods consumed, and $\sum\sum p_j\, x_{mj}$ is the production cost necessary for the provision of all the highway services.

At this stage, it would be necessary to clarify the expression of the marginal cost function with the variable $\mathbf{Q_m}$ instead of $\mathbf{g_m}$. By definition, the value of $\alpha_t g_{mt}$ is equal to that of Q_{mt}. Therefore, equation (7) can alternatively be expressed as:

$$MC_{mt}\,(\mathbf{Q_m}/\mathbf{\acute{a}};\mathbf{p}) = AUC_{mt}(\mathbf{Q_m}/\mathbf{\acute{a}};\mathbf{p}) + MCC_{mt}(\mathbf{Q_m}/\mathbf{\acute{a}};\mathbf{p}) \quad \cdots\cdots (29)$$

where $\mathbf{Q_m}/\mathbf{\acute{a}} \equiv (Q_{m1}/\alpha_1, Q_{m2}/\alpha_2, \cdots, Q_{mT}/\alpha_T)$. The specific expression of the marginal congestion cost function MCC_{mt} is:

$$MCC_{mt}\,(\mathbf{Q_m}/\mathbf{\acute{a}};\mathbf{p}) \equiv Q_{mt} \left.\frac{\partial F_{mt}(Q_{mt}/\alpha_t; y_m)}{\partial Q_{mt}}\right|_{y_m = \hat{y}_m(\mathbf{g};\mathbf{p})} \quad \cdots\cdots (30)$$

We are now in a position to formulate the social welfare maximization problem defined in equation (25). This decision problem involves choosing optimal values of $\mathbf{x_m}$, y_m $\mathbf{C_m}$, $\mathbf{R_m}$ for every m, $\mathbf{h}$ and $\mathbf{v}$. The Lagrangian of this welfare problem is:

$$\begin{aligned} SW_1 \equiv \max W\,(u^1, u^2, \cdots, u^I) &+ \sum_m \phi_m \big(G_m(\mathbf{x_m}) - y_m\big) \\ &+ \sum_m \sum_t \lambda_{mt} \alpha_t \big(C_{mt} - F_{mt}(Q_{mt}/\alpha_t; y_m)\big) \quad \cdots\cdots (31) \\ &+ \omega\Big(\sum_i h^i + \sum_m \sum_t R_{mt} Q_{mt} + \sum_j v_j Z_j - \sum_m \sum_j p_j x_{mj}\Big) \end{aligned}$$

where ϕ_m, λ_{mt} and ω are nonnegative Lagrangian coefficients.

On arranging the first-order conditions with respect to various decision variables yields the two types of relationships characterizing the optimal solution. One group of relationships is the *efficient production conditions*, which characterize the decision rules for the optimal highway capacity. Another group involves the *optimal pricing and taxation rules* associated with the provision of highway services.

The efficient production conditions correspond to the expressions characterizing the optimal capacity. One alternative to expressing the conditions could be:

$$\lambda_{mt}/\omega = Q_{mt}/\alpha_t, \qquad \forall t \in \langle 1, T\rangle \quad (32)$$

$$\frac{\partial SC_m}{\partial y_m} = -\sum_t Q_{mt}\frac{\partial F_{mt}}{\partial y_m}, \quad \forall t \in \langle 1, T\rangle \quad (33)$$

where SC_m refers to the supplier cost function defined in equation (4) (See Appendix B).

These conditions represent the relationships between the marginal supplier costs and the marginal user costs. Specifically, equation (32) implies that one unit reduction in average user costs at the t th period calls for the Q_{mt}/α_t units of additional investment to increase capacities. On the other hand, equation (33) shows that the marginal supplier cost with respect to y_m should equal the marginal user cost saving with respect to the same variable, for every m.

The optimal pricing and taxation rules show the structures of the optimal congestion charges and excise taxes. One alternative of expressing the rules could be the equation system similar to equation (27), namely:

$$\begin{bmatrix} \mathbf{S_{qq}} & \mathbf{S_{qz}} \\ \mathbf{S_{zq}} & \mathbf{S_{zz}} \end{bmatrix}\begin{bmatrix} \mathbf{R-MCC} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad (34)$$

where $[\mathbf{S_{qq}}] \equiv \sum[\mathbf{s^i_{qq}}]$, $[\mathbf{S_{qz}}] \equiv \sum[\mathbf{s^i_{qz}}]$, $[\mathbf{S_{zq}}] \equiv \sum[\mathbf{s^i_{zq}}]$, $[\mathbf{S_{zz}}] \equiv \sum[\mathbf{s^i_{zz}}]$, and $\mathbf{R-MCC} \equiv (R_{11}-MCC_{11}, R_{12}-MCC_{12}, \cdots, R_{MT}-MCC_{MT})$ (See Appendix B).

The optimal congestion charge **R** and excise tax **v** should be simultaneous solutions to two different equation systems (27) and (34). Also, the two equation systems are homogeneous functions of degree one, with respect to their left-side column vectors. Therefore, one possible solution to these equation systems is:

$$\beta\begin{bmatrix} \mathbf{C+R} \\ \mathbf{p+v} \end{bmatrix} = (1+\beta)\begin{bmatrix} \mathbf{R-MCC} \\ \mathbf{v} \end{bmatrix}$$

$$\text{or} \quad \begin{bmatrix} \mathbf{R} \\ \mathbf{v} \end{bmatrix} = \begin{bmatrix} (1+\beta)\mathbf{MCC}+\beta\mathbf{C} \\ \beta\mathbf{p} \end{bmatrix} \quad (35)$$

where $\beta \geq 0$ refers to the uniform excise tax rate.

According to equation (35), the uniform excise tax rate β should be applied to all the highway services as well as the consumption goods. This result coincides with the optimal excise taxation rule for consumption goods[7]. Further, this analysis result shows that the tax base for the highway services should include the user cost in the trip. This uniform taxation rule confirms that the excise taxation has to maintain the proportionality of price to marginal cost for all uses of resources[8].

Subsequently, we show that a set of optimal solutions to the welfare maximization problem in equation (31) fulfills the conditions for the first-best social optimality. Our analysis for this issue pinpoints the derivation of the expressions relevant to the characterization of the optimal taxation scheme and resource allocation efficiency.

First, marginal social welfare with respect to income (or income tax) of every individual i is identical, as shown below:

$$MSW^i \equiv \frac{\partial W}{\partial u^i}\frac{\partial u^i}{\partial m^i} = -\frac{\partial W}{\partial u^i}\frac{\partial u^i}{\partial h^i} = \frac{\omega}{1+\beta}, \quad \forall i \in \langle 1, I\rangle \quad (36)$$

where MSW^i stands for marginal social welfare with respect to the income of the consumer i (See Appendix C).

Equation (36) shows the characteristics of the optimal taxation rule for the first-best social optimality. According to this equation, the optimal taxation should fulfill the condition that marginal social welfare of each individual with respect to his income is identical to others. This marginal social welfare is equal to $\omega/(1+\beta)$, where ω is the marginal social welfare with respect to government expenditures, and β represents the uniform excise tax rate applied to all goods and services.

Second, the optimal congestion charge and excise tax satisfy the resource allocation efficiency, namely:

$$\frac{MSW^i_{mt}}{MSW^h_j} \equiv \frac{\partial W}{\partial u^i}\frac{\partial u^i}{\partial q^i_{mt}}\Bigg/\left(\frac{\partial W}{\partial u^h}\frac{\partial u^h}{\partial z^h_j}\right) = \frac{(1+\beta)(C_{mt}+MCC_{mt})}{(1+\beta)p_j} = \frac{MC_{mt}}{MC_j}, \quad \forall i,h \in \langle 1, I\rangle \quad (37)$$

where MC_j refers to the marginal cost of good j supplied by perfectly competitive producers (See Appendix C).

Equation (37) represents the Pareto efficiency condition for highway services. The term MSU^i_{mt}/MSU^h_j refers to the marginal social rate of substitution (*MSRS*) of q_{mt} for z_j, for every m, t and j. The last term presents the rate of product transformation (*RPT*) of q_{mt} for z_j. Therefore, equation (37) depicts the Pareto efficiency that the value of MSRS between q_{mt} and z_j is equal to the value of RPT between them.

Finally, we schematically illustrate the implication of the first-best social optimality conditions depicted above under a number of strong assumptions. First, the economy consists of one person. Second, there is only one input to the highway service, denoted by z. Third, the output, the average trip demand rate, is constant across all periods, and is denoted by q. Fourth, the highway cost function has the structure depicted in Figure 1.

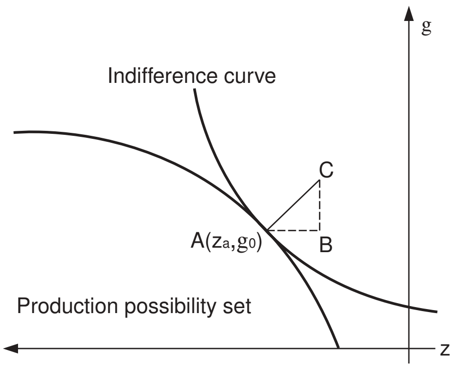

Fig. 3 Schematic illustration of first-best social optimality conditions

For this hypothetical economy, Figure 3 depicts the first-best social optimality conditions. The frontier of the production possibility set in the figure corresponds to the optimal solution to the cost minimization problem in equation (6). Specifically, every point on the frontier represents the optimal value of z necessary for the production of the optimal capacity y, and this capacity has a linear relationship with the given output q in equation (14). The indifference curve corresponds to the social indifference curve for the economy consisting of one person. The point $A \equiv (z_0, q_0)$ depicts the social optimal solution that can attain the maximum social welfare.

At the point A, the production possibility frontier is tangent to the social indifference curve. This condition implies that the values of RPT and $MSRS$ satisfy the following equalities:

$$\frac{RPT\ (q\ for\ z)}{MSRS\ (q\ for\ z)} = \frac{MC_q / MC_z}{MU_q / MU_z} = \frac{BC/AB}{BC/AB} = 1.0 \quad \text{........ (38)}$$

where $MU_q \equiv \partial u/\partial q$ and $MUz \equiv \partial u/\partial z$.

4. THE SECOND-BEST SOCIAL OPTIMALITY

In this section, we assess the congestion charge program to impose a charge upon a single highway link. We approach this policy question with the amended versions of the social welfare maximization problem examined above. We construct the amended maximization problems by modifying the constraint in a manner to reflect the different decision environment. Note also that, for the simplicity of analysis, we formulate the optimization problems under the assumption that the average demand flow rate is constant over the entire period.

Suppose that the government imposes an additional congestion charge R_1 on a major urban arterial link '1' as a measure for alleviating a congestion problem. Suppose also that the government levies the user charges R_m^0, for every $m \in \langle 2, M\rangle$, on other highway links. Assume further that the user charges of other links, R_m^0, do not include any direct charge such as an expressway toll, but includes indirect user charges such as a fuel tax.

Then, the decision problem of the government is to choose optimal values of $\mathbf{x_m}$, y_m, for every m , $\mathbf{h}$ and R_1, for the following welfare maximization problem:

$$\begin{aligned} SW_2 \equiv \max W(u^1, u^2, \cdots, u^I) &+ \sum_m \phi_m\big(G_m(y_m; \mathbf{x_m}) - y_m\big) \\ &+ \sum_m \lambda_m\big(C_m - F_m(Q_m; y_m)\big) \qquad \text{.........................(39)} \\ &+ \omega\Big(\sum_i h^i + R_1 Q_1 + \sum_{m=2}^{M} R_m^0 Q_m - \sum_m \sum_j p_j x_{mj}\Big) \end{aligned}$$

where the last term represents the government's budget constraint.

Proceeding with the analysis using equation (38) along the same lines that lead to equations (32)~(33) yields the following efficient production conditions.

$$\lambda_m/\omega = Q_m \ \text{ and } \quad \text{.. (40)}$$

$$\frac{\partial SC_m}{\partial y_m} = Q_m \frac{\partial F_m}{\partial y_m}, \quad \forall m \in \langle 1, M\rangle \quad \text{.............................. (41)}$$

Similarly, the analysis used to lead to (35) results in the following pricing rule:

$$(R_1 - MCC_1)S_{11} + \sum_{m=2}^{M}(R_m^0 - MCC_m)\, S_{m1} = 0 \quad \text{............... (42)}$$

where S_{m1} refers to the Slusky substitution term.

Accordingly, the optimal value of R_1 can be estimated using the following equation:

$$\begin{aligned} R_1 &= Q_1 \frac{\partial F_1}{\partial Q_1} + \sum_{m=2}^{M} \frac{S_{m1}}{S_{11}}\Big(Q_m \frac{\partial F_m}{\partial Q_m} - R_m^0\Big) \\ &\cong MCC_1 + \sum_{m=2}^{M} \rho_{m1}\,(MCC_m - R_m^0) \end{aligned} \quad \text{...................... (43)}$$

where $\rho_{m1} \equiv \dfrac{\partial Q_m}{\partial R_1} \Big/ \dfrac{\partial Q_1}{\partial R_1} \cong \dfrac{\Delta Q_m}{\Delta R_1} \Big/ \dfrac{\Delta Q_1}{\Delta R_1} \langle\, 0.$ Here, ΔQ_1 represents the amount of decrease in Q_1 by an increase in R_1 of ΔR_1, and ΔQ_m corresponds to the increase of Q_m, for every $m \in \langle 2, M\rangle$, increased by the increase in R_1.

Therefore, ρ_{m1} can be interpreted as the diversion ratio of ΔQ_1 to other competing highway services.

The optimal congestion charge in equation (43) differs from the one derived from the net benefit maximization problem for a highway link in other studies[1,2]. The congestion charge estimated in those previous studies does not include the second term on the right side of equation (43), $\sum \rho_{m1}(MCC_m - R_m^0)$. This additional term can be zero, only if the user charge R_m^0, for every $m \in \langle 2, M \rangle$, is equal to MCC_m, the optimal user charge.

To obtain more specific information, we make another assumption that the traffic diversion to public transit is negligible. Then, the sign of $\sum \rho_{m1}(MCC_m - R_m^0)$ is surely negative, since the ρ_{m1} values for competing highway links are negative, and the $MCC_m - R_m^0$ values are positive. Accordingly, the congestion charge imposed only upon highway link '1' should be smaller than the MCC_1 value, as long as alternative highway links are available to the trip-makers.

This interpretation of equation (43) could be applicable for assessing the two alternative implementation programs of an urban congestion toll. One alternative is to impose the congestion toll on the user of one or two from the many arterial roads connecting to the CBD. Another alternative is to collect the congestion toll from trip-makers using all the roads to the CBD, and/or to add the congestion toll on the parking fee inside the CBD.

The first alternative has a high potential to yield an optimal congestion toll smaller than the *MCC* value with a large margin, as indicated earlier. Therefore, the net welfare gain of this congestion toll program would not be significant. Also, the congestion toll equal to the *MCC* value would cause an excessive amount of diverted traffic to other arterial roads exempted from the congestion toll, and thus the consequence would be worse-off than the optimal toll estimated by using equation (43).

On the other hand, the second alternative could have the optimal congestion toll, which approximates the *MCC* value. Such a congestion toll program has the potential to improve the resource allocation efficiency close to the maximum level. Therefore, it would be fair to say that the second alternative is a more effective policy measure to improve economic efficiency.

6. CONCLUDING REMARKS

So far, we have examined the two different social welfare maximization problems, in order to obtain the decision rules for the optimal investment and pricing of highway systems in the presence of congestion. One welfare maximization problem involves searching for the decision rules that can attain the first-best social optimality. Another is to find the second-best optimal solutions, which can provide useful information about the congestion charge scheme to impose congestion tolls on a highway link.

We have confirmed that the analysis outputs are compatible with the results for a highway link in previous studies dealing with a congestion pricing scheme. More importantly, we have been able to obtain more precise expressions about the optimal decision rules with the analysis framework of general equilibrium. Associated with this advantage, we introduce some of the findings and their economic implications below.

One of the main analysis topics involves efficient production conditions, which characterize the decision rule for the optimal highway capacity. The analysis result shows that the efficient production conditions for the first-best solution are identical with those for the second-best solutions. These conditions are:

- Providing that the optimal capacity is chosen, one unit reduction in average user costs at the t th period calls for the Q_t/α_t units of monetary investments to increase the highway capacity; and
- The marginal supplier cost with respect to the capacity of a certain highway link should be equal to the marginal user cost saving attainable by one unit increase in the capacity.

Another analysis topic has been the pricing rule for highway services. This analysis involves searching for the congestion charge that can reach maximum social welfare, under a certain constraint associated with financial requirements. Of course, the analysis result for the different social welfare maximization problems yields the different expressions for the optimal congestion charges. In spite of that, all the optimal user charges share the following common property: they are the functions of the marginal congestion cost, *MCC* ($\equiv Q \cdot \partial F/\partial Q$).

Specifically, the user charge for the first-best social optimality equal to $(1 + \beta)$ *MCC* + βC for all highway services, where β is an uniform excise tax rate applied to highway services as well as ordinary goods, and C is the average user cost. This result suggests that the government would be better to impose the excise tax, β (*MCC* + C), on all the highway services to enhance resource allocation efficiency.

Also, the second-best optimal congestion charge examined in Section 4 is expressed with the *MCC* terms of all the available highway links. Like the first-best optimal solution, this optimal congestion charge is the outcome of analyses comprehending their impacts on the demands for other competing routes. Accordingly, the pricing rules listed in Section 4 could be more appropriate guidelines for the corresponding policy questions.

In addition, the second-best pricing rule in equation (43) provides a useful guideline in designing the implementation program of congestion charges. According to this result, the congestion toll imposed on one or two links from many arterial roads connecting to the CBD would not be an effective measure to improve the resource allocation efficiency. To avoid this problem, it would be desirable to impose the congestion toll upon the users of all roads inside the congested area.

REFERENCES

1. Keeler, T.E. and Small, K.A. Optimal peak-load pricing, investment, and service levels on urban expressways. "Journal of Political Economy" 85(1): pp. 1-25 (1977).
2. Small, K.A. Urban Transportation Economics. Harwood Academic Publishers pp. 106-111 (1995).
3. Mohring, H. The peak load problem with increasing returns and pricing constraints. "American Economic Review" 60: pp. 693-705 (1970).
4. Diamond, P.A. and Mirrles, J.A. Optimal taxation and public production I: production efficiency. "American Economic Review" 61: pp. 8-27 (1971).
5. Transportation Research Board. Highway Capacity Manual (HCM 2000). Washington D.C. Chapter 30 (2000).
6. Manheim M.L. Fundamentals of transportation systems analysis. The MIT press, Cambridge MA (1982).
7. Dixit, A. On the optimal structure of commodity taxes. "America Economic Review" 60: pp. 295-301 (1970).
8. Lerner, A.P. On the optimal taxes with an untaxable sector. "American Economic Review" 60: pp. 284-294 (1970).

APPENDIX: ESTIMATION OF THE FIRST-BEST OPTIMAL SOLUTIONS

Here, we show the detailed derivation of optimal pricing and investment decision rules for multi-route highway systems from the social welfare maximization problem in equation (31).

A. *Proof of Equation (27)*

The utility maximization problem of consumer i is:

$$U^i \equiv \max\ u^i(\mathbf{q^i}, \mathbf{z^i}) + \eta^i [m^i - h^i - \sum_m \sum_t (C_{mt} + R_{mt}) - \sum_j q^i_{mt} p_j z^i_j] \qquad (26)$$

The first-order conditions of this maximization problem are:

$$\frac{\partial u^i}{\partial q^i_{mt}} - \eta^i (C_{mt} + R_{mt}) = 0,\ \forall m \in \langle 1, M\rangle,\ \forall t \in \langle 1, T\rangle \qquad (A.1)$$

$$\frac{\partial u^i}{\partial z^i_j} - \eta^i P_j = 0 \qquad \forall m \in \langle 1, M\rangle \qquad (A.2)$$

$$m^i - h^i - \sum_m \sum_t (C_{mt} + R_{mt})\, q^i_{mt} - \sum_j p_j z^i_j = 0 \qquad (A.3)$$

Solving the above equation system yields the demand functions for q^i_{mt} and z^i_j, denoted by $\hat{q}^i_{mt}$ and $\hat{z}^i_j$, respectively. Substituting these demand functions into the utility function, and differentiating this indirect utility function obtained in the previous step with respect to the government's decision variables results in:

$$\frac{\partial u^i}{\partial h^i} = -\eta^i,\ \frac{\partial u^i}{\partial R_{mt}} = -\eta^i q^i_{mt},\ \frac{\partial u^i}{\partial v^i} = -\eta^i z^i_j \qquad (A.4)$$

for every m, t and j.

On the other hand, substituting the demand functions $\hat{q}^i_{mt}$ and $\hat{z}^i_j$ into equation (A.3), and differentiating the resulting equation with respect to the government's decision variables yields:

$$1 + \sum_m \sum_t (C_{mt} + R_{mt}) \frac{\partial \hat{q}^i_{mt}}{\partial h^i} + \sum_j P_j \frac{\partial \hat{z}^i_j}{\partial h^i} = 0, \quad \forall i \in \langle 1, I\rangle \qquad (A.5)$$

$$\hat{q}_{kl}^i + \sum_m \sum_t (C_{mt} + R_{mt}) \frac{\partial \hat{q}_{mt}^i}{\partial R_{kl}} + \sum_j P_j \frac{\partial \hat{z}_j^i}{\partial R_{kl}} = 0, \quad \forall k \in \langle 1, M \rangle, \forall l \in \langle 1, T \rangle \qquad \text{(A.6)}$$

$$\sum_m \sum_t (C_{mt} + R_{mt}) \frac{\partial \hat{q}_{mt}^i}{\partial v_n} + \hat{z}_n^i + \sum_j P_j \frac{\partial \hat{z}_j^i}{\partial v_n} = 0, \quad \forall n \in \langle 1, J \rangle \qquad \text{(A.7)}$$

Further, these 3 sets of equation systems can be rearranged as:

$$\begin{aligned}
&(\text{A.6}) - \hat{q}_{kl}^i \times (\text{A.5}) \\
&= \sum_m \sum_t (C_{mt} + R_{mt}) \left(\frac{\partial \hat{q}_{mt}^i}{\partial R_{kl}} - \hat{q}_{kl}^i \frac{\partial \hat{q}_{mt}^i}{\partial h^i} \right) \\
&+ \sum_j P_j \left(\frac{\partial \hat{z}_j^i}{\partial R_{kl}} - \hat{q}_{kl}^i \frac{\partial \hat{z}_j^i}{\partial h^i} \right) = 0 \qquad \text{(A.8)}
\end{aligned}$$

$$\begin{aligned}
&(\text{A.7}) - \hat{z}_n^i \times (\text{A.5}) \\
&= \sum_m \sum_t (C_{mt} + R_{mt}) \left(\frac{\partial \hat{q}_{mt}^i}{\partial v_n} - \hat{z}_n^i \frac{\partial \hat{q}_{mt}^i}{\partial h^i} \right) \\
&+ \sum_j P_j \left(\frac{\partial \hat{z}_j^i}{\partial v_n} - \hat{z}_n^i \frac{\partial \hat{z}_j^i}{\partial h^i} \right) = 0 \qquad \text{(A.9)}
\end{aligned}$$

Expressing equations (A.8) and (A.9) in matrix, we have equation (27).

B. Proof of Equations (32)~(34)

The welfare maximization problem to be analyzed is:

$$\begin{aligned}
SW_1 \equiv \max\ & W(u^1, u^2, \cdots, u^I) + \sum_m \phi_m \left(G_m(\mathbf{x_m}) - y_m \right) \\
&+ \sum_m \sum_t \lambda_{mt} \alpha_t \left(C_{mt} - F_{mt}(Q_{mt} / \alpha_t; y_m) \right) \qquad \text{(A.10)} \\
&+ \omega \left(\sum_i h^i + \sum_m \sum_t R_{mt} Q_{mt} + \sum_j v_j Z_j - \sum_m \sum_j p_j x_{mj} \right)
\end{aligned}$$

For this maximization problem, the first-order conditions with respect to x_{mj} and y_m are:

$$\phi_m \frac{\partial G_m}{\partial x_{mj}} - \omega p_j = 0, \qquad \text{(A.11)}$$

$$\phi_m + \sum_t \lambda_{mt} \alpha_t \frac{\partial F_{mt}}{\partial y_m} = 0, \qquad \text{(A.12)}$$

The first-order conditions with respect to h^i, R_{kl}, C_{kl}, and v_n are:

$$\begin{aligned}
\frac{\partial SW_1}{\partial h^i} &= \frac{\partial W}{\partial u^i} \frac{\partial u^i}{\partial h^i} - \sum_m \sum_t \lambda_{mt} \alpha_t \frac{\partial F_{mt}}{\partial Q_{mt}} \frac{\partial \hat{q}_{mt}^i}{\partial h^i} \qquad \text{(A.13)} \\
&+ \omega \left(1 + \sum_m \sum_t R_{mt} \frac{\partial \hat{q}_{mt}^i}{\partial h^i} + \sum_j v_j \frac{\partial \hat{z}_j^i}{\partial h^i} \right) = 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial SW_1}{\partial R_{kl}} &= \sum_i \frac{\partial W}{\partial u^i} \frac{\partial u^i}{\partial R_{kl}} - \sum_m \sum_t \lambda_{mt} \alpha_t \frac{\partial F_{mt}}{\partial Q_{mt}} \frac{\partial Q_{mt}}{\partial R_{kl}} \qquad \text{(A.14)} \\
&+ \omega \left(Q_{kl} + \sum_m \sum_t R_{mt} \frac{\partial Q_{mt}}{\partial R_{kl}} + \sum_j v_j \frac{\partial Z_j}{\partial R_{kl}} \right) = 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial SW_1}{\partial C_{kl}} &= \sum_i \frac{\partial W}{\partial u^i} \frac{\partial u^i}{\partial C_{kl}} + \lambda_{kl} \alpha_t - \sum_m \sum_t \lambda_{mt} \alpha_t \frac{\partial F_{mt}}{\partial Q_{mt}} \frac{\partial Q_{mt}}{\partial C_{kl}} \qquad \text{(A.15)} \\
&+ \omega \left(\sum_m \sum_t R_{mt} \frac{\partial Q_{mt}}{\partial C_{kl}} + \sum_j v_j \frac{\partial Z_j}{\partial C_{kl}} \right) = 0,
\end{aligned}$$

$$\begin{aligned}
\frac{\partial SW_1}{\partial v_n} &= \sum_i \frac{\partial W}{\partial u^i} \frac{\partial u^i}{\partial v_n} - \sum_m \sum_t \lambda_{mt} \alpha_t \frac{\partial F_{mt}}{\partial Q_{mt}} \frac{\partial Q_{mt}}{\partial v_n} \\
&+ \omega \left(\sum_m \sum_t R_{mt} \frac{\partial Q_{mt}}{\partial v_n} + Z_n + \sum_j v_j \frac{\partial Z_j}{\partial v_n} \right) = 0 \qquad \text{(A.16)}
\end{aligned}$$

Replacing $\partial u^i / \partial C_{kl}$ and $\partial Q_{mt} / \partial C_{kl}$ in equation (A.15) with $\partial u^i / \partial R_{kl}$ and $\partial Q_{mt} / \partial R_{kl}$, respectively, and subtracting the resulting equation from equation (A.14) yields:

$$\lambda_{mt} / \omega = Q_{mt} / \alpha_t, \ \forall m \in \langle 1, M \rangle, \ \forall t \in \langle 1, T \rangle \qquad \text{(32)}$$

Substituting equations (32) and equation (A.10) into equation (A.11) yields:

$$p_j \Big/ \frac{\partial G_m}{\partial x_{mj}} = -\sum_t Q_{mt} \frac{\partial F_{mt}}{\partial y_m} \qquad \text{(A.17)}$$

On the other hand, the first-order condition of the supplier decision in equation (4) satisfies the following:

$$p_j \Big/ \frac{\partial G_m}{\partial x_{mj}} = \frac{\partial SC_m}{\partial y_m} \quad \text{(A.18)}$$

where SC_m refers to the supplier cost function. Thus, equation (34) can readily be obtained by combining equations (A.17) and (A.18).

Using equations (A.4) and (33), the optimal conditions in equations (A.13), (A.14) and (A.16) can be restated as:

$$\frac{\partial SW_1}{\partial h^i} = -\frac{\partial W}{\partial u^i}\eta^i + \omega\Big[1 + \sum_m\sum_t (R_{mt} - MCC_{mt})\frac{\partial \hat{q}^i_{mt}}{\partial h^i} + \sum_j v_j \frac{\partial \hat{z}^i_j}{\partial h^1}\,\omega\Big] = 0, \quad \text{(A.19)}$$

$$\frac{\partial SW_1}{\partial R_{kl}} = -\frac{\partial W}{\partial u^i}\eta^i \hat{q}^i_{kl} + \omega\Big[Q_{kl} + \sum_m\sum_t (R_{mt} - MCC_{mt})\frac{\partial Q^i_{mt}}{\partial R_{kl}} + \sum_j v_j \frac{\partial Z_j}{\partial R_{kl}}\Big] = 0 \quad \text{(A.20)}$$

$$\frac{\partial SW_1}{\partial v_n} = -\frac{\partial W}{\partial u^i}\eta^i z^i_n + \omega\Big[\sum_m\sum_t (R_{mt} - MCC_{mt})\frac{\partial Q_{mt}}{\partial R_{kl}} + Z_n + \sum_j v_j \frac{\partial Z_j}{\partial v_n}\Big] = 0 \quad \text{(A.21)}$$

The above three sets of equations can be rearranged as follows:

$$\begin{aligned}
&\text{(A.20)} - \sum_i \hat{q}^i_{kl} \times \text{(A.19)} \\
&= \sum_m\sum_t (R_{mt} - MCC_{mt})\cdot\sum_i\left(\frac{\partial \hat{q}^i_{mt}}{\partial R_{kl}} - \hat{q}^i_{kl}\frac{\partial \hat{q}^i_{mt}}{\partial h^1}\right) \\
&\quad + \sum_i\sum_j v_j\left(\frac{\partial \hat{z}^i_j}{\partial R_{kl}} - \hat{q}^i_{kl}\frac{\partial \hat{z}^i_j}{\partial h^i}\right) = 0 \qquad \text{(A.22)}
\end{aligned}$$

$$\begin{aligned}
&\text{(A.21)} - \sum_i \hat{z}^i_n \times \text{(A.19)} \\
&= \sum_m\sum_t (R_{mt} - MCC_{mt})\cdot\sum_i\left(\frac{\partial \hat{q}^i_{mt}}{\partial v_n} - \hat{z}^i_n\frac{\partial \hat{q}^i_{mt}}{\partial h^i}\right) \\
&\quad + \sum_i\sum_j v_j\left(\frac{\partial \hat{z}^i_j}{\partial v_n} - \hat{z}^i_n\frac{\partial \hat{z}^i_j}{\partial h^i}\right) = 0 \qquad \text{(A.23)}
\end{aligned}$$

Expressing these two equation systems in matrix becomes equation (34).

C. Proof of Equations (36) and (37)

Firstly, the proof of equation (36) is worked out by combining the previous analysis results, as shown below:

$$\begin{aligned}
MSW^i &\equiv \frac{\partial W}{\partial u^i}\frac{\partial u^i}{\partial m^i} = -\frac{\partial W}{\partial u^i}\frac{\partial u^i}{\partial h^i} \\
&= \omega\left(1 + \sum_m\sum_t (R_{mt} - MCC_{mt})\frac{\partial \hat{q}^i_{mt}}{\partial h^i} + \sum_j v_j \frac{\partial \hat{z}^i_j}{\partial h^i}\right) \\
&= \omega\left[1 + \frac{\beta}{1+\beta}\left(\sum_m\sum_t (C_{mt} + R_{mt})\frac{\partial \hat{q}^i_{mt}}{\partial h^i} + \sum_j P_j \frac{\partial \hat{z}^i_j}{\partial h^1}\right)\right] \\
&= \frac{\omega}{1+\beta},\ \forall i \in \langle 1, I\rangle \qquad \text{(36)}
\end{aligned}$$

In the above, the first equality is obtained by substituting equation (A.19). The second and last equalities are derived by using equations (24) and (A.5) respectively.

Subsequently, the proof of equation (37) is approached in the following manner:

$$\begin{aligned}
\frac{MSW^i_{mk}}{MSW^h_j} &\equiv \frac{\partial W}{\partial u^i}\frac{\partial u^i}{\partial q^i_{mt}}\Big/\left(\frac{\partial W}{\partial u^h}\frac{\partial u^h}{\partial z^h_j}\right) \\
&= \frac{\partial W}{\partial u^i}\eta^i (C_{mt} + R_{mt})\Big/\left(\frac{\partial W}{\partial u^h}\eta^h P_j\right) \\
&= \frac{(C_{mt} + MCC_{mt})}{p_j} \qquad \text{(37)} \\
&= \frac{MC_{mt}}{MC_j}
\end{aligned}$$

Here, the first equality is obtained by substituting equation (A.1)