The Charging Principle for the Development and Maintenance of Transport Infrastructure

Toshinori NEMOTO

Working Paper Series No. 141
Graduate School of Commerce and Management
Hitotsubashi University

October 2012

The Charging Principle for
the Development and Maintenance of Transport Infrastructure

Toshinori Nemoto (Hitotsubashi University)

Abstract:

With the onset of an 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 life in rural areas and provide infrastructure that strengthens global competitiveness of large cities has become more vital than before. Moreover, costs concerning maintenance and renewal of transport infrastructure are also expected to be needed (MLIT, 2012). Under such a situation, what kind of charging principle should be introduced in order to ensure revenue sources for maintenance and renewal of transport infrastructure, such as roads and railways? The objectives of this paper are: 1) to clarify definitions of short run marginal cost pricing and short run average cost pricing (charging principles 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 short run marginal cost pricing through a planning process that gradually increases or reduces transport infrastructure capacity.

1. Transport System Costs

It is necessary to define transport system costs to design charging 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 costs which are not related to traffic volume, such as construction costs. Variable costs are costs determined according to the amount of traffic, such as maintenance costs. The value which divides traffic volume from the aggregate total cost (the sum of fixed costs and variable costs) is the average cost; and the value which differentiates variable cost from traffic volume is the marginal cost.

It is also important to pay attention to stakeholders that generate costs. Broadly speaking, the stakeholders consist of that of the transport infrastructure manager doing construction and maintenance, the user using the transport infrastructure, and other stakeholder which can be subjected to positive and/or negative impacts arising from the use of the transport infrastructure although not directly related to the provision and use of transport infrastructure. Transport infrastructure managers and users are assumed to form a pseudo market, and their costs are called internal costs, while costs arising on other stakeholder are called external costs.

Although users can be divided into transport companies (e.g. bus companies) who provide transport services and directly use the transport infrastructure, and service users who purchase these transport services, this paper focuses on the former type of users who directly use the transport infrastructure, and looks at how taxes and charges are collected for maintenance and renewal of transport infrastructure. Since transport companies do not exist for "road use by private cars", the transport system is simpler and it is easier to appreciate the distribution of benefits and costs. 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 taxes and charges which 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 are cancelled resulting in no social costs. The ultimate costs remaining are A, B, C, D, G, and H. This sum functions as an aggregate total cost.

The charges and taxes which the users pay are not all appropriated for the maintenance of the transport infrastructure. Historically, there exist an earmarked financing system which uses taxes and charges from users of specific transport infrastructure for its maintenance. However, the number of countries using this system has been reduced at the moment. Some examples that have exceptionally remained are the federal gasoline tax of the U.S., the energy tax of Germany, and the axle tax of France. While new road construction becomes difficult partly owing to people’s environmental concerns, the Ministry of Finance, which wants to secure the favourably increasing automobile-related taxes as a source of revenue, abolished the earmarked financing system, and has been using them as a general revenue source. In the newest European transport white paper (EC, 2011), it has been specified that “it is important to establish a financial system in which revenue from transport users is used for the improvement of transport". This is in agreement with the position of this paper which aims at coinciding both benefit and cost.

Furthermore, it is necessary to distinguish between short run and long run costs. Short run costs are costs 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.

Table 1. Road system costs

		Fixed costs	Variable costs (related to traffic volume)
Internal Cost	Road Manager	A: Construction cost	B: Maintenance cost
	Road User (i.e., E and F are taxes/charges transferred to the road manager)	C: Vehicle cost	D: Time cost, Fuel cost
		E: Vehicle ownership charges	F ： Gasoline tax, toll charges, distance-based charges
External Cost	Other Stakeholder	G: Improvement or destruction of scenery	H: Congestion, Air pollution

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 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. convex form).

The marginal cost is the increase in total social cost caused by unit increases in traffic. Since the increase in time cost of the society caused by the additional users at 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 will be 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, it can be seen that the long run average cost decreases.

For transport infrastructure, it has been pointed out that ‘indivisibilities’, which means that the number of lanes of a road must be an integar, 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 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 maximum roadway capacity can be secured and designed under a certain width of the street facility.

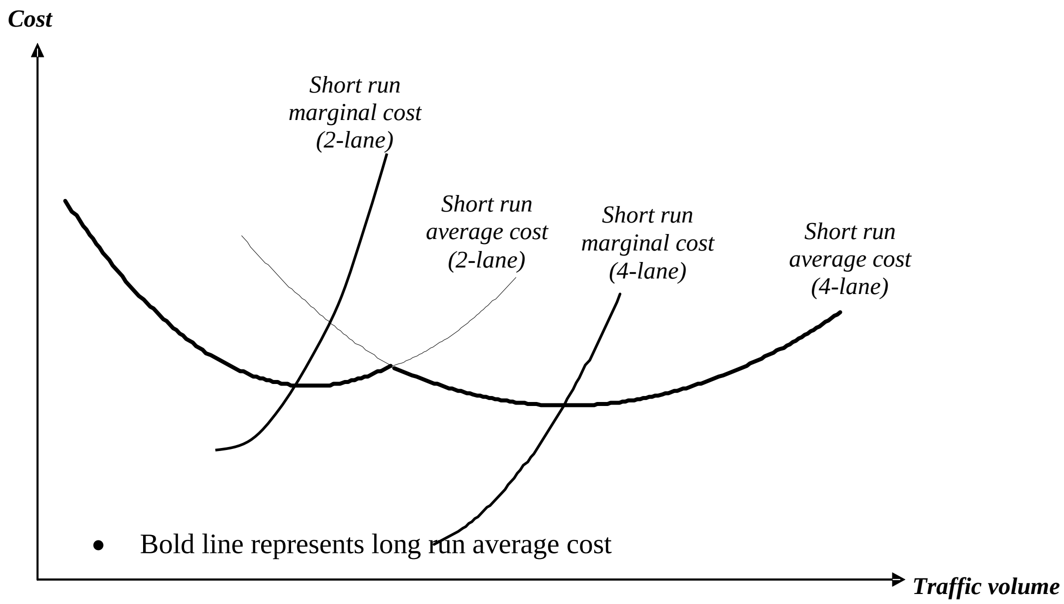

Figure 1. Short run average cost, Short run marginal cost, and Long run average cost

2. 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, it is understood that price setting by short run marginal cost (or for accuracy, short run marginal social cost containing external cost), can realize optimal usage of a transport infrastructure. That is, it is believed that traffic volume that maximizes social benefits can be realized using road capacity at that time.

Marginal cost pricing is explained using Fig. 2. Let us recall the two-lane short run average cost and the short run marginal cost curves which were shown in Fig. 1. There, a short run user average cost curve is added. The user average cost is the value equal to the user cost (C and D of Table 1) divided by 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 (a 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 cost that society additionally pays. To realize this amount of traffic, a deviated portion of the short run marginal cost and the short run user average cost is charged, so that the users would face 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 be all 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 infrastructure, 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 compensate with other taxes the fund shortage arising from marginal cost pricing. Although supplementation by poll tax that does not distort resource allocation is recommended, its introduction is generally difficult from the viewpoint of fairness. Under a situation where priority must be given to expenditures on social security and pension, expenditures on transport infrastructure from general revenue cannot be expected.

In addition, the European White Paper on Fair Payment for Infrastructure Use (EC, 1998) indicated "promotion of fair competition" and "revenue generation by internalizing external costs" as two reasons to support marginal cost pricing. The former is a proposal made for the adoption of marginal cost charging as a standard for EU in order to prevent international transport service providers to receive disadvantageous treatment on competition. At present each country imposes different taxes and charges for each transport mode, which influences the international transport service providers favourably or unfavourably. However, the peripheral states of EU such as Eastern Europe oppose the standardization of charges on congestion, one of the typical externalities. They regard the congestion charges as unfair because the trucks belonging to these countries have to pay the charges when passing through the severely congested central part of Europe (Nash, 2008).

The latter is a proposal which indicates that external costs of congestion should be charged such that infrastructure costs of the whole transport system can be provided. In the EU, since border customs clearance was abolished in 1993 and international truck traffic (ton-km) has increased more than the GDP growth, it is expected that external costs of congestion and environment are no doubt larger than the cost provision of transport infrastructure (including railways). However, criticisms have been raised such as those of Rothengatter (2003) who stated that "assistance on railway 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 other mode (or area) does not guarantee that the required amount can be secured, and the improved efficiency due to marginal cost pricing during the charging period also entails risks of being ruined in the expenditure period. Incidentally, the term called marginal cost pricing had not been used in the subsequent transport White Papers (EC, 2001; EC, 2011).

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

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

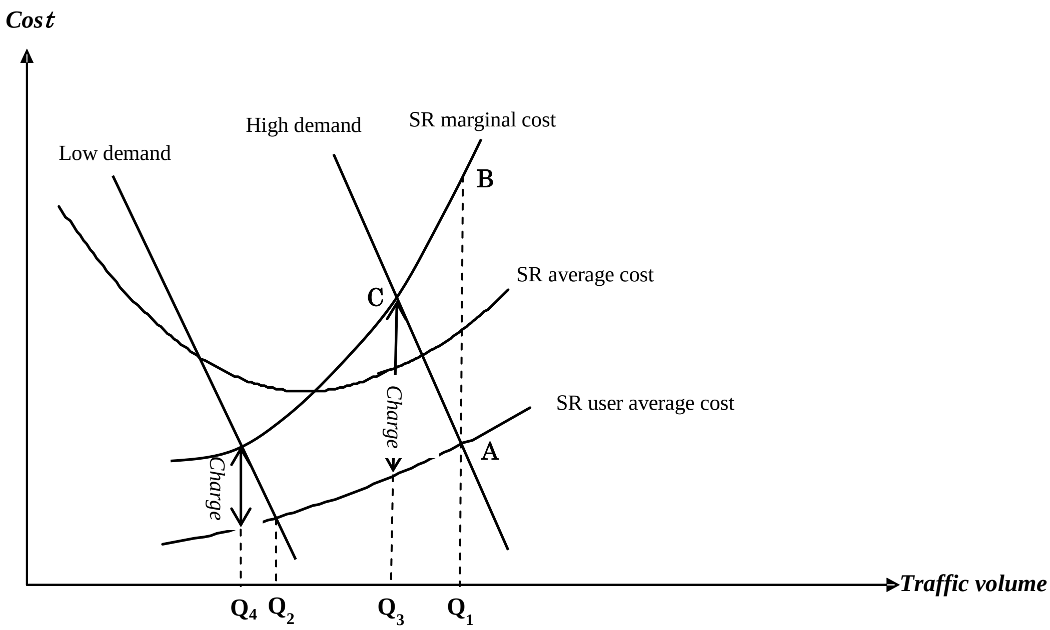

Figure 2. Short run marginal cost pricing

3. 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 Q5, and the traffic could not be reduced to the optimal volume. Conversely, at low demand, since the charge becomes higher, traffic volume smaller than the optimal one can be realized (Q6). In other words, in the marginal cost pricing at low demand as previously explained, it can be considered that 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 expenses arising from other stakeholder are collected, it is assumed for simplification that the collected portion of costs arising from other stakeholder is outlaid for the costs of relieving externalities (e.g. spending revenue corresponding to noise externalities in construction of noise barrier).

The problem of average cost pricing is that the charge becomes high in areas with little demand. Transport demand may decrease against the traffic infrastructure manager's expectation due to factors such as population decrease and factory relocation, among others. If there are few users and the charge is expensive, traffic will furthermore decrease. Supposing that excessive transport infrastructure exists in the area where transport demand decreased, the following discussion will introduce essential solutions to reduce transport infrastructure capacity. Solutions using charges are limited on the premise of maintaining continuously the present transport infrastructure.

Maintaining the idea of average cost pricing to recover total costs, two-part tariff and Ramsey pricing which do not reduce traffic greatly in a low demand area are attractive charging methods which deserve introduction. Even though the social benefits do not necessarily increase dramatically, their implementation should be considered if it will be clarified that little increases in social benefits are attainable through these charging methods.

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 with maintained percent composition of users, by making light the burden of users who are sensitive to charging (e.g. passenger car is more sensitive if comparing between passenger car and truck), and making heavy the burden of users who are not sensitive to charging. However, the charge for those who have no alternative mode of transport and then are not sensitive becomes high. From the viewpoint of fairness, therefore, almost no actual application exists.

It is understood that two-part tariff does not need to reduce much traffic either. Two-part tariff refers to a method of collecting fixed costs as basic charges and collecting variable costs as per unit charges. 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 are not purchased if basic charges are priced expensively. The basic charges influence the amount of traffic. It seems unreasonable that the same basic charges are set to all users. It is possible to change basic charges according to the amount of use. The impact of basic charges on making decision whether or not to contract could be examined for each actor (fixed charges elasticity on subscription), and the method of increasing the amount of use is proposed.

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

At present we are not good at using either principle. That is, as mentioned above, the road user has already paid various taxes and charges, and there is high possibility that more than the average cost has been paid. But the taxes and charges are not related with the cost of congestion and air pollution. Therefore, on the average, taxes and charges have become excessive, and have been unnecessarily reducing traffic. On the other hand, there is high possibility that excessive traffic has been happening because even though externalities such as congestion and air pollution exist, they are not reflected in the present taxation and charging system (EC, 2001).

The proper use of the two charging methods has been proposed. While the European Commission is doing 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, the maximum amount of congestion charging is explained as the value determined by subtracting the average cost of transport infrastructure from the external costs of congestion (Nishikawa and Kon, 2011). This is recognized as selecting the effective charging principle, which acknowledges average cost pricing to ensure revenue sources during non-congested hours, and recommending marginal cost pricing to realize optimal traffic during congested hours.

There is a limit in our arguments so far, 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 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 a case where road capacity can be changed.

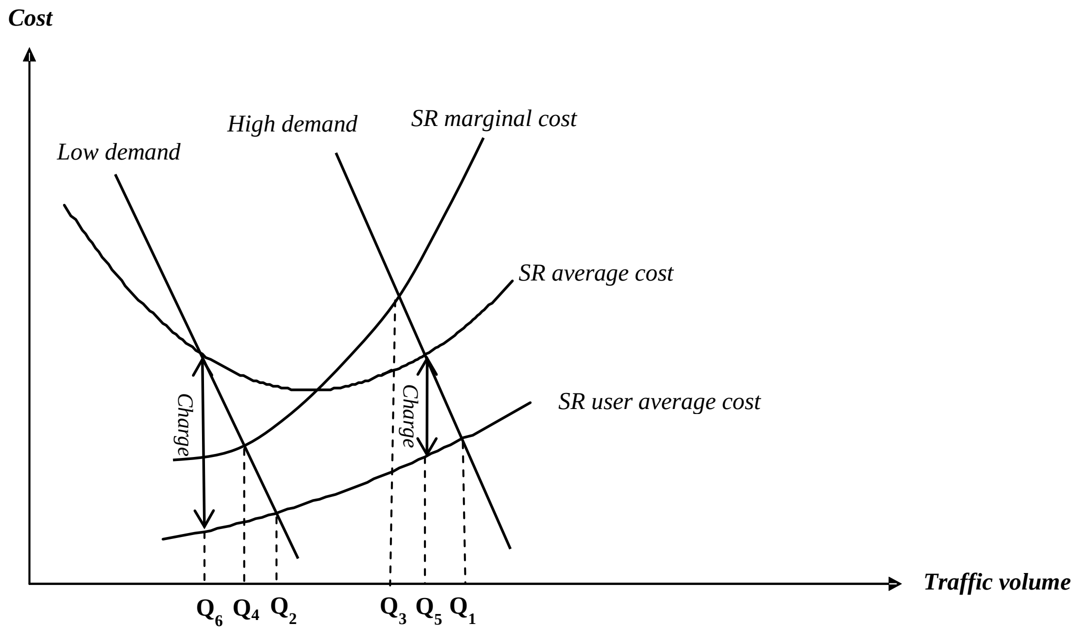

Figure 3 Short run average cost pricing

4. Infrastructure Capacity Optimization using Long Run Marginal Cost

A road is composed of various facilities having different life spans. Included in it for examples are a traffic information system usable for 10 or 20 years and a bridge usable 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 forecasted to become 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, road capacity with the transport demand in the long run.

Let us discuss long run marginal cost pricing that optimizes 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 when the capacity of the transport infrastructure increases, it has been argued whether the long run average cost continues to gradually decrease or whether it increases if certain road capacity is exceeded (whether or not diseconomy of scale arises). As a conclusion, it is sufficient to think that it continues to gradually decrease as long as existing and planned infrastructure capacity in Japan is concerned. As for the basis, since the transport infrastructure are exclusively supplied and the transport service cannot be moved to other areas, naturally there is a limit in 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.

Let us tackle the case of the 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 optimal level, i.e. 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, 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 processing traffic Q_2, twice as much traffic as Q_1, is cheaper than the long run average cost C_1 processing traffic Q_1, since we need not 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.

Highways having large widths and large airports exist 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 worry about diseconomies of scale in the number of lanes of a highway 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 imbalance of demand and supply. 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 differs in marginal cost pricing and average cost pricing.

Long run marginal cost represents 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 at optimized condition. Moreover, this becomes equivalent to $(C_3 * Q_3)$.

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 are willingness-to-pay of consumers to a transport service. The willingness-to-pay can be presumed by observing consumer behaviour (e.g. change in traffic before and after charge increases). Moreover, since the willingness-to-pay against the same transportation service is provided by the same transport infrastructure, willingness-to-pay of consumers can be aggregated and the total is a consumer benefit.

The demand used in long run analysis becomes the willingness-to-pay amount against the transport service provided by the transport infrastructure that does not exist. Therefore, the long run marginal benefit must be identified using a different method with the short run demand. Since observed data on user behaviour cannot be used, a carefully designed questionnaire survey is necessary. In particular, when asking 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 subsidy to the construction project.

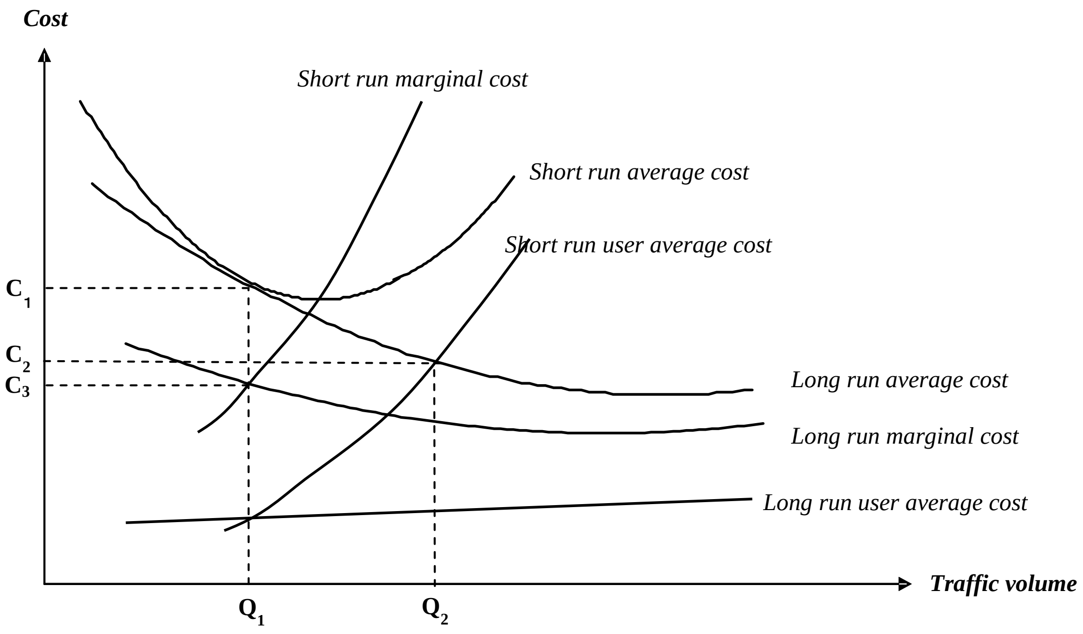

Figure 4. Long run average cost and long run marginal cost

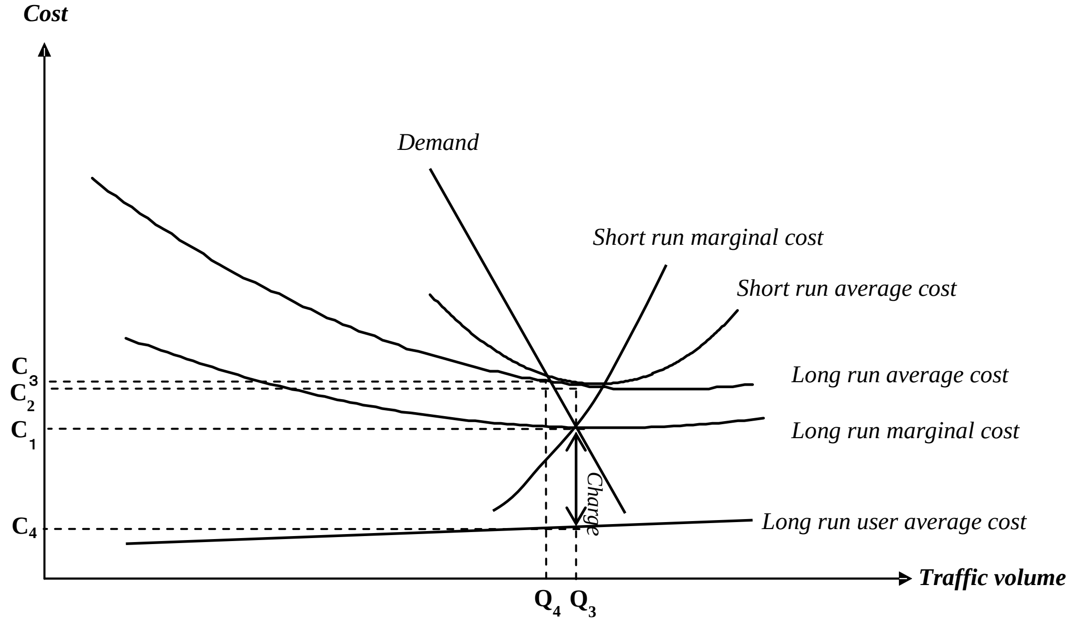

Figure 5. Long run marginal cost pricing

5. Expansion and Reduction of Infrastructure Capacity by Successive Short Run Marginal Cost Pricing

In this section, the most suitable method of charging is examined that will optimize the capacity when road capacity is too high or too little compared to demand, on the premise that economies of scale exist in road service production. Fig. 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 seem more important to analyze dynamic decision making processes of the introduction of charges by examining whether it is useful as incentives for getting closer to the optimal condition or whether there is social acceptance and institutional compatibility (Rothengatter, 2003).

First, let us look at the charging method to attain 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 in a condition 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, 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 concerned infrastructure (Mohring, 1976; Kanemoto, 1997).

A dynamic interpretation has been added to the static analysis of this optimal condition. Verhoef and others suggested that marginal capacity expansion of specific portions can be justified when congestion tax revenues from there can provide increases in capacity (CE Delft, 2002). Furthermore, from the simulation analysis of road development and maintenance procedures, it was 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, 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 is kept the same, the charge will become cheaper because 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 turned into new investment, and eventually converges to the road capacity 'after optimization' as shown in the Figure.

On the contrary, at uncongested roads where demand is less than 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 road capacity. However, as the gap between the road capacity and traffic demand becomes narrow, the charge increases and reaches to the amount 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 long run by the 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. We do not discuss the charging principle for roads with small traffic to guarantee people's mobility in rural areas. It seems difficult to apply the same charging principle to them. We have to make consensus on 'universal service' of arterial roads, and introduce a different principle led by the government.

Next, let us explore the case when 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 it, AB can be turned into a new investment. Although road capacity gradually increases due to new investments, it will be balanced before reaching the optimal capacity which is different in the process from 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 marginal cost intersects with the 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 a little bit 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 Q3, 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 the one smaller than the optimal road capacity. This equilibrium is interpreted as a situation in which road capacity is kept smaller than the optimal capacity resulting in a little bit congestion so that road manager can balance the road expenditure with the charge revenue, otherwise might suffer from deficit shown in Fig. 5.

Road expansion with borrowed money seems attractive financing alternative, although this paper examines successive charging and continuous expansion. We could enjoy time saving benefit if the roads are constructed and congestion is reduced earlier. It would be a good exercise when and how much we should expand the roads with debt in accordance to the future traffic demand, and how we should repay the debt with the user charges. It is noted, however, the charge by the 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 smaller than the optimal one in order to make the charge higher, or we may add interest of debt to the charge determined by short run marginal cost pricing.

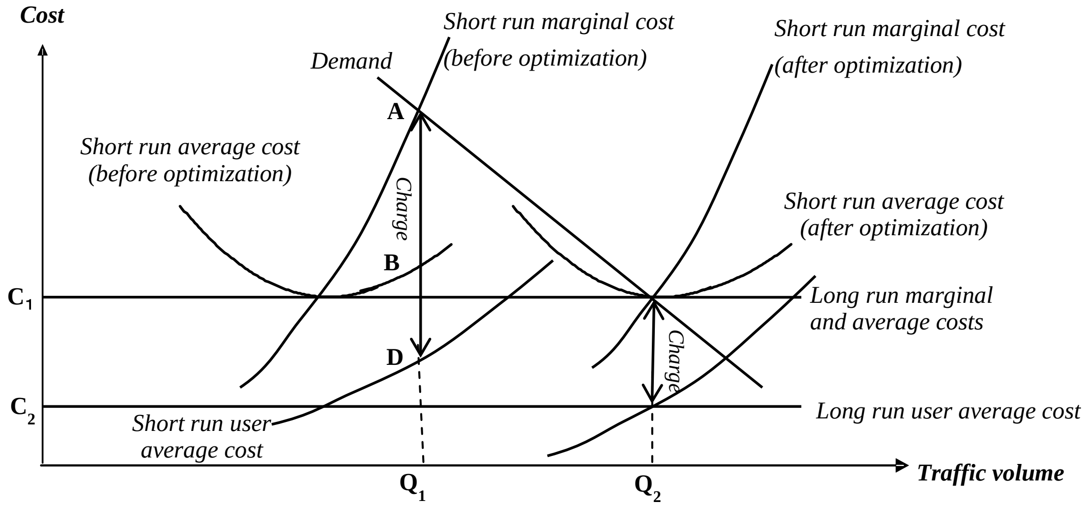

Figure 6. Road capacity increase by successive short run marginal cost pricing (constant returns to scale)

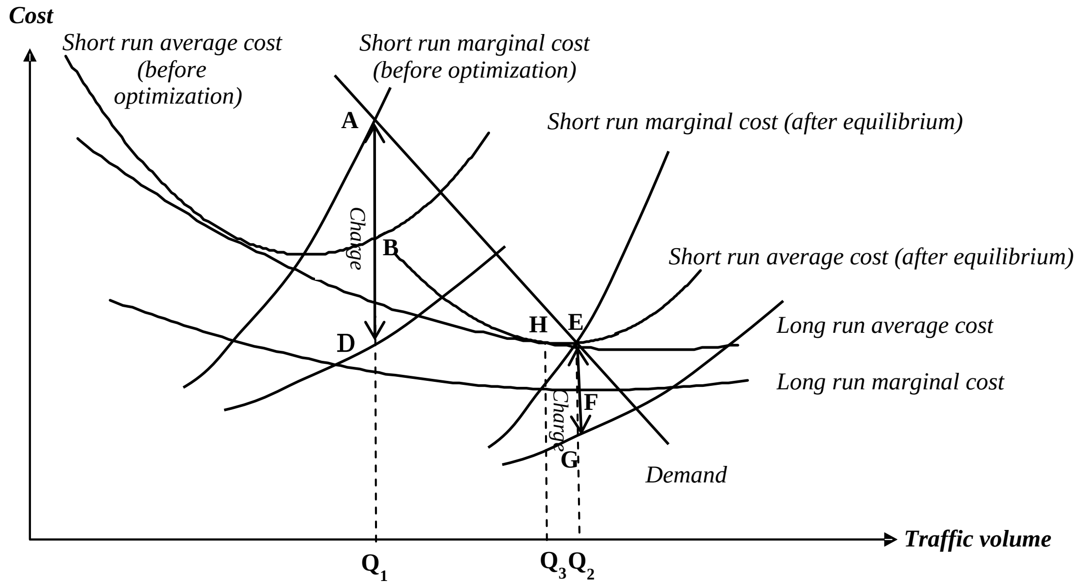

Figure 7. Road capacity increase by successive short run marginal cost pricing (economies of scale)

6. Conclusion

This paper proposes road planning and financing system in which 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 the charges on congestion. When the capacity is more than the traffic demand, we have to reduce the capacity in order to balance renewal costs and the revenue.

We regard important the charging principle aiming at transportation demand management. It should be utilized positively 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 makes transport demand shift to other time periods and other routes. And a low-priced charge could induce transport at non-congested sections. However, transportation demand management is strictly a short run measure. As we can increase or reduce road capacity in the long run, it is normal that the necessity of these transportation demand management may become lost. It seems that the planning process of increasing and decreasing transport infrastructure capacity gradually by short run marginal cost charging shown in this paper is more effective in that sense.

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 revenue of congestion charges. On the other hand, road users who have enjoyed cheap charges on non-congested roads must get to agree to reduce road capacity when the facilities get to reach their end-of-life and need to be updated.

In the age of decreasing population, it is necessary to analyze the economic aspect of road demand and supply. In areas with decreasing population, there is high possibility that economic activities are reduced at increased rates and thus also reduced transport demand. As for the optimal road capacity in each area, it is desired that it is not decided independently but decided simultaneously with the benefits and burdens, particularly with the burden on costs 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) The Role of Track Usage Fees in the Vertical Separation of Railways in Europe, *Proceedings of the 71st Annual Conference of the Japan Society of Transportation Economics* (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