Hitotsubashi Journal of Commerce and Management 47 (2013), pp.17-31.

THE CHARGING PRINCIPLE FOR THE DEVELOPMENT AND MAINTENANCE OF TRANSPORT INFRASTRUCTURE

Toshinori Nemoto*

I. *Introduction*

With the onset of the aging society and low birth rates, coupled with tight financial conditions being experienced by central and local governments, the need to ensure transport services that protect people's lives in rural areas and provide infrastructure that strengthens the global competitiveness of large cities has become increasingly important. In addition, there are high costs associated with the maintenance and renewal of transport infrastructure such as roads and railways (MLIT, 2012). Considering these pressures, a charging principle could be introduced to provide a revenue source for maintaining and renewing this transport infrastructure.

The objectives of this paper are: 1) to clarify the effects of short-run marginal cost pricing and short-run average cost pricing when transport infrastructure capacity is given, 2) to explain how to obtain long-run marginal cost pricing assuming a condition that can optimize transport infrastructure capacity to meet transport demand, and 3) to demonstrate the validity of successive short-run marginal cost pricing through a planning process that gradually increases or reduces transport infrastructure capacity.

II. *Transport System Costs*

It is necessary to define transport system costs to design a charging principle in transport. Transport system costs refer to the costs arising from the construction and maintenance of transport infrastructure and its use. When considering a cost item, it is necessary to differentiate first between fixed and variable costs. Fixed costs are not related to traffic volume, such as construction costs, whereas variable costs are determined according to the amount of traffic, such as maintenance costs. The average cost is determined by dividing the aggregate total cost (the sum of fixed costs and variable costs) by the traffic volume ; and the marginal cost differentiates the variable cost from the traffic volume.

It is also important to pay attention to which stakeholder bears the costs. Broadly speaking, the stakeholders consist of the transport infrastructure manager conducting construction and maintenance, the user using the transport infrastructure, and other stakeholders that can be subjected to positive and/or negative impacts arising from the use of the transport infrastructure although not directly related to its provision and use. Transport infrastructure

* Graduate School of Commerce and Management, Hitotsubashi University

TABLE 1. ROAD SYSTEM COSTS

<table>
<tr><th></th><th></th><th>Fixed costs</th><th>Variable costs
(related to traffic volume)</th></tr>
<tr><td rowspan="3">Internal Costs</td><td>Road Manager</td><td>A: Construction cost</td><td>B: Maintenance cost</td></tr>
<tr><td rowspan="2">Road User (i.e., E and F are taxes/charges transferred to the road manager)</td><td>C: Vehicle cost</td><td>D: Time cost, fuel cost</td></tr>
<tr><td>E: Vehicle ownership charges</td><td>F: Gasoline tax, toll charges, distance-based charges</td></tr>
<tr><td>External Costs</td><td>Other Stakeholders</td><td>G: Improvement or destruction of scenery</td><td>H: Congestion, air pollution</td></tr>
</table>

managers and users are assumed to form a quasi-market, and their costs are regarded as internal costs, while costs arising on other stakeholders are regarded as external costs.

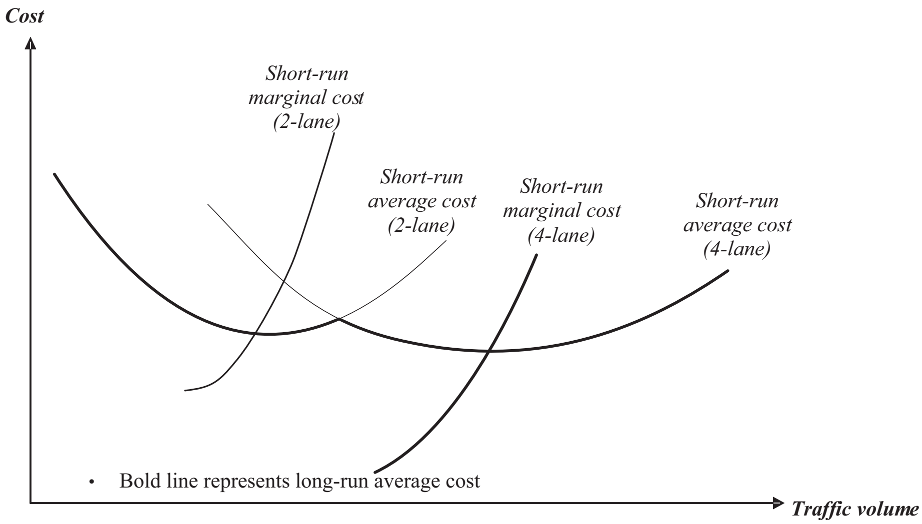
Users of transport infrastructure include the transport companies (e.g. bus companies) that provide transport services to people, where direct users are transport companies and indirect users are people purchasing the services. However when we use roads by private cars, we are direct users and service purchasers at the same time. The transport system is simpler, so that it is easier to look at how taxes and charges are collected for maintenance and renewal of the transport infrastructure. Therefore, this paper concentrates on the case of “road use by private cars.”

Let us suppose that items of transport system costs on the road (i.e. road system costs) can be specifically observed. The road user may use the road by paying C, D, E, and F. Among these, E and F are the taxes and charges that the road user pays for car ownership and use. These are transferred to the road manager, and are appropriated for the maintenance of roads. These charges serve as revenue to the road manager, and then can be cancelled resulting in no social costs. The ultimate costs remaining are A, B, C, D, G, and H. This sum is the total costs.

The charges and taxes that the users pay are not all appropriated for the maintenance of the transport infrastructure. Historically, there exists an earmarked financing system that uses taxes and charges from users for its maintenance and renewal. However, the number of countries using this system has been reduced. Some examples that remain are the federal gasoline tax of the U.S., the energy tax of Germany, and the axle tax of France. While new road construction is difficult partly owing to people’s environmental concerns, the Ministry of Finance, which wants to increase vehicle-related taxes as a source of revenue, abolished the earmarked financing system, and has been using vehicle-related taxes as a general revenue source. The most recent European transport white paper (EC, 2011) states that “it is important to establish a financial system in which revenue from transport users is used for the improvement of transport.” This paper concurs with that position and aims at matching each stakeholder’s benefits and burdens.

Furthermore, it is necessary to distinguish between short-run and long-run costs. Short-run costs arise when the given transport infrastructure does not change. Table 1 distinguishes between fixed costs and variable costs, although these are both for the short-term. In the long run, wherein transport infrastructure capacity may change through road widening and road network extension, all costs become variable costs. Therefore, for short-run optimization, given the present transport infrastructure, traffic can be rationalized by introducing charges. On the other hand, for long-run optimization, a policy is examined to rationalize transport infrastructure capacity given future traffic volumes.

Fig. 1 Short-Run Average Cost, Short-Run Marginal Cost, and Long-Run Average Cost



The cost functions for two-lane and four-lane roads connecting two cities are illustrated in Fig. 1. As mentioned above, the average cost is the value equal to the sum of the fixed cost and variable cost divided by traffic volume. The fixed cost divided by traffic volume decreases as traffic volume increases. However, since congestion arises and time cost increases as traffic increases, the average cost also increases (i.e. downward convex).

The marginal cost is the increase in the total social cost caused by unit increases in traffic. Since the increase in time cost of society caused by additional users in congested periods may surpass by far the increase in the concerned user's time cost, the marginal cost curve rapidly increases after crossing the lowest point of the average cost curve.

Although the average cost curves of two-lane and four-lane roads are of the same shape, the cost curve for the four-lane road is shifted towards the lower right. As mentioned later, economies of scale may be assumed at these numbers of lanes. The long-run average cost curve is an envelope curve connecting the bottoms of the short-run average cost curves. If the number of lanes and road capacity increases, the long-run average cost decreases.

For transport infrastructure, 'indivisibilities,' which means that the number of lanes of a road must be an integer, may become obstacles during road capacity optimization. In Fig. 1, the long-run average cost cannot be differentiated at the intersection. However, as far as the road is concerned, the road capacity is not only a function of the number of lanes, but also a function of the width of each lane and the width of the shoulder. Conversely, it seems that the road capacity can be increased almost continuously if the maximum roadway capacity can be secured and designed under a certain width of the street facility.

III. *Short-Run Marginal Cost Pricing that Realizes Optimal Traffic Given the Transport Infrastructure*

Let us examine the charging principle given the transport infrastructure. Corresponding to the short-run decision-making problem in economics, price setting by short-run marginal cost (or for accuracy, short-run marginal social cost containing external costs) can realize optimal usage of a transport infrastructure. That is, traffic volume that maximizes social benefits can be realized under the road capacity at that time.

Marginal cost pricing is explained using Fig. 2. In the case of the two-lane short-run average cost and the short-run marginal cost curves shown in Fig. 1, a short-run user average cost curve can be added. The user average cost is equal to the user cost (C and D of Table 1) divided by the traffic volume. Of the user average cost, the time cost required for travel becomes the major cost (average time value of the road user is about 40 yen/minute). Furthermore, the high and low demand curves are added.

If there are no taxes and charges, since the road can be used by payment of the short-run user average cost, the road may be used up to Q_1 corresponding to intersection A of the high demand curve (traffic volume is equivalent to Q_2 at low demand). At this traffic volume, however, total cost increases by the short-run marginal cost shown in the figure due to the increase in the number of concerned road users. Social loss is produced between the short-run marginal cost curve and the demand curve (the triangle ABC).

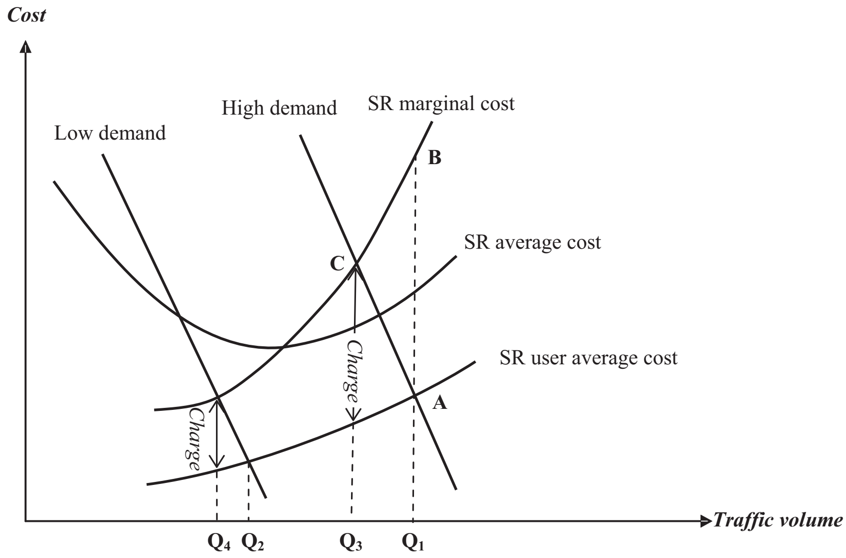
The amount of traffic decided by the intersection of the demand curve and the short-run marginal cost curve (Q_3 in the case of high demand and Q_4 in the case of low demand) becomes the optimal amount of traffic. At this point, the benefit measured in willingness-to-pay is equal to the additional cost that society pays. To realize this amount of traffic, the deviating portion of the short-run marginal cost and the short-run user average cost is charged, so that the users would face an upwards-shifted short-run user average cost curve. This charging principle is called marginal cost pricing.

The weakness of short-run marginal cost pricing is that the costs of the transport infrastructure cannot all be provided. As shown in the figure, although the charge is higher than the short-run average cost at high demand, it is smaller than the short-run average cost at low demand. As with many transport infrastructures, congested time periods and areas are fewer than non-congested time periods and areas. As a result, the transport infrastructure manager will suffer big deficits with marginal cost charging.

It will be necessary to use other taxes to compensate for the fund shortage arising from marginal cost pricing. Although supplementation by a poll tax that does not distort resource allocation is recommended, its introduction is generally difficult from the viewpoint of fairness. Where priority is given to expenditures on social welfare and pensions, there is little prospect of increased expenditure on transport infrastructure from general revenue.

The European White Paper on Fair Payment for Infrastructure Use (EC, 1998) proposes the "promotion of fair competition" and "revenue generation by internalizing external costs" as two reasons for supporting marginal cost pricing. The former is the proposal that marginal cost charging should be adopted as a standard for the EU in order to prevent international transport service providers from suffering a competitive disadvantage because of their nationality. At present, each country imposes different taxes and charges for each transport mode, which

FIG. 2 SHORT-RUN MARGINAL COST PRICING



influences the international transport service providers favorably or unfavorably. However, the peripheral states of the EU such as those in Eastern Europe oppose the standardization of charges on congestion, a typical externality. They regard the congestion charges as unfair because the trucks belonging to these countries have to pay the charges when passing through central Europe, which experiences severe congestion (Nash, 2008).

The latter is the proposal that external costs of congestion should be charged such that the infrastructure costs of the whole transport system can be provided. In the EU, since border customs clearance was abolished in 1993 and international freight traffic (ton-km) has outpaced GDP growth, external costs of congestion and the environment are expected to be larger than the cost for provision of transport infrastructure (including railways). However, criticisms have been raised such as those of Rothengatter (2003) who state "assistance on railways from road pricing revenues does not necessarily bring about optimization of the whole transport system." Surely, the method of appropriating the revenues of a congested mode (or area) to another mode (or area) does not guarantee that the required amount of money can be secured, and the improved efficiency due to marginal cost pricing during the charging stage also entails risks of being ruined in the expenditure stage. Incidentally, the term marginal cost pricing was not used in the subsequent transport White Papers (EC, 2001; EC, 2011).

Marginal cost pricing may be used for certain modes of transportation due to political considerations. In Sweden and the Netherlands, rail is favorably treated to increase its modal share so that railway companies only pay track access fees equal to the marginal cost to the railway infrastructure company. The access fees are only about 5% of total costs. The remaining fixed costs are paid by the government (Ozawa and Nemoto, 2012).

Japanese freight railway company also receives preferential treatment. Freight railway company accepts charges, set by the passenger railway companies that own the railway infrastructure, for the use of the tracks using the avoidable cost principle. Avoidable costs are costs that would not be spent if the freight train were not run. Specifically, the increased portion of the variable cost related to railway infrastructure is computed using the standard unit formula (track access fee unit price multiplied by train-kilometers), and does not interfere substantially with the marginal cost pricing. Such a measure is taken to increase the profitability of freight railways. However, resource allocation will be distorted with the measure in situations wherein passenger railways and other modes have not carried out marginal cost pricing.

IV. *Short-Run Average Cost Pricing that Aims at Balanced Budget Given Transport Infrastructure*

Let us look at average cost pricing as shown in Fig. 3. At high demand, the charge for average cost pricing becomes lower than marginal cost pricing. Traffic volume remains at Q_5, and the traffic could not be reduced to the optimal volume. Conversely, at low demand, since the charge becomes higher, smaller than optimal traffic volume can be realized (Q_6). In other words, in the marginal cost pricing at low demand as previously explained, total cost recovery can be conceded, and traffic can be induced by giving a discount to the marginal cost.

The merit of average cost pricing is that the revenues and expenses of the road manager are balanced. In addition, although it becomes profitable for the road manager if total costs including the external costs against other stakeholders are collected. In this paper, it is assumed for simplification that the collected portion of external costs is outlaid for the costs of relieving externalities (e.g. spending revenue corresponding to noise externalities in construction of noise barriers).

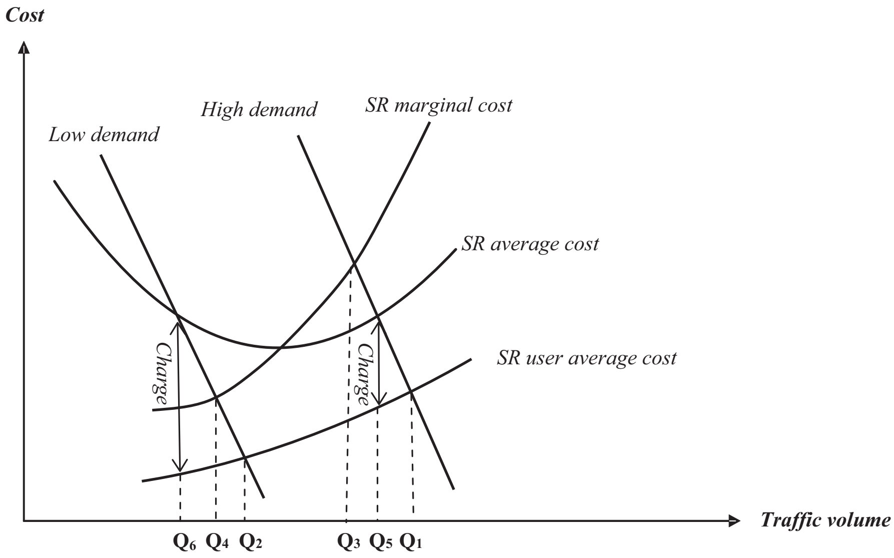
The problem of average cost pricing is that the charge becomes high in areas with little demand. Transport demand may decrease against the transport infrastructure manager's expectations due to factors such as population decrease and factory relocation. If there are few users and the charge is high, traffic will furthermore decrease. In the case that excessive transport infrastructure exists in an area where transport demand has decreased, there are essential solutions for reducing the transport infrastructure capacity. Solutions using charges are limited on the premise of continuously maintaining the present transport infrastructure.

A two-part tariff and Ramsey pricing, which do not reduce traffic greatly in a low demand area, are attractive charging methods to recover total costs even though these methods may not increase social benefit dramatically.

Ramsey pricing is a method of distributing the fixed costs to the users inversely proportional to the charge elasticity of demand after making each user pay the marginal cost calculated according to the user. Fixed costs can be recovered without significantly reducing the total number of users whilst maintaining the percentage composition of users, by reducing the burden of users who are sensitive to charging (e.g. a passenger car is more sensitive than a truck), and increasing the burden of users who are not sensitive to charging. However, the charge for those who have no alternative mode of transport and thus are not sensitive becomes high. From the viewpoint of fairness, therefore, there are almost no practical applications.

The two-part tariff refers to a method of collecting fixed costs as basic charges and collecting variable costs as per unit charges, and does not need to reduce traffic much either.

FIG. 3 SHORT-RUN AVERAGE COST PRICING



After paying basic charges, the user will decide the amount of transport infrastructure use in consideration of the per unit charges. Traffic will be decided by the same mechanism as marginal cost pricing consequently if the per unit charges are set as marginal costs. That is, people willing to pay more than the marginal cost are buying the transportation service. However, the right to use this transport infrastructure primarily is not purchased if basic charges are high. Thus the amount of the basic charge influences the amount of traffic. It seems unreasonable to set the same basic charges for all users, but we could change basic charges according to the amount of use. The impact of basic charges on deciding whether to subscribe could be examined for each actor (fixed charges elasticity on subscription), so that we could increase the amount of use.

As mentioned above, although the merits and demerits of short-run marginal cost pricing and average cost pricing are understood, is it possible to discriminate the effective use of these two charging principles?

At present we are not good at applying either principle. That is, as mentioned above, the road user already pays various taxes and charges, and there is a high possibility that more than the average cost has been paid. Furthermore the taxes and charges are not related with the cost of congestion, air pollution and so on. Therefore, on average, taxes and charges have become excessive, and have been unnecessarily reducing traffic. At the same time, there is a high possibility that there is excessive traffic along congested roads because externalities such as congestion are not reflected in the present taxation and charging system (EC, 2001).

In reality the proper use of the two charging methods has been proposed. While the European Commission is conducting research on the method of estimating and internalizing external costs in transport (IMPACT: Internalisation Measures and Policies for All external Cost of Transport), as part of that, they propose to introduce average cost pricing to ensure revenue sources independently of congestion. Then defining maximum amount of congestion charge as the value determined by subtracting the average cost of transport infrastructure from the external costs of congestion, they propose to add the congestion charge to the average cost during congested hours (Nishikawa and Kon, 2011). This is recognized as selecting the effective charging principle.

There is a limit to our arguments, however, given the transport infrastructure. Short-run marginal cost charging or short-run average cost charging produce inconvenience associated with the supply-demand gap. If the road capacity is adjusted and harmonized with transport demand in the long run, we could avoid such inconvenience; continuously charging road users with excessive short-run marginal costs at congested sections and continuously imposing excessive short-run average costs at non-congested sections. Henceforth, we would like to see a charging principle that assumes the case where the road capacity can be changed.

V. *Infrastructure Capacity Optimization Using Long-Run Marginal Cost*

A road is composed of various facilities having different life spans such as a traffic information system usable for 10 or 20 years and a bridge usable for 100 or more years. This means that within a certain fixed period the road capacity can be increased or reduced, and can be made to conform to the transport demand at that time. In 2050, the Japanese population is predicted to decline to 70 percent of the current population. Furthermore, it is assumed that areas with increased transport demand can be separated from areas with decreasing demand, so that it is necessary to harmonize, whether to increase or reduce, the road capacity with the transport demand in the long run.

First, we consider long-run marginal cost pricing, which optimizes the road capacity. The long-run average cost curve, as shown in Fig. 1, is an envelope curve connecting the minimum values of the short-run average costs. On the other hand, the long-run marginal cost curve is a line that connects the values of the short-run marginal cost at the time of the optimal production capacity at each production quantity. The short-run average cost curve corresponding to the optimal scale of production at Q_1 in Fig. 4 touches the long-run average cost curve, and the short-run marginal cost C_3 is also the long-run marginal cost at Q_1.

In cases where the capacity of the transport infrastructure increases, it is unclear whether the long-run average cost continues to gradually decrease or whether it increases if a certain road capacity is exceeded (whether or not diseconomy of scale arises). In conclusion, it is sufficient to assume that it continues to decrease gradually as far as the existing and planned infrastructure capacity in Japan is concerned. As for the basis, since the transport infrastructure is exclusively supplied and the transport service cannot be moved to other areas, naturally there is a limit to the transport demand of the area concerned. Therefore, when the traffic volume is less than the minimum capacity of the transport infrastructure (e.g., single-track railway, two-lane highway, etc.), the average cost gradually decreases until the capacity is reached.

Let us consider the case of an appropriate increase in road capacity that is compatible with the increase in demand. In the long-run assumption, it may be satisfactory to think that congestion may be controlled up to the optimal level, i.e. travel time cost does not increase that much even if traffic increases. Among the total cost, the cost of construction and maintenance increases depending on the extent of the rate of increase of the width and the number of lanes,

Fig. 4 Long-Run Average Cost and Long Run Marginal Cost

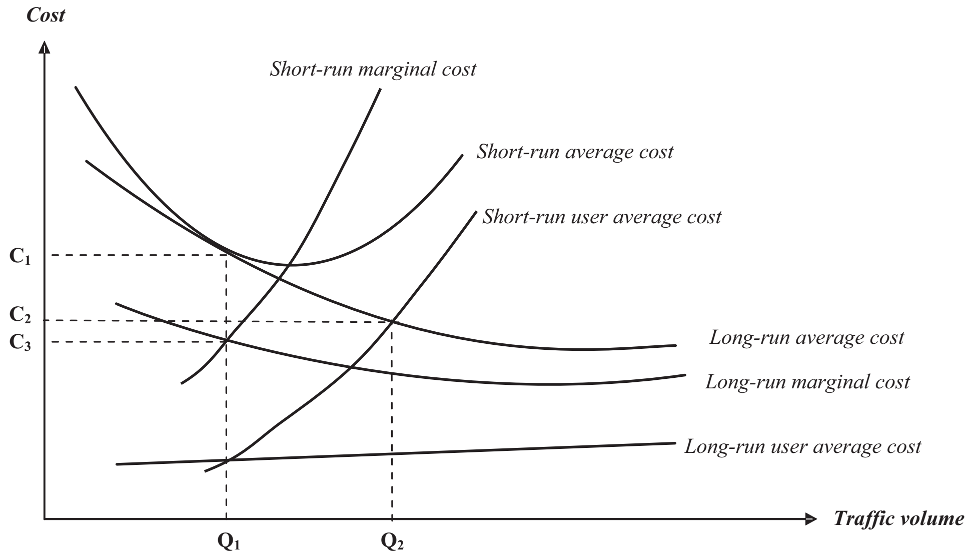
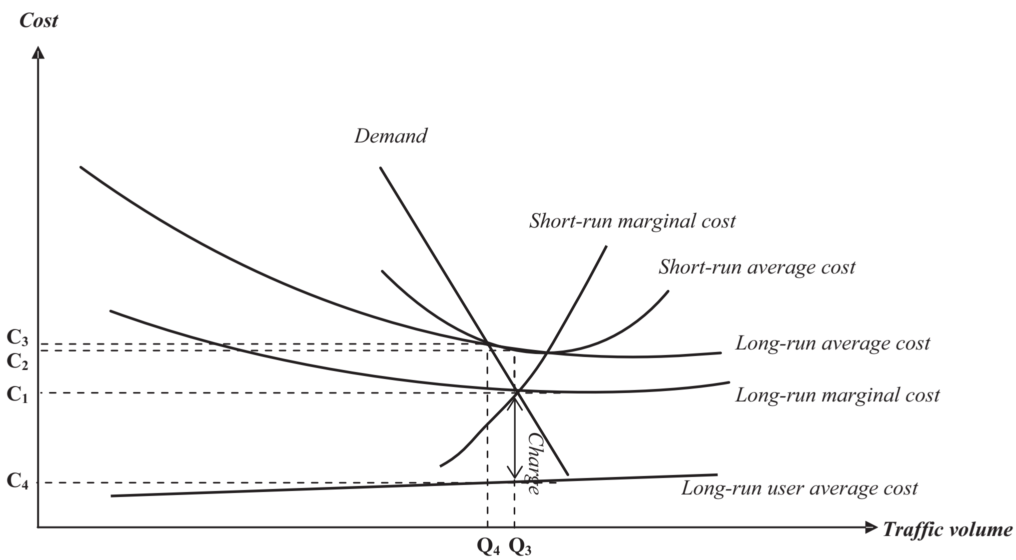


Fig. 5 Long-Run Marginal Cost Pricing



while traffic capacity increases more than the rate of physical increase. This is why the average cost decreases. Based on Fig. 4, the long-run average cost C_2 at traffic Q_2, twice as much traffic as Q_1, is cheaper than the long-run average cost C_1 at traffic Q_1, since we do not need twice as many lanes in order to double the capacity. Quinet (1997) reviewed existing researches and indicated that economies of scale exist up to 4 lane-roads (per direction), although constant returns to scale or diseconomies of scale exist for those with more than 4 lanes.

There are wide highways and large airports in the United States. If the diseconomy of scale is large, the transport infrastructure is divided and construction of large capacity transport infrastructure can be avoided. Japan, at least, does not need to consider diseconomies of scale in the number of lanes of the highways that it provides (at most about four lanes per direction).

Long-run marginal cost pricing, under the assumption that transport infrastructure capacity is variable, specifies a charging principle that determines the optimal traffic volume Q_3 and the charge (C_1 -C_4), which is the point where the marginal benefit (willingness-to-pay for the marginal use of transport service) becomes equal to the long-run marginal cost (Fig. 5). At the same time as deciding the optimal traffic volume, we decide the optimal road capacity to produce the traffic efficiently. The short-run average cost curve and the short-run marginal cost curve at the optimal road capacity Q_3, are as shown in the figure. Therefore, the charge decided by long-run marginal cost pricing becomes the same as the charge decided by short-run marginal cost pricing at the optimal road capacity (Jansson, 1997).

As in the short-run case, since the long-run average cost C_2 at the optimized condition is larger than C_1, a deficit ((C_1 -C_2) *Q_3) is produced, which needs governmental subsidy (McCarthy, 2001). However, unlike the short-run case, cost deviation is not produced because of diseconomies of scale due to congestion and economies of scale due to excess capacity under the imbalance of supply and demand. Moreover, when there is a certain amount of demand and both curves of long-run marginal cost and long-run average cost approaches (relatively small economies of scale condition) as shown in the figure, Q_3 at long-run marginal cost pricing does not deviate greatly from the capacity Q_4 in the case of long-run average cost pricing. Incidentally, during short-run analysis, the charges and the equilibrium traffic greatly differ in marginal cost pricing and average cost pricing.

Long-run marginal cost represents an increase of construction (renewal) and maintenance costs associated with the unit increase of road capacity. Therefore, the integrated value of the long-run marginal cost from the starting point up to Q_3 can be interpreted as the maintenance and renewal costs of transport infrastructure under optimized conditions.

However, it is difficult to measure demand or marginal benefit in the long-run analysis. It may be necessary to compare it with the demand used in short-run analysis. The demand used in Figs. 2 and 3 represents the willingness-to-pay of consumers to a transport service. The willingness-to-pay can be presumed by observing consumer behavior (e.g. change in traffic before and after charge increases). Moreover, since the willingness-to-pay against the transportation service is provided by the same existing transport infrastructure, the willingness-to-pay of consumers can be aggregated and the total is a consumer surplus.

The demand used in long-run analysis becomes the willingness-to-pay amount against the imaginary transport service provided by the future transport infrastructure. Therefore, the long-run marginal benefit must be identified using a different method with the short-run demand. Since observed data on user behavior cannot be used, a carefully designed questionnaire survey is necessary. In particular, when asking about the willingness-to-pay of a transport service assuming new construction of transport infrastructure, the reply may be subjected to certain biases influenced by the likely charging methods and governmental subsidies to the construction project.

VI. *Expansion and Reduction of Infrastructure Capacity by Successive Short-Run Marginal Cost Pricing*

In this section, the most suitable method of charging to optimize the capacity when the road capacity is too large or too little compared to demand is examined on the premise that economies of scale exist in road service production. Figure 5 in the previous section shows the charge required for the continuous maintenance and renewal of road capacities during long-run optimization, which describes the static equilibrium condition. It is more important to analyze the dynamic decision making processes of the introduction of charges by examining whether it is useful as an incentive for approaching the optimal condition or whether there is social acceptance and institutional compatibility (Rothengatter, 2003).

First, let us look at the charging method for attaining the optimal condition when there are constant returns to scale in road service production. As shown in Fig. 6, at constant returns to scale, the long-run average cost becomes horizontal and corresponds with the long-run marginal cost. Moreover, we could assume that the long-run user average cost is also horizontal. The road capacity 'before optimization' shown in the figure is lower than the demand. If the road capacity is expanded to the intersection point where the demand curve crosses the long-run marginal cost curve (i.e. in the optimal capacity), the cost C_1 at that time likewise corresponds with the short-run average cost and the short-run marginal cost. That is, the charge by short-run marginal cost pricing (short-run marginal cost minus short-run user average cost) corresponds with the cost (short-run average cost) for the continuous maintenance and renewal of the infrastructure concerned (Mohring, 1976; Kanemoto, 1997).

A dynamic interpretation has been added to the static analysis of this optimal condition. Verhoef and others suggested that the marginal capacity expansion of specific portions could be justified when congestion tax revenues from those portions can provide increases in capacity (CE Delft, 2002). Furthermore, simulation analysis of road development and maintenance procedures has shown that charging revenues by successive short-run marginal cost pricing promote the gradual expansion of roads, and realize optimal road capacities (Nemoto and Misui, 2008; Nemoto et al., 2009).

Specifically, the charge of short-run marginal cost pricing before optimization is AD as shown in Fig. 6. Of this charge, the value BD, which is the short-run average cost minus the short-run user average cost, becomes the road manager's cost for maintenance and renewal of the existing roads. The remaining amount of charge AB can then be turned into new investments, such as road expansion and network extension. From this new investment, the road capacity will gradually increase in the next period, and the short-run average cost and short-run marginal cost curves will be shifted parallel towards the right. If demand remains the same, the charge will become cheaper because the road capacity will increase. However, as long as the intersection with the short-run marginal cost curve is above the short-run average cost curve, a portion of the charge can then be diverted into new investments, and eventually converges to the road capacity 'after optimization' as shown in the figure.

On the contrary, on uncongested roads where demand is less than the road capacity, the charge becomes small. Therefore, renewal is terminated to the extent that the marginal cost falls below the short-run average cost, resulting in gradual reduction of the road capacity. However, as the gap between the road capacity and traffic demand becomes narrow, the charge increases

FIG. 6 ROAD CAPACITY INCREASE BY SUCCESSIVE SHORT-RUN MARGINAL COST PRICING
(constant returns to scale)

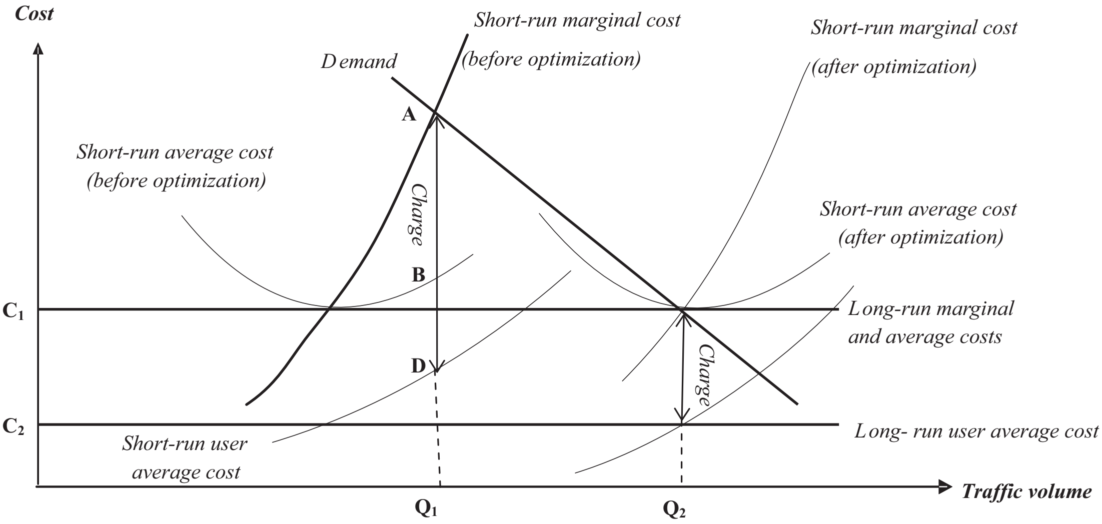
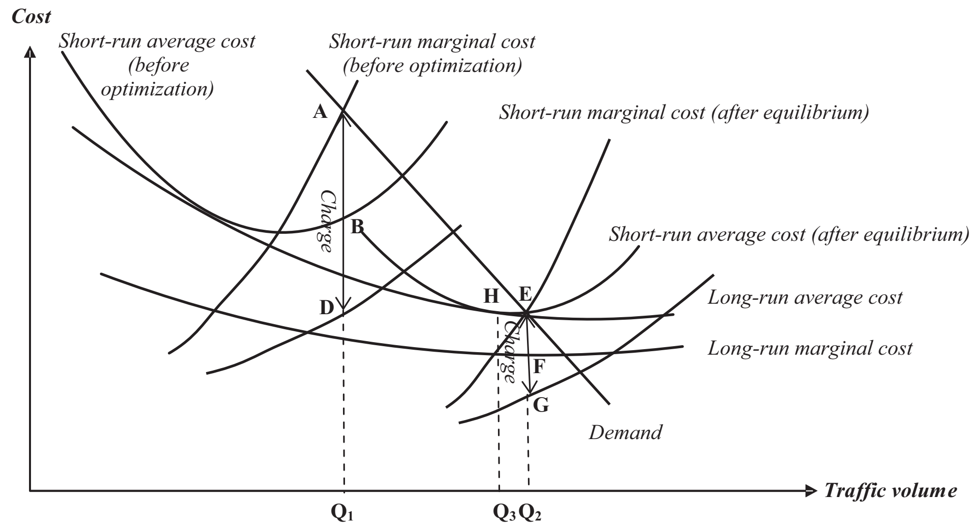


FIG. 7 ROAD CAPACITY INCREASE BY SUCCESSIVE SHORT-RUN MARGINAL COST PRICING
(economies of scale)



and reaches an amount sufficient to cover all maintenance and renewal costs.

In both cases of over- and under-capacity against traffic demand, we could adjust the capacity to the traffic demand in the long run by successive short-run marginal cost pricing.

This paper examines the user-pay principle of arterial roads whose capacity can be expanded or reduced according to the traffic demand. This charging principle should not be applied to roads with small traffic in rural areas, as this would penalize rural mobility. We should form a consensus on the provision of a 'universal service' of local roads, and introduce a different principle led by the government.

Next, let us explore the case where the production of road service has economies of scale (Fig. 7). As is the case for constant returns to scale, the charge before optimization is AD, and of this, AB can be diverted into new investments. Although the road capacity gradually increases due to new investments, it will be balanced before reaching the optimal capacity, which is different to the process shown in Fig. 6. Specifically, the capacity is determined corresponding to the intersection of the demand curve, the short-run average cost curve and the short-run marginal cost curve (E). Please note that the short-run marginal cost intersects with the short-run average cost at the bottom of the curve. The charge (EG) here corresponds to the cost of maintenance and renewal of this road capacity. The short-run average cost slightly exceeds the long-run average cost. Although the revenue and necessary expenditure is balanced, it is not optimized in the long run as shown in Fig. 5. The road capacity at the equilibrium state is Q_3, and the short-run average cost curve touches the long-run average cost curve at H. Thus, when economies of scale exist, the road capacity will be balanced with a smaller than the optimal road capacity. This equilibrium is interpreted as a situation in which the road capacity is kept smaller than the optimal capacity resulting in slight congestion so that the road manager can balance the road expenditure with the charge revenue, and thus avoid a deficit as shown in Fig. 5.

Road expansion with borrowed money seems an attractive financing alternative (Small and Verhoef, 2007), although this paper examines successive charging and continuous expansion. Time would be saved if the roads were constructed and congestion reduced sooner, so it may be optimal to expand the roads with debt in accordance with the predicted future traffic demand and repay the debt with the user charges. However, the charge by short-run marginal cost pricing after optimization is not sufficient to repay the debt after the congestion disappears. Therefore, we may increase the road capacity to below optimal levels in order to make the charge higher, or we may add the debt interest to the charge determined by short-run marginal cost pricing.

VI. *Conclusion*

This paper proposes a road planning and financing system in which the road capacity is expanded or reduced to meet traffic demand with the revenue of charges by short-run marginal cost pricing. When the capacity is less than the traffic demand, we could expand the capacity with charges on congestion. When the capacity is more than the traffic demand, we have to reduce the capacity in order to balance the renewal and maintenance costs and the revenue.

The charging principle aiming at transportation demand management is important, and should be utilized depending on the situation of the area. Road capacity cannot be increased or decreased in the short term. In order to utilize the present road capacity to the maximum, it is possible that a high-priced charge based on short-run marginal cost at congested sections would shift transport demand to other time periods and other routes. And a low-priced charge could induce transport at non-congested sections. However, transportation demand management is a strictly short-run measure. As we can increase or reduce the road capacity in the long run, the need for transportation demand management may diminish. The planning process of increasing and decreasing transport infrastructure capacity gradually by the short-run marginal cost charging shown in this paper should be more effective.

In addition, the planning process seems effective if also seen from the incentives of stakeholders concerned. That is, when the road is congested, road users who have paid high charges will demand expansion of road capacities. The road manager can also implement measures with the revenue from congestion charges. On the other hand, road users who have enjoyed low charges on non-congested roads must agree to reduce the road capacity when the facilities approach the end of their life span and need to be updated.

In an age of decreasing population, it is necessary to analyze the economic aspects of road supply and demand. In areas with decreasing population, economic activities are likely to diminish, leading to reduced transport demand. The optimal road capacity should not be decided independently but decided simultaneously with the acceptable sharing of burdens for maintenance and renewal.

References

CE Delft (2002) *Return on Roads*.

European Commission (1998) *Fair Payment for Infrastructure Use: A phased approach to a common transport infrastructure charging framework in the EU*, *White Paper;* COM, 466.

European Commission (2001), *White Paper; European Transport Policy for 2010*, COM (2001) 370.

European Commission (2011) *White Paper; Roadmap to a Single European Transport Area*, COM (2011), 144.

Kanemoto, Yoshitugu (1997) *Urban Economics*, Toyo-Kezai-Shinposha (in Japanese)

Ministry of Land, Infrastructure, Transport and Tourism (2012) *Plan for the Improvement of Social Infrastructure* (in Japanese).

McCarthy, P. S., (2001), *Transportation Economics*, Blackwell Publisher Ltd.

Mohring, H. (1976) *Transportation Economics*, Cambridge Mels.

Nash, C. (2008) Transport Infrastructure Pricing: A European Perspective, *Pricing in Road Transport*, edited by Verhoef et al., Edward Elgar.

Nemoto, T. and Misui, Y. (2008) *Distance-Based Taxation for Road Maintenance*, Keiso-Shobo (in Japanese).

Nemoto, T., Misui, Y., and Kajiwara, A. (2009) Optimal Road capacity Building: Road Planning by Marginal Cost Pricing, *Journal of Infrastructure Systems*, ASCE, pp.290-296.

Nishikawa, R. and Kon, N. (2011) Trend of Eurovignette Directive on Road Use Charges for Heavy Goods Vehicles and its Implication for Japan, *Institute for Transport Policy Studies*, Vol. 14, No. 1.

Ozawa, S. and Nemoto, T. (2013) Track Access Charge Roles in Railway Vertical Separation in the European Union, *The Japanese Journal of Transportation Economics, No 56* (in Japanese).

Quinet, E. (1997), Full Social Cost of Transportation in Europe, *The Full Costs and Benefits of*

Transportation, edited by D. L. Greene et al., Springer.
Rothengatter, W. (2003), How good first best? Marginal cost and other pricing principles for user charging transport, *Transport Policy*, 10.
Small, K. A. and Verhoef, E. T. (2007), *The Economics of Urban Transportation*, Routledge.