Vol. 25 No. 2 (1988. 9)

OPTIMAL CONGESTION TAX OF EXPRESSWAY: A. A. WALTERS RE-EXAMINED, P. K. ELSE RE-APPRAISED, AND DEMAND-SURFACE PARADIGM RE-CONSIDERED

Tatsuhiko Kawashima

CONTENTS

ABSTRACT

This paper will take a sceptical view of the conventional wisdom of the notion: "The theory of marginal cost pricing suggests that the desirable level of traffic flow should always be kept below the highway's capacity for traffic flow." This notion, originating with A. A. Walters' seminal 1961 article, is based on the interpretation that the marginal social cost is infinite when traffic flow reaches the highway capacity. However, this interpretation is disputable, as earlier pointed out by P. K. Else.

Challenging the conventional notion, the paper will try to demonstrate that the marginal social cost actually has a positive constant value when traffic flow is at the level of highway capacity. This implies the existence of a possible case that the optimal traffic flow level can coincide with the highway capacity.

The paper will also try to apply the marginal social cost pricing principle to the paradigm

* Economics Department at Gakushuin University in Tokyo. The author is indebted to the participants of the Tenth Pacific Regional Science Conference, Pusan, Republic of Korea, July 6–11, 1987, as well as to those of the Thirty-fourth North American Meetings of the Regional Science Association, Baltimore, Maryland, U.S.A., November 6–8, 1987, for the helpful comments on earlier drafts of this paper.

of demand-surface analysis through which we can theoretically investigate how the marginal social benefit for highway users is effected by the quality level of highway services which would vary depending on the level of traffic demand. The final section of this paper attempts to synthesize (1) the approach developed for reformulating Walters' notion and (2) the demand-surface approach to get a tiny but new insight into the optimal congestion tax problems.

KEY WORDS

Congestion Tax, Demand-flow Conversion, Demand-surface, Else (P. K.), Optimal Traffic Flow, and Walters (A. A.)

1. Introduction

Highway traffic congestion in urbanized areas has long been recognized as a technological external agglomeration diseconomy which generates a clear discrepancy between the average social cost and the marginal social cost[1]. Under a free market system with neutral taxation, this kind of discrepancy usually brings about inefficient allocation of a fixed highway capacity. For this type of market failure, it has been shown that a Parato improvement in the efficiency of utilizing a highway can be achieved if on all motorists[2] are levied congestion tolls[3] equal to the difference between the average social cost and the margianl social cost corresponding to the demand level which satisfies the condition that the marginal social benefit[4] is equal to the marginal social cost.

The essence of the theory of marginal cost pricing, which serves as an important basis for this argument, was originally conceptualized by Pigou (1920). His marginal principle framework has been theoretically elaborated and empirically examined by a number of transportation economists and transportation engineers[5] for the account of highway congestion. Among those studies that have applied the marginal principle is one by Walters (1961) in which he attempted to apply the principle of marginal cost fare-setting to the services of congested highways in the U.S. For this purpose, he set up a special analytical apparatus in which the average social cost is expressed as a function of the volume of traffic flow. Diagrammatically speaking, this setting makes the average social cost curve have its upper part backward-sloping. Partly due to the curious characteristics of the average social cost curve with a backward sloping part on it and partly due to the theoretical importance of and interest in the marginal cost pricing principle, Walters' work has widely stimulated highway researchers into various types of studies dealing with the optimal congestion tolls as a function of traffic flow. In this sense, Walters has certainly made a fair contribution to the development of analytical scope for the studies on the phenomena of highway congestion.

Walters' work, however, suffers from two deficiencies. Firstly, his work shows that the value of marginal social cost which is expressed as a function of traffic flow becomes infinite if traffic flow reaches the highway capacity[6]. But this shape of marginal social cost function seems inadequate if we employ it for the purpose of finding optimal congestion tolls. Secondly, the grounds for his criticism against Meyer, Peck, Stenason, and Zwick (1959), who proposed that fares on urban roads in the U.S. should be reduced, is rather insufficient. This is because Walters' criticism is not a little based on the conception accruing from his misleading marginal social cost function. It is also because the long-term problem of how best to allocate limited transportation resources which Meyer and his collaborators tackled, is no less important than the short-term efficiency problem of allocating a fixed highway capacity which Watlers tried to investigate in his work.

This paper, in Section 2, illuminates the first deficiency found in the work of Walters, and theoretically rectifies it without losing his original idea of investigating the optimal level of traffic flow in terms of the marginal social cost which is expressed as a function of traffic flow. In this section, the demand-flow conversion function plays an important role. Section 3 sheds diagrammatic light on the same issue as examined in Section 2, while a numerical example is given to clarify the basic concepts canvassed in Sections 2 and 3. Section 5 discusses the paradigm of demand-surface analysis. This approach explicitly draws into the analysis the quality level of road services for a given highway which would vary depending on the level of traffic demand. In the concluding Section is made an attempt to synthesize the demand-flow conversion approach and the demand-surface approach.

2. Marginal Social Cost as a Function of Traffic Flow: Logic

Consider a mile-long, one-way expressway with a fixed capacity, located in an urbanized area. This expressway has neither entrance nor exit but at its two ends. For this setting, let us assume the followings:

Assumption 1

Traffic is homogeneous, which implies that, with a given volume of traffic flow, each vehicle driving along the expressway experiences exactly the same speed and cost.

Assumption 2

Traffic density is the same all the way along the expressway at any point of time, which implies that the dynamic adjustments of traffic density are made instantaneously.

Under these assumptions, from the relationship between density and speed can be derived the curve XX' in Figure 1 as shown by Walters (1961, p. 679), presenting the relationship between traffic flow and time of trip-mile. We can interpret this curve as the average social cost curve[7] if we employ a unit time to measure "the time of trip-mile" as cost-indicator. From this average social cost curve XX' together with the following formula, Walters derived the marginal social cost curve shown by the broken line *MSC* in Figure 1:

$$\text{marginal social cost} = \text{average social cost} \times (1 + \text{elasticity of average social cost}). \quad \cdots\cdots\cdots\cdots(1)$$

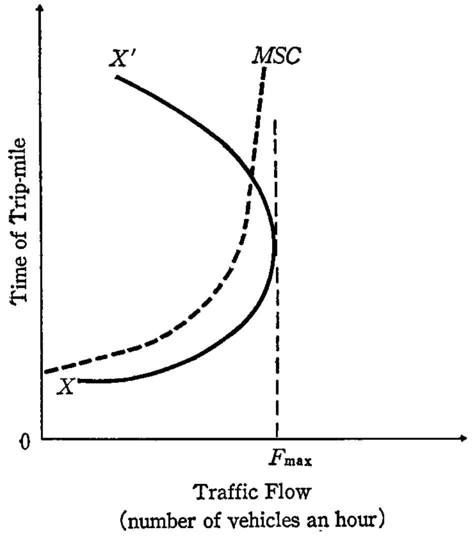
About the broken line curve *MSC*, Walters stated:

> "It (*i.e.*, the curve *MSC*) rises above the unit cost curve (*i.e.*, the average social cost curve *XX'*)......; but when flow reaches $F_{\max}$ the marginal social cost is *infinite*. On the backward sloping part of the unit cost curve the marginal social cost is *not defined* since the change in output is negative. ......The theory of marginal cost pricing suggests that taxes be levied to reduce demand until traffic flow is at a level where private unit cost (with tax) is equal to marginal social cost. Thus *traffic flow should always be kept below capacity*[8]."

This interpretation seems inadequate. My opinion goes:

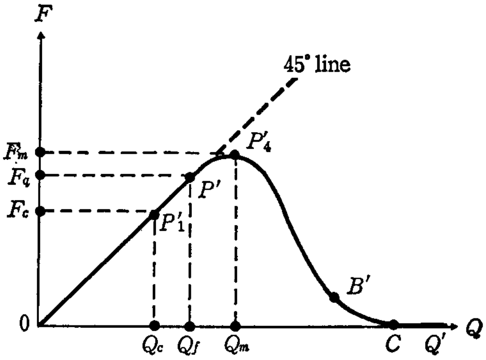
> When flow is at the level of $F_{\max}$, the marginal social cost has a *positive constant* value. On the backward sloping part of the average social cost curve, the marginal social cost is *defined*. Accordingly, it is possible for the *optimal traffic flow*[9] *to coincide with capacity.*

In order to furnish the grounds for this opinion of mine, observe that Equation



Source: Walters (1961, p. 679)

Figure 1 Traffic Flow and Trip Time



[Note] Q: traffic demand
F: traffic flow

Figure 2 *Q-F* Conversion Function

1 follows the Pigou condition:

$$MC(Q)=AC(Q)+Q\times dAC(Q)/dQ, \quad \cdots\cdots(2)$$

which provides

$$MC(Q)=AC(Q)\times[1+\{Q/AC(Q)\}\times\{dAC(Q)/dQ\}]. \quad \cdots\cdots(3)$$

where Q, $AC(Q)$, and $MC(Q)$ respectively mean demand level[10], average cost, and marginal cost.

The principle of marginal cost fare-setting suggests that the best allocation of a fixed highway capacity can be achieved only for such traffic demand level that the marginal social benefit is equal to the marginal social cost. This traffic demand level shall be referred to as the optimal demand level. Then, the optimal congestion tolls that would bring about the best allocation of a fixed highway capacity should be equal to the optimal demand level multiplied by the value of derivative of the average social cost with respect to demand level (for the formulation of Equation 2) or equal to the average social cost at the optimal demand level multiplied by the value of its elasticity (for the formulation of Equation 3). It should be noted that the derivative in the second term of the right side of Equation 2 is the derivative of the average cost *with respect to Q* (instead of *with respect to traffic flow F*), and that the elasticity in Equation 3 is the elasticity of the average cost *with respect to Q* (instead of *with respect to traffic flow F*).

The aforementioned would tell us that, for our purpose of finding optimal congestion tolls, the marginal social cost curve which corresponds to the average social cost curve XX' in Figure 1 should be obtained through the following formulation:

$$MSC_f(F)=ASC_q\{G(F)\}+G(F)\times dASC_q\{G(F)\}/dG(F)$$

or

$$MSC_f(F)=ASC_f(F)+G(F)\times dASC_q\{G(F)\}/dG(F), \quad \cdots\cdots(4)$$

where $MSC_f(\cdot)$: marginal social cost as a function of F,
$ASC_f(\cdot)$: average of social cost as a function of F,
$ASC_q(\cdot)$: average social cost as a function of Q,
Q: traffic demand,
F: traffic flow,
$Q=G(F)$: Q–F conversion function as shown, for instance, by Figure 1.

The broken line curve MSC in Figure 1 is likely to have been obtained through the following formulation:

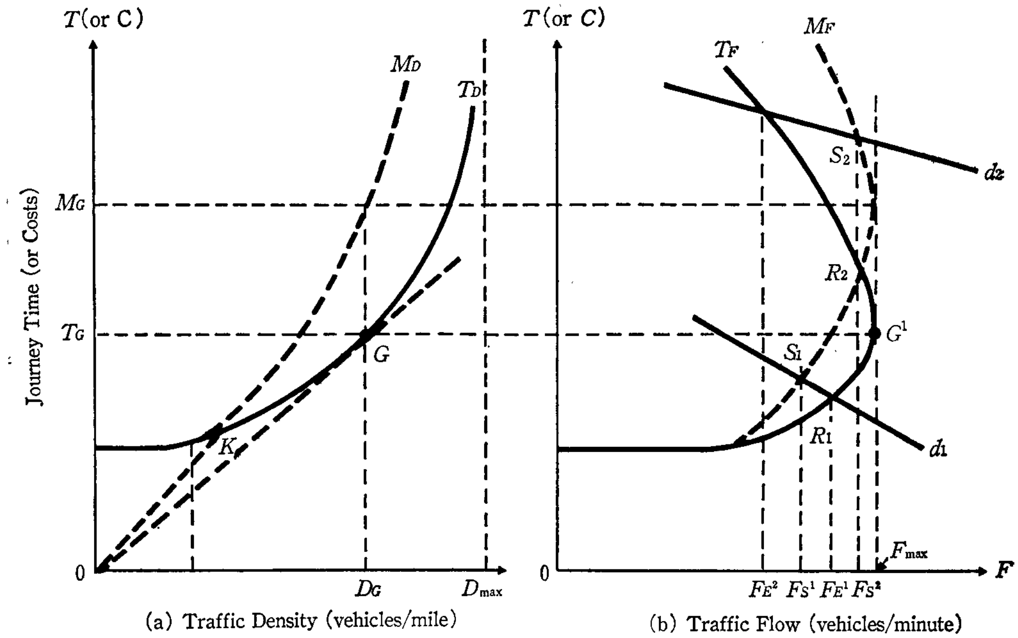
$$MSC_f(F)=ASC_f(F)+F\times\{dASC_f(F)/dF\}. \quad \cdots\cdots(5)$$

It can be easily noticed that there is a clear difference between the second term of the right-hand side of Equation 4 and that in Equation 5[11]. The formulation given by Equation 5, therefore, seems to be false for the analysis of the optimal congestion tax.

Meanwhile Else (1981) posed the same kind of question as have been addressed in this paper, on Walters' argument that the optimum traffic flow can not occur at the maximum traffic flow level of the expressway (*i.e.*, at the level of the traffic flow capacity of the expressway) since the marginal social cost is infinite. Else remarked[12]:

> "(Walters' theory) presupposes that the cost of adding to the traffic flow is the relevant measure of marginal social cost. This may be doubted. ...... the optimal level of traffic should be defined in terms of the marginal cost of adding to the number of vehicles rather than the costs of adding to the traffic flow. ...... (It is because) a decision by an individual to use a road is essentially a decision to add to the number of vehicles on that road, but whether or not it increases the traffic flow depends on the volume of traffic already on the road ...... (The marginal social costs is supposed to) reflect the additional costs imposed by an additional vehicle on a traffic already on the road."

Based on this consideration, Else derived the marginal social cost curve which is labelled M_F as shown in Figure 3(b). This curve can be constructed, from the



Source: Else (1981, p. 220)

Figure 3 Journey Time, Traffic Density, and Traffic Flow engineering relationship between traffic density D and journey time T as given by Figure 3(a), in a way clearly indicated by the relations between the diagram in Figure 3(a) and the diagram in Figure 3(b). Note that the curve M_F slopes backwards over the range where both traffic density level and traffic flow level are increasing, but that it bends backwards as traffic flow level falls back from its maximum capacity with the traffic density level continuing to increase.

Nash (1982) has argued that criticisms raised by Else are for the most part unjustified and that Else's redefinition of marginal social cost relative to numbers of vehicles rather than flow can not be acceptable within the conventional theoretical framework of economics since all demands relate to flows, not stocks. "Following this argument" said Button and Pearman[13], "(they) support the conventional position that a social optimum with flow at greater than the maximum is unattainable, and thus it seems that Nash has re-established the authority of the conventional wisdom in this case."

I would think, however, that Else's criticisms arguing that Walters' analytical approach incorrectly defines the marginal social cost curve is appropriate and so is his rejoinder (1982) to Nash's criticism (1982), and that Else should have the credit of clearly pinpointing the shortcomings of Walters' approach as well as of having derived the backward-bending marginal social cost curve M_F in Figure 3. But it should be noted that I would have some reservations about Else's assumption that an additionally generated traffic demand can be instantaneously absorbed into the volume of traffic already on the road without any delay.

Despite the fact that Else derived the curve M_F based on this debatable assumption, I would think that the bending-backwardness of the marginal social cost curve in the F–C dimension of Figure 3(b) still remains valid. The reasons to support the validity of the bending-backwardness of the marginal social cost curve, have already been abstractly discussed through Equation 4 in this Section, and shall be furthermore discussed in a more concrete way in the following.

3. Marginal Social Cost as a Function of Traffic Flow: Diagrammatic Approach

Now we are in a position to ask ourselves how to draw diagrammatically the appropriate marginal social cost curve which corresponds to the average social cost curve XX' in Figure 1. In order to answer this question, let us introduce a relationship between traffic demand Q and traffic flow F as given by the bell-shaped Q–F conversion curve OP'_4CQ'[14] in Figure 2. The rationale of this curve would be:

(1) If the level of traffic demand which is generated for the road services of the expressway for a specific period of time is low, say below Q_c, then the vehicle can drive along the expressway at the speed of the lowest cost[15].

At the same time, the vehicle does not have to wait at the entrance of the expressway at all. We shall assume that Q_c is the traffic demand level at which traffic congestion begins to result in the slowing down of vehicle speed.

(2) If the level of traffic demand is between Q_c and Q_f, then the vehicle speed continues to slow down as the level of traffic demand increases, but the vehicle does not have to wait at the entrance of the expressway. We shall assume that Q_f is the level of traffic demand at which the queue of vehicles starts to be observed at the entrance.

(3) If the level of traffic demand is between Q_f and Q_m, then the vehicle speed continues to slow down as Q increases. The queue of vehicles is increasingly observed. We shall assume that Q_m is the level of traffic demand at which the level of traffic flow is maximized.

(4) If the traffic demand level is beyond Q_m, then the level of traffic flow continues to reduce as Q increases. The rather long queue of vehicles is meanwhile observed. The traffic flow is almost at a standstill when the level of traffic demand approaches C.

It should be mentioned that the above-mentioned rationale for the shape of Q–F conversion curve should require the following two assumptions in addition to our previous Assumptions 1 and 2:

Assumption 3

All traffic demand will be uniformly generated along time-axis during a specific period of time.

Assumption 4

There is no traffic congestion either in access roads to or in egress roads from the expressway.

As to the cost element, we set the following two assumptions:

Assumption 5

The average social cost for the traffic demand level Q is the fuel cost necessary for the vehicle to drive along the expressway from its entrance to exit at the speed corresponding to the level of traffic flow which changes according to the level of traffic demand Q[16)].

Assumption 6

Time loss, uncomfortability and other types of non-monetary external diseconomies accruing from the highway congestion, shall be explicitly excluded from the components of the average social cost[17)].

Under our assumptions, it would perhaps be instructive to construct Figure 4. Suppose we have, for the cost side, the average social cost curve ASC_q as shown in the second quadrant. Suppose furthermore that we have the Q–F conversion curve as shown in the fourth quadrant. From the curve ASC_q, we obtain in the second

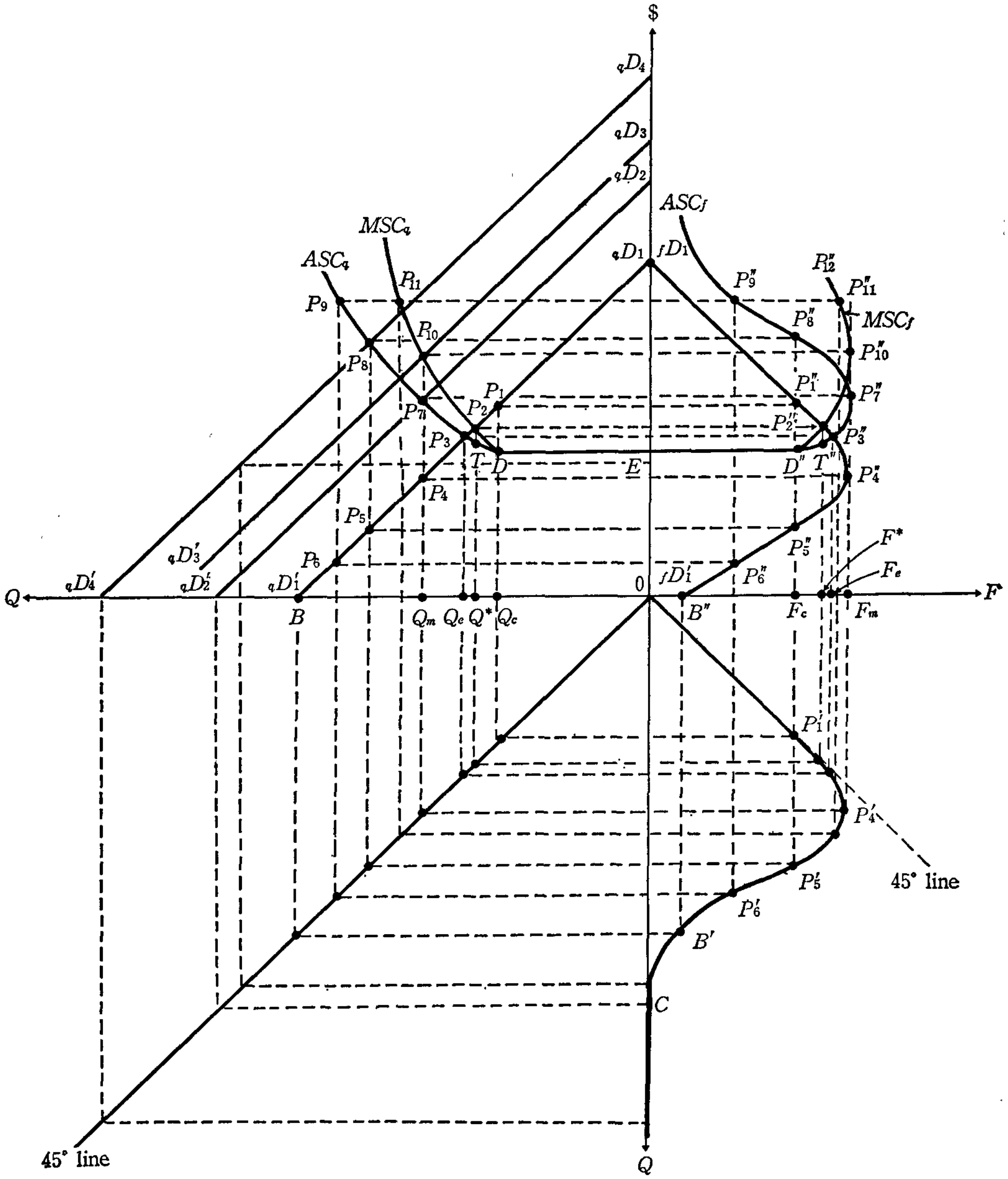


Figure 4 Traffic Demand (Q), Traffic Flow (F), Average Social Cost (ASC), and Marginal Social Cost (MSC)

Demand Curve	Order-Relationships
${}_qD_1\ {}_qD'_1$	$F^*<F_e<F_m$
${}_qD_2\ {}_qD'_2$	$F^*<F_e=F_m$
……	$F^*=F_e<F_m$
……	$F_e<F^*<F_m$
${}_qD_3\ {}_qD'_3$	$F_e<F^*=F_m$
${}_qD_4\ {}_qD'_4$	$F_e<F^*<F_m$

[Note] F_e: equilibrium level of traffic flow
F^*: optimal level of traffic flow
F_m: maximum level of traffic flow

Table 1 Order-Relationships among F_e, F^*, and F_m

quadrant the marginal social cost curve MSC_q through Equation 1. Now we can draw curves of ASC_f and MSC_f as shown in the first quadrant from the curves of ASC_q and MSC_q respectively, by use of the 45-degree line in the third quadrant and the Q–F conversion curve in the fourth quadrant. The curve of ASC_f is the average social cost curve which is expressed as a function of traffic flow F, while the curve of MSC_f is the marginal social cost curve which is also expressed as a function of F. From this setting can be clearly seen the following characteristics which contradict the interpretation by Walters.

(1) Marginal social cost curve MSC_f has the positive constant value at pont P''_{10}. when the traffic flow is at its maximum level F_m to which the traffic demand level Q_m correspond.

(2) On the backward sloping part of the average social cost curve ASC_f, the marginal social cost curve is defined as the curve of $P''_{10}\ P''_{11}\ P''_{12}$.

For the demand side, suppose we have a traffic demand curve ${}_qD_1\,{}_qD'_1$ as shown in the second quadrant. From this curve we obtain the curve ${}_fD_1\,{}_fD'_1$ in the first quadrant. This obtained curve is the traffic demand curve which is expressed as a function of traffic flow F. Note that *no part* of the traffic demand curve presented in the F–\$ dimension shall appear, unlike the traffic demand curve shown in Figure 3(b) by Else as well as in a number of other literatures, when the traffic flow level is greater than the flow capacity F_m of the expressway, though this would not discount Else's credit that he has suggested that the marginal social cost would have the shape of bending-backwardness as shown by curve M_F in Figure 3.

Be that as it may, it is now easy for us to identify the equilibrium level of traffic demand Q_e, the optimal level of traffic demand Q^*, the equilibrium level of traffic flow F_e, the optimal level of traffic flow F^*, as well as the optimal level of congestion tolls P_2T (in the second quadrant) or P''_2T'' (in the first quadrant). Therefore, it turns out that F^* is the lowest among F_e, F^* and F_m if the demand curve is given by ${}_qD_1\,{}_qD'_1$. To explore other possible order-relationships among F_e, F^*, and F_m, let us introduce in the second quadrant of Figure 3 additional three alternative traffic demand curves; ${}_qD_2{}_qD'_2$, ${}_qD_3{}_qD'_3$ and ${}_qD_4{}_qD'_4$ Careful diagrammatic investigation of those three curves and the demand curve ${}_qD_1{}_qD'_1$ as well, would provide us with Table 1 which furnishes the information on order-relationships among F_e, F^* and F_m for various demand curves. This table shows us the following third characteristic that contradicts the interpretation by Walters:

(3) For the traffic demand curve ${}_qD_3{}_qD'_3$, the optimal level of traffic flow is identical to the maximum level of traffic flow.

4. Marginal Social Cost as a Function of Traffic Flow: Numerical Example Approach

Having shown both theoretical and diagrammatical explanations on the characteristics of the marginal social cost function (as derivative of the total social cost with respect to traffic demand) expressed as a function of traffic flow, we now proceed to a numerical example in order to see such characteristics more tangibly. We shall begin by setting our notational definitions;

x : level of traffic demand to be generated per unit of time,
y : level of monetary amount, representing average cost, marginal cost, or price,
z : level of traffic flow per unit of time,

where $0 \leqq z \leqq x$.

Suppose that the average social cost function and the traffic demand function are given in the x-y dimension respectively by:

$$y-x=0. \quad \cdots\cdots (6)$$
$$y+x-5a/4=0 \quad \cdots\cdots (7)$$

Suppose furthermore that the x-z conversion function (*i.e.*, demand-flow conversion function) in the x-z dimension is given by:

$$\left.\begin{array}{ll} z+x(x-a)=0 & (\text{for } 0 \leqq x \leqq a) \\ z=0 & (\text{for } x>a) \end{array}\right\} \quad \cdots\cdots (8)$$

where $0<a\leqq 1$[18].

From Equation 6, we can derive the marginal social cost function in the x-y dimension:

$$y-2x=0. \quad \cdots\cdots (9)$$

In Figure 5 are shown line OA for Equation 6, line DE for Equation 7, curve OBC for Equation 8, and line OM for Equation 9. By use of the x-y conversion function as expressed by Equation 8, we can obtain from Equation 6 the average social cost function for the y-z dimension:

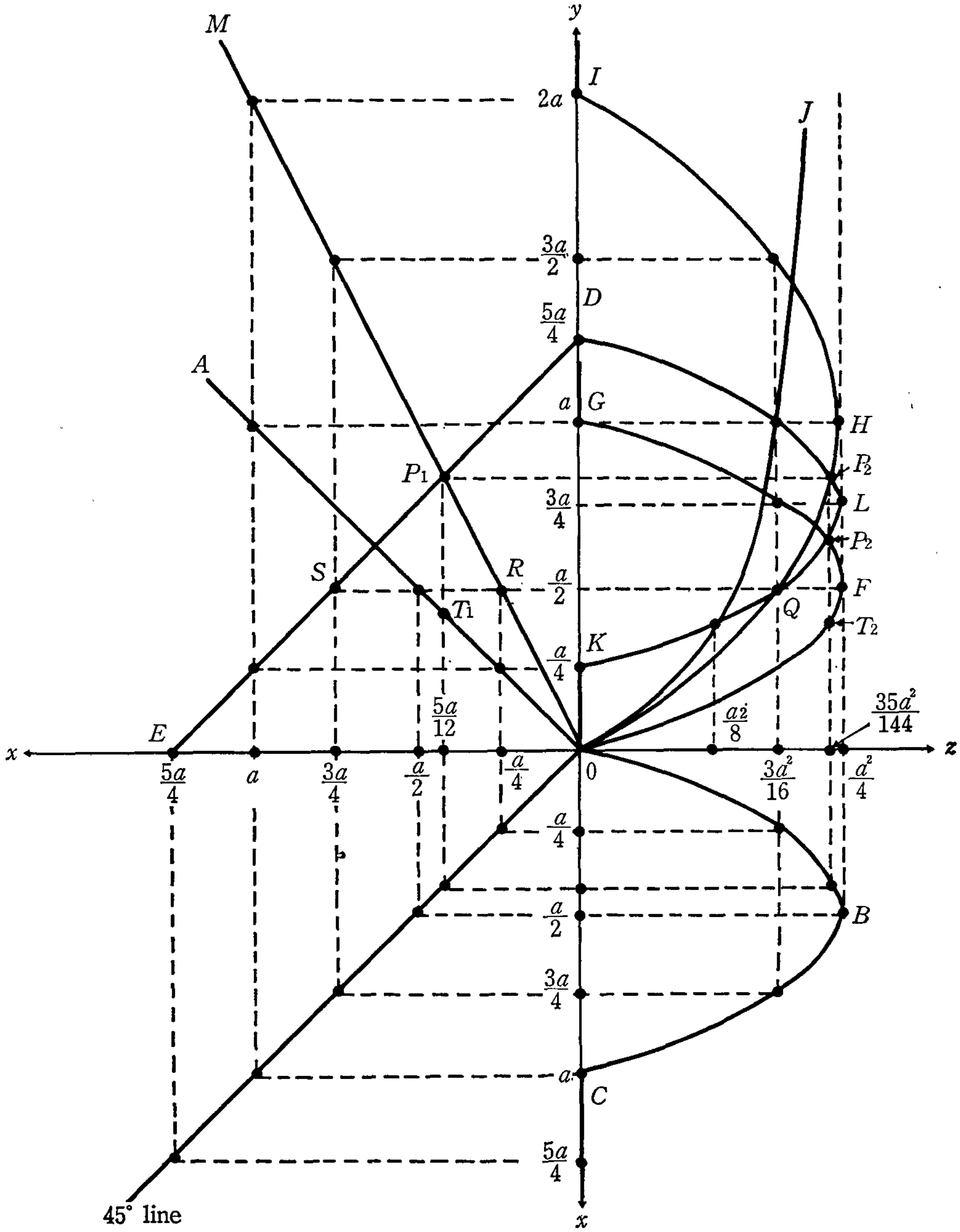


Figure 5 Role of the x-z Conversion Curve

$$z+y(y-a)=0 \qquad (0<a\leqq 1) \qquad \cdots\cdots (10)$$

which is graphically shown by curve *OFG* in Figure 5. Caution is advised here: the average social cost y expressed by Equation 10 as a function of the level of traffic flow z, means the division of "the total fuel cost for the traffic demand level x" by "the traffic demand level x" when the traffic flow level z corresponds to the traffic demand level x through the x-z conversion function expressed by Equation 8.

Also by means of Equation 8, we can obtain from Equation 9 the marginal social cost function for the y-z dimension:

$$4z+y(y-2a)=0 \qquad (0<a\leqq 1). \qquad \cdots\cdots (11)$$

which is graphically shown by curve *OHI* in Figure 5. It should be noted that the marginal social cost y expressed by Equation 11 as a function of traffic flow level z is the derivative of "the total fuel cost y for the traffic demand level x" with respect to "the traffic demand level x" when the traffic flow level z corresponds to the traffic demand level x through Equation 8. At the same time, it should be clearly kept in mind that Equation 11 does *not* imply "the derivative of the total fuel cost y with respect to the level of traffic flow z[19]" that is given by:

$$y-[\{a-(a^2-4z)^{1/2}\}/2+z(a^2-4z)^{-1/2}]=0 \qquad \cdots\cdots (12)$$

which is graphically shown by curve *OJ* in Figure 5. The function expressed by Equation 12 is a fallacious marginal social cost curve in light of our discussion in Section 2. Comparing, in the y-z dimension, the y-value of the positive-slope part of the marginal social cost curve *OHI* with the y-value of the "fallacious" marginal social cost curve *OJ*, it can be pointed out that:

(1) The value of the "fallacious" marginal social cost curve *OJ* is always greater than that of the "legitimate" marginal social cost curve *OHI* for any positive level of traffic flow.

(2) The ratio of the y-value of the "fallacious" marginal social cost against the y-value of the "legitimate" marginal social cost, increases as the level of traffic flow increases. Their ratio would be, for example, 1.00 for $z=0+dz$, 1.04 for $z=a^2/16$, 1.10 for $z=a^2/8$, 2.00 for $z=3a^2/16$, and infinity for $z=a^2/4$.

From Equation 7, we can obtain by use of Equation 8 the traffic demand function for the y-z dimension:

$$\left.\begin{array}{ll} z+(y-a/4)(y-5a/4)=0 & (\text{for } a/4\leqq y\leqq 5a/4) \\ z=0 & (\text{for } 0\leqq y<a/4 \text{ or } y>5a/4 \end{array}\right\} \cdots\cdots (13)$$

which is graphically shown by curve *OKLD* in Figure 5. As can be clearly indicated, the optimal levels of traffic demand and traffic flow are equal to $5a/12$ and $35a^2/144$

respectively. The optimal level of congestion tax is equal to the length of line segment P_1T_1 or P_2T_2. Therefore, the use of the fallacious marginal social cost curve for the estimation of the optimal congestion tolls would lead us to the unreasonably high overestimation for the optimal level of the congestion tolls, especially when the level of traffic flow is at or around the maximum capacity of the expressway.

5. Demand-surface Analysis of Traffic Congestion

Thus far, we have examined the marginal social cost function associated with the use of expressway, assuming that the demand curve for expressway services is fixed to be independent of the change in the level of traffic demand. It would be, however, more reasonable for us to consider the situation where the demand curve would shift downward as the level of traffic congestion (*i.e.*, traffic demand level) increases.

To meet this analytical rationale, it will be helpful to construct the *demand-surface* as illustrated by Figure 6. In this diagram, the \$-axis represents the cost for consuming expressway services, the Q-axis the level of *actual* demand, and the M-axis the level of *premised* demand. The actual demand Q has the same concept as the demand level usually defined in the conventional textbook. The preposed word "actual" is attached to this terminology simply in order to emphasize the difference of the usual demand from the premised demand. The premised demand is a sort of presumde demand, serving as an instrument to enable the (actual) demand curve to shift flexibly as the level of traffic congestion varies.

For any given level of Q, the height of the demand-surface with the shape of a quarter-cut bullet remains the same for the case the presumed demand level M falls within the range from zero to M_1, and gradually reduces as the level of M icreases from M_1 to F. If the level of M is greater then F, then no demand-surface exists. For any given level of M, on the othre hand, the height of the demand-surface becomes lower as the level of Q increases. On the top of these, it should be noted that the curve DBD'_2 is the locus of the point (on the demand-surface) satisfying the condition that Q is equal to M. In other words, the projection of the curve DBD'_2 against the Q–M plane forms the 45 degree line OED'_2 whose point E, for example, corresponds to point B on the demand-surface. This point B in turn assures that the actual demand level Q_1 be equal to the premised demand level M_1.

The verbal implications of these characteristics of the demand-surface given by Figure 6 are (i) that the traffic runs smoothly before the demand level Q reaches Q_1 where the traffic congestion will start, and (ii) that the dead-congestion with solid mass of immobile vehicles would take place when the demand exceeds the level of F. More specifically speaking:

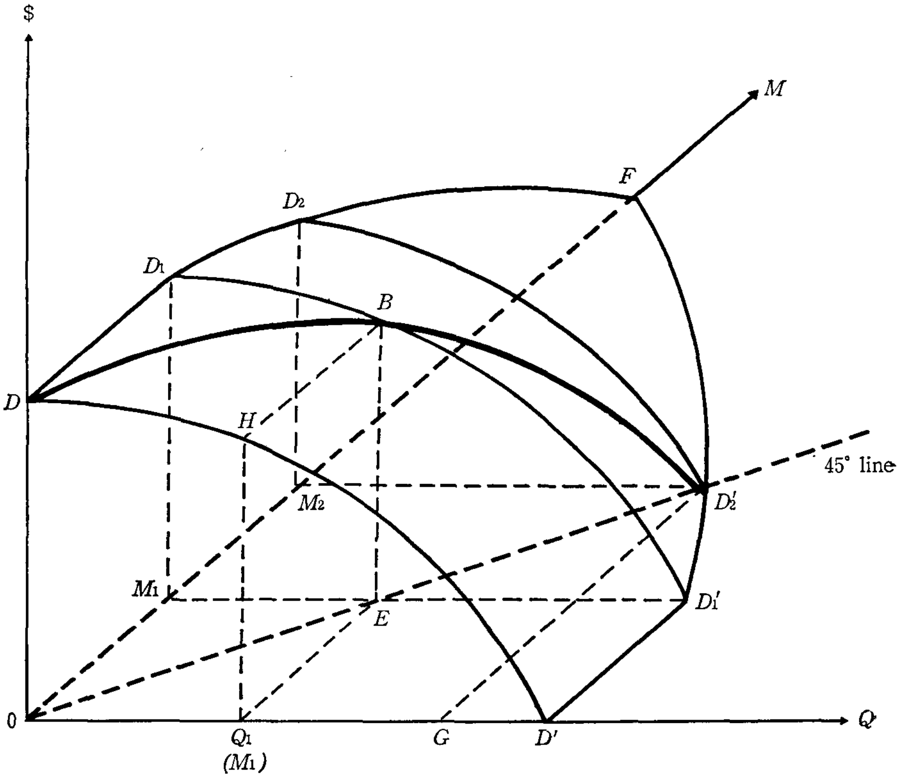


Figure 6 Demand-surface with Traffic Congestion

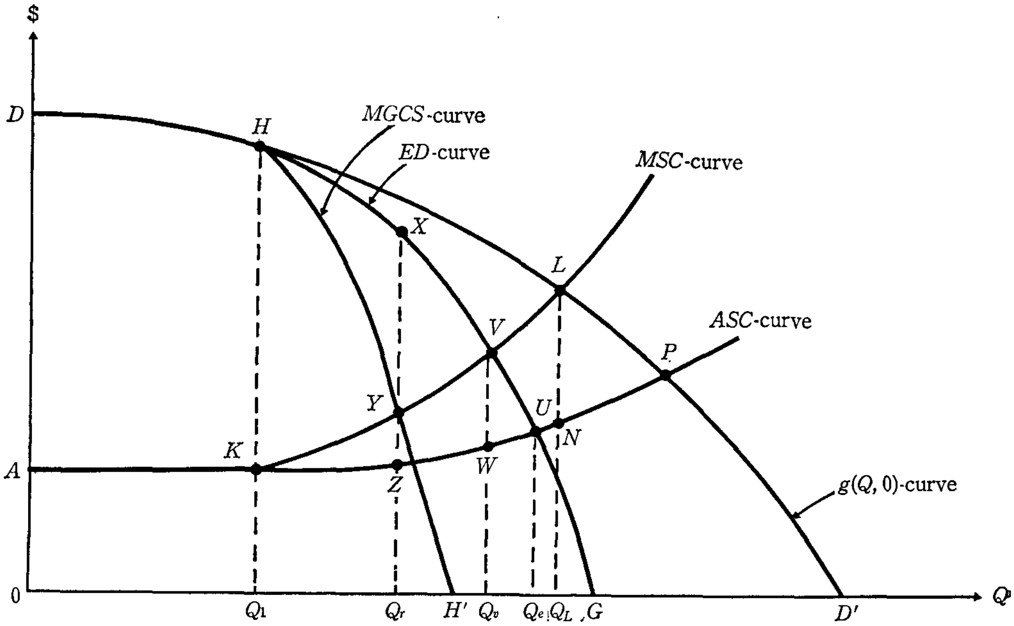


Figure 7 Various Curves related with Demand-surface

(1) If each driver premises that the demand level for expressway services would be zero (or nearly zero) and therefore can expect no traffic congestion, then the (actual) demand curve is given by curve DD'.

(2) If each driver premises that the demand level would be M_1 and therefore still can expect no traffic congestion, then the demand curve is given by curve $D_1D'_1$ which is identical with curve DD'.

(3) If each driver premises that the demand level would be M_2 and therefore should anticipate some degree of congestion, then the demand curve is given by curve $D_2D'_2$ the height of which at any level of Q is lower than that of curve DD'.

(4) If each driver premises that the demand level would be F and therefore should anticipate no movement of traffic on the jammed expressway, then there emerges no demand at all.

It is obvious that, as discussed in detail by Kawashima (1975, 1981)[20], only the points situated on the curve DBD'_2 along the demand-surface can satisfy the necessary condition to become an equilibrium point for a given cost-surface[21] in the Q–M–\$ space of Figure 6. In this sense, the projection of the curve DBD'_2 against the Q–\$ plane shall be called *equilibrium demand curve.* Figure 7 presents, by curve DHG with the name of ED-curve, the equilibrium demand curve which corresponds to curve DBD'_2 in Figure 6. The demand curve DD' in Figure 6 which can be conceived when the premised demand level is between zero and M_1, is also shown in Figure 7 by curve DHD' with the name of $g(Q, 0)$-curve.

We have now arrived at the inevitable position of asking ourselves what kind of shape the marginal gross consumers surplus ($MGCS$) curve[22] would have in the Q–\$ plane of Figure 7 for the demand-surface given by Figure 6, since the $MGCS$ curve plays a key role together with the marginal social cost curve when we attempt to find out the optimal level of congestion tax. In considering that the $MGCS$ curve is logically identical with the demand curve in the conventional approach of transportation economics[23], and furthermore in light of the fact that each of the ED-curve and the $g(Q, 0)$-curve is anyway considered as a sort of demand curve, we may at this point artlessly wonder whether the $MGCS$ curve would be identical with either curve DHG or curve DHD' in Figure 7. However, this naive inference is misleading as indicated in the following.

Suppose that a demand-surface function is given by:

$$\$=g(Q, M). \quad \cdots\cdots (14)$$

The equilibrium demand function, $ED(Q)$, for this demand-surface is then expressed as:

$$ED(Q)=g(Q, Q) \quad \cdots\cdots (15)$$

In this setting, the gross consumers surplus function, $GCS(Q)$, can be obtained through integrating $g(q, Q)$ with respect to q from zero to Q:

$$GCS(Q)=\int_0^Q g(q, Q)\,dq^{24)}. \quad (16)$$

Hence, the marginal gross consumers surplus function, $MGCS(Q)$, is given by:

$$MGCS(Q)=dGCS(Q)/dQ=d\left(\int_0^Q g(q, Q)dq\right)/dQ \quad (17)$$

One of the conceivable curves for the marginal gross consumers surplus expressed by Equation 17, is provided in Figure 7 by curve DHH' with the name of $MGCS$-curve. Let us withal introduce, in addition to the ED-curve, $g(Q, 0)$-curve, and $MGCS$-curve in Figure 7, the average social cost curve $ASC(Q)$ given by curve AKP and its associated marginal social cost curve $MSC(Q)$ given by curve AKL. From this arrangement, it can be pointed out:

(1) The curve of the marginal gross consumers surplus is identical with the equilibrium demand curve when the traffic demand level is under Q_1. It is because no externality[25] exists at such a level of traffic demand.

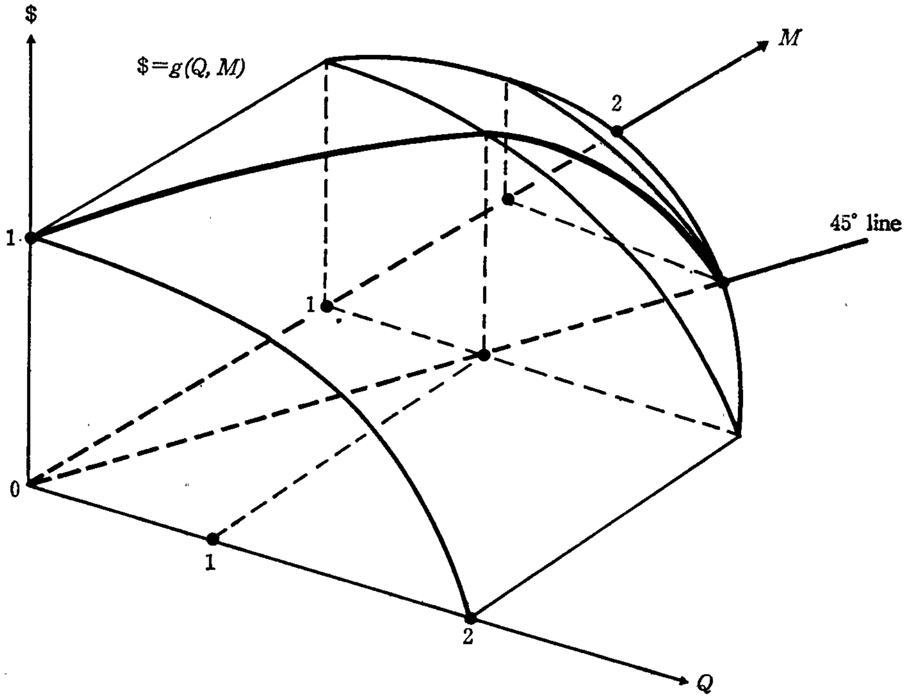
(2) The curve of the marginal gross consumers surplus begins to situate below the equilibrium demand curve when the traffic demand exceeds Q_1, with the separation between these two curves expanding as the level of traffic demand increases. It is because external agglomeration diseconomies start to emerge when the traffic demand level exceeds Q_1, with the magnitude of diseconomies enlarging as the traffic demand level increases.

(3) By the same reason as for the relationship *between* the curve of the marginal gross consumers surplus *and* the equilibrium demand curve, the marginal social cost curve is identical with average social cost curve when the traffic demand level is under Q_1, while they start to separate from each other when the traffic demand level exceeds Q_1.

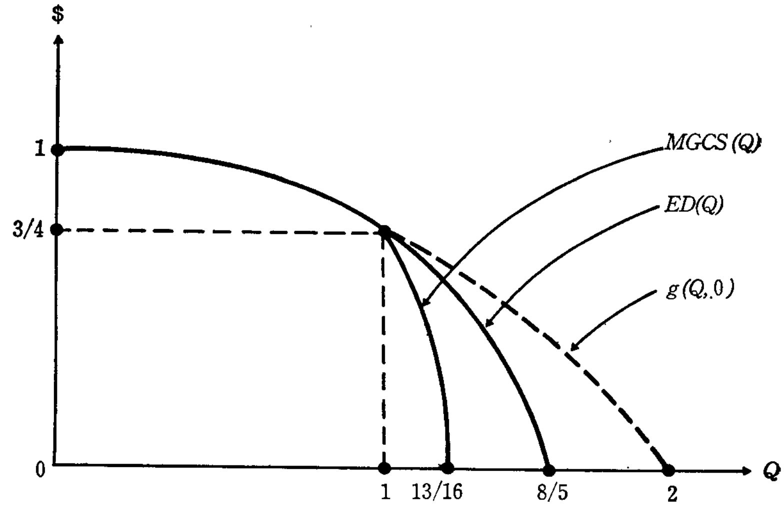
(4) A kind of agreeable symmetricalness[26] is recognized with respect to the cost side curves and the demand side curves as to the discrepancy of the marginal value from its companion curve[27].

(5) The net consumers surplus is maximized when the marginal social cost curve AKL cuts the marginal gross consumers surplus curve DHH' at point Y. Accordingly, the optimal level of congestion tax is equal to the distance between X and Z[28].

(6) In case we assume that the demand curve which we conventionally use is considered as, in a very strict sense, the equilibrium demand curve, then the optimal level of congestion tax obtained through this conventional demand curve is equal to the length between V and W[29]. This length is considerably shorter than the correct optimal level of congestion tax expressed by the length between X and Z.



(a) Demand-surface



(b) Curves derived from Demand-surface

Figure 8 Numerical Example of Demand-surface Approach

For the sake of ascertaining the reasonability of the above understandings of ours, consider the numerical example for the demand-surface as given by:

$$g(Q,M)\begin{cases}=-Q^2/4+1 & (0\leqq M\leqq 1)\\ =-Q^2/4+1-(M-1)^2 & (1\leqq M\leqq 2)\end{cases}$$

From this, we can derive three major functions peculiar to the demand-surface framework:

(1) function for equilibrium demand: $ED(Q)=g(Q,Q)$

$$ED(Q)\begin{cases}=-Q^2/4+1 & (0\leqq Q\leqq 1)\\ =-5Q^2/4+2Q & (1\leqq Q\leqq 8/5)\end{cases}$$

(2) function for gross consumers surplus: $GCS(Q)=\int_0^Q g(q,Q)dq$

$$GCS(Q)\begin{cases}=-Q^3/12+Q & (0\leqq Q\leqq 1)\\ =-13Q^3/12+2Q^2 & (1\leqq Q\leqq 24/13)\end{cases}$$

(3) function for marginal gross consumers surplus: $MGCS(Q)=dGCS(Q)/dQ$

$$MGCS(Q)\begin{cases}=-Q^2/4+1 & (0\leqq Q\leqq 1)\\ =-13Q^2/4+4Q & (1\leqq Q\leqq 13/16)\end{cases}$$

Figure 8 diagrammatically illustrates three curves: equilibrium demand curve, curve of marginal gross consumers surplus, and traffic demand curve for the zero premised demand level. These results of our numerical example[30] for demand-surface reflecting the existence of external diseconomies show:

(1) The *MGCS*-curve is identical with the *ED*-curve when the traffic demand level is below 1.

(2) The *MGCS*-curve begins to situate below the *ED*-curve when the traffic demand level exceeds 1, with the discrepancy between them expanding as the traffic demand level increases.

(3) The $g(Q,0)$-curve is identical with both *MGCS*-curve and *ED*-curve when the traffic demand level is below 1. Both *MGCS*-curve and *ED*-curve begin to situate below the $g(Q,0)$-curve when the traffic demand level exceeds 1.

6. Conclusion: Synthesis

In the preceding sections, we have first studied how both the average social cost function and the marginal social cost function for expressway services can be presented in terms of the traffic flow level. The *Q–F conversion function* has been found to be of significant assistance to that kind of study.

We have then discussed the *demand-surface method* to be employed, in order to take explicit account of the effects of traffic congestion as external agglomeration

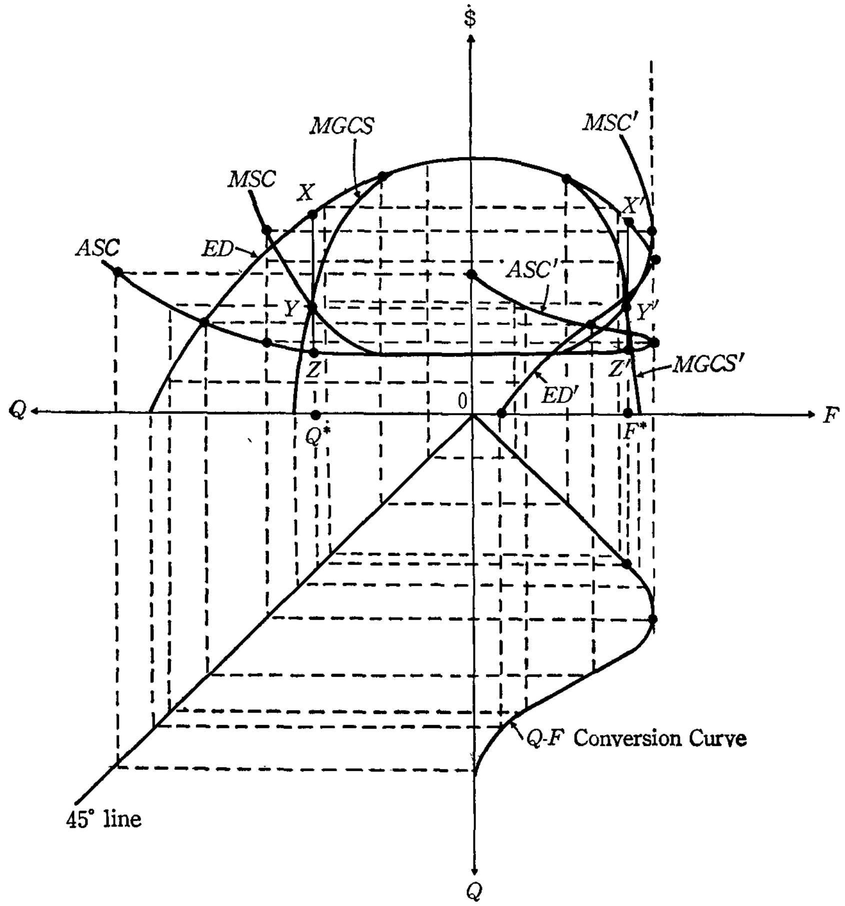


Figure 9 Synthesis

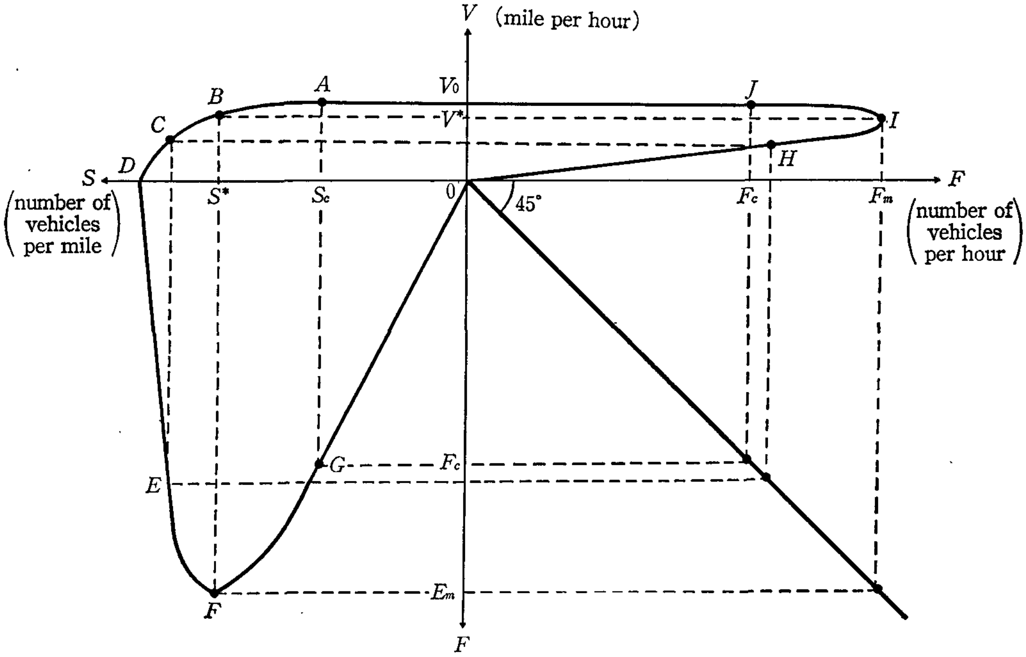
diseconomies upon the demand function for expressway services. One of the striking characteristics of the demand-surface approach is that it would enable us to enjoy the economic analysis of traffic congestion in a scientifically aesthetic way. That is, from the demand-surface approach emerges interesting symmetricalness with respect to the side of cost function and the side of demand function as to the relative position of the marginal curve against its corresponding average curve[31]. It has been demonstrated that, due to this symmetricalness, the optimal level of congestion tax derived through the demand-surface approach is larger than that derived through the conventional approach.

Pursuing the task of synthesizing outcomes from our twin investigations (*i.e.*, one based on *Q–F* conversion function approach and the other based on demand-surface approach) in this paper, we construct Figure 9 which shows how the four curves in the *Q–$* dimension (*ASC* for average social cost, *MSC* for marginal social cost, *ED* for equilibrium demand, and *MGCS* for marginal gross consumers surplus) look like once they are transformed into the *F*–$ dimension. Their corresponding curves in the *F*–$ dimension are, *ASC′* for *ASC*, *MSC′* for *MSC*, *ED′* for *ED*, and *MGCS′* for *MGCS*. It follows from this diagram that the level of the optimal congestion tax in the *Q*–$ dimension is equal to the distance between *X* and *Z*. This level of congestion tax would bring about the optimal traffic demand Q^* which corresponds to point *Y* where the *MSC* curve cuts the *MGCS* curve. It also follows that in the *F*–$ dimension, the optimal level of congestion tax is equal to the distance between *X′* and *Z′* bringing about the optimal traffic flow F^* which corresponds to point *Y′* where the *MSC′* curve cuts the *MGCS′* curve.

In concluding this paper, the scientific stimulation which I have received especially from papers by Walters (1961) and Else (1981) is gratefully acknowledged.

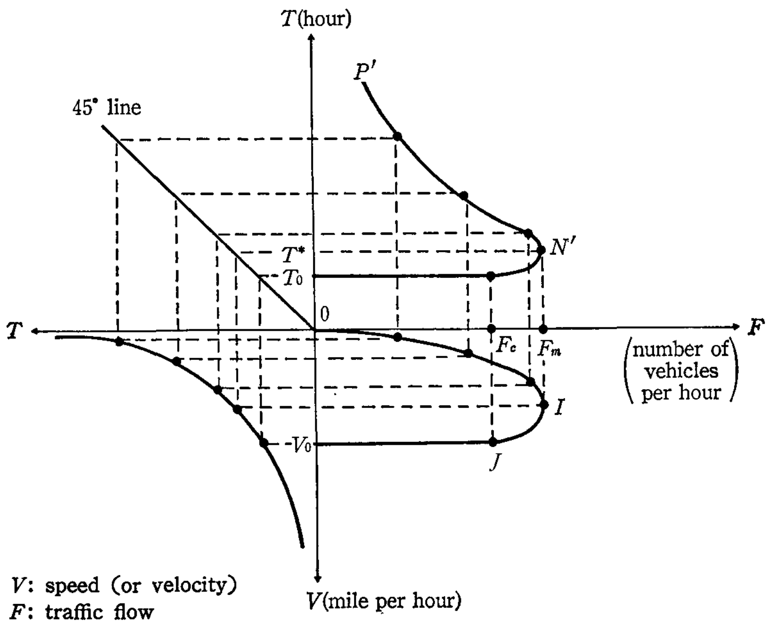
NOTES

1) In this paper, the term of average social cost could be interchangeably used with such words often appearing in transportation literature on highway traffic congestion as average user cost (or simply, average cost), average consumers cost, private unit cost (or simply, unit cost), or marginal private cost. The term of marginal social cost could, on the other hand, be interchangeably used with that of "marginal social cost for users."
2) Throughout this paper, the term of motorist could be interchangeably used with the words of highway user, vehicle operator, car driver, vehicle, car, or automobile.
3) The term of congestion tolls could be interchangeably used in this paper with that of congestion taxes.
4) In the first four Sections of this paper, the marginal social benefit curve is treated as to be identical to the conventional demand curve.
5) For theoretical discussion of optimal congestion tolls, see for example, Knight (1924), Walters (1954, 1961), Beckmann, McGuire, and Winsten (1955), Buchanan (1956), Meyer, Peck, Stenason, and Zwick (1959), Mohring (1964), Johnson (1964), Vickrey (1965), Jansson (1969), Else (1981), and Button and Pearman (1983). For empirical works relating to the estimation of optimal congestion tolls, see for example Mohring (1964), Johnson (1964), Walters (1961), Smeed (1968), Jansson (1969) and Dewees (1979).
6) The highway capacity here would mean the physically possible maximum level of traffic flow of the highway under consideration.
7) The process to derive this bending-backward average social cost curve, can be diagrammatically explained through Figures N-1(a), N-1(b), and N-2. See Walters (1961), Vickrey (1965), and Button and Pearman (1983) for reference.
8) Walters (1961, p. 680).
9) The optimal traffic flow should be regarded as the level of traffic flow at which, the marginal social benefit with respect to traffic demand is equal to the marginal social cost with respect to traffic demand. The marginal social benefit could be defined as



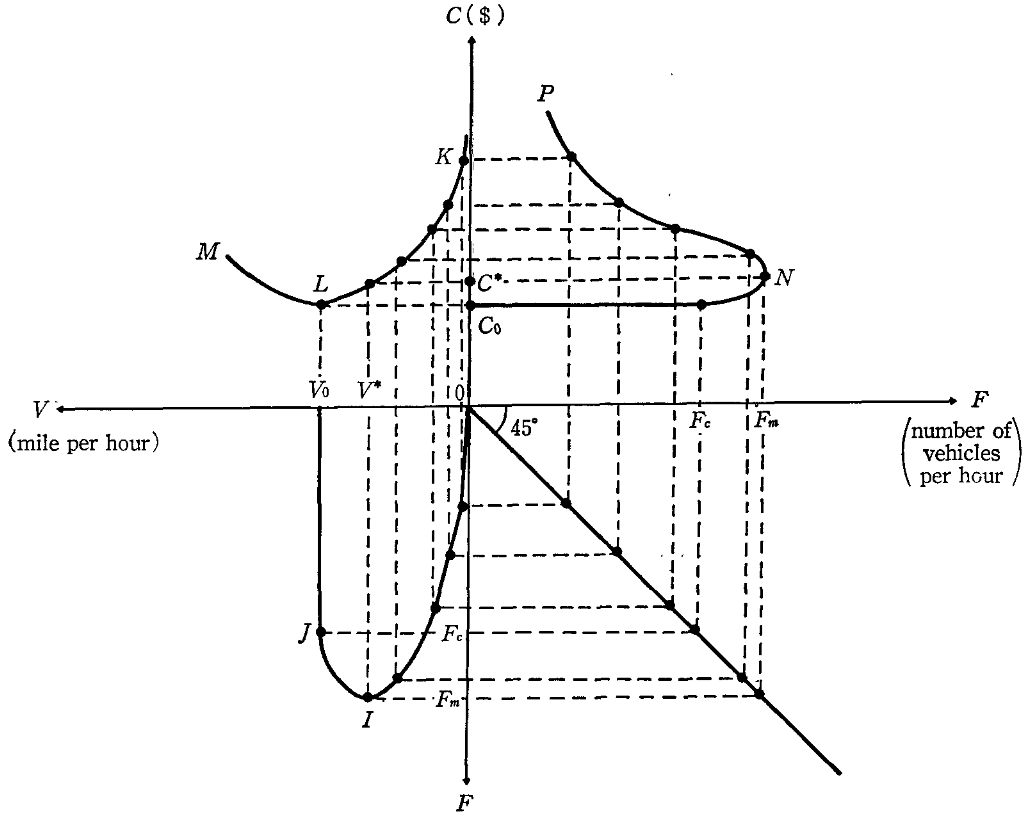
[Note] V: speed (or velocity)
S: traffic density
F: traffic flow

Figure N-1 (a) Speed, Density, and Flow



[Note] V: speed (or velocity)
F: traffic flow
T: trip time

Figure N-1 (b) Traffic Flow, Trip Time, and Speed



[Note] C: fuel cost
V: speed (or velocity)
F: traffic flow
C_0NP: average social cost curve

Figure N-2 Traffic Flow and Fuel Cost

marginal gross consumers surplus.

10) This demand level could be replaced by, for instance, production level or sales level depending on the purpose of study.

11) This difference can also be exhibited by rewriting Equations 4 and 5 as follows:

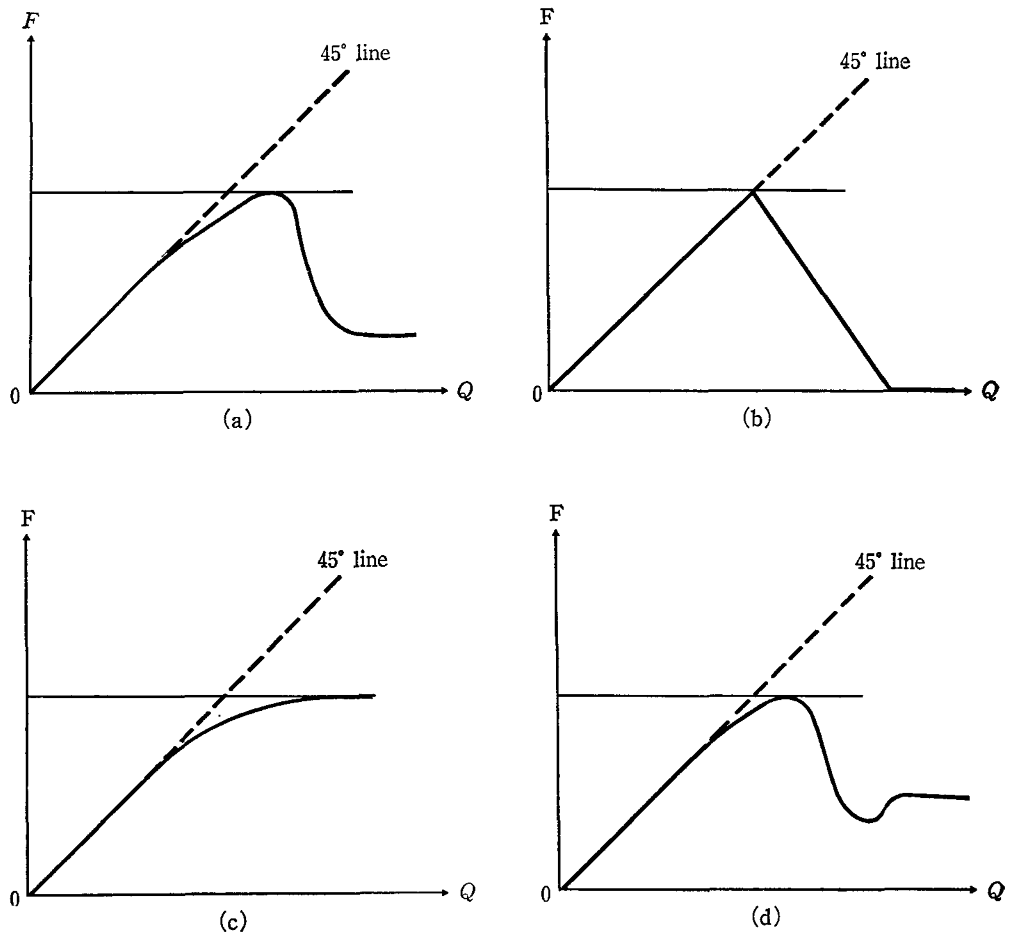
$$MSC_f(F) = ASC_f(F) \times [1 + G(F)/ASC_f(F) \times dASC_q\{G(F)\}/dG(F)] \cdots\cdots (4')$$
$$MSC_f(F) = ASC_f(F) \times \{1 + F/ASC_f(F) \times dASC_f(F)/dF\} \cdots\cdots (5')$$

Equation 5′ is analogous to Equation 1. But this analogy is deceptive, and it is illegitimate to apply Equation 1 to calculate the marginal social cost. Only exception for this is the case when F is always equal to Q, implying that all traffic demand can be smoothly absorbed into traffic flow without any queue at the entrance point of the expressway.

12) Else (1981, pp. 220–221).

13) Button and Pearman (1983, p. 21).

14) Among other possible forms of the Q–F conversion curve are those as presented in



〔Note〕 Q: traffic demand
F: traffic flow

Figure N-3 Possible Q-F Conversion Curves

Figure N-3. The author is presently conducting a primitive but fundamental simulation experiment to identify the suitable functional form for a standard Q–F conversion curve.

15) We shall assume that the maximum speed limit of the expressway is such that the fuel cost is minimized. Therefore, if the speed reduces from the maximum speed limit, the fuel cost is postulated to increase.

16) If we want to include into the average social cost the "fuel cost for idling the engine of the vehicle" necessary for waiting at the entrance point of the expressway when Q (*i.e.*, traffic demand generated per unit of time) is greater than F (*i.e.*, traffic flow observed per unit of time), then the average social cost shall be given by:

"average fuel cost necessary for driving when the traffic demand level is Q"

plus "$C_I \times W(Q)$"

where C_I: fuel cost for idling the engine per unit of time,

$W(Q)$: expected waiting time at the entrance of the expressway when the traffic demand level is Q.

17) These external diseconomies are exclusively reflected by the demand-surface which will be discussed in the later section. As can be seen from that section, the approach demonstrated in this paper is different from the ordinary approach which requires both average cost curve and demand curve to reflect such diseconomies.

18) This condition for the value of "a" is necessary to ensure the traffic flow level z to be always less than or equal to the traffic demand level x.

19) "The derivative of the total fuel cost y with respect to the traffic demand level x" is identical to "the derivative of the total fuel cost y with respect to the traffic flow level z" only when all the volume of generated demand traffic can be instantaneously absorbed into the volume of traffic flow.

20) He developed these papers under the stimulus of Buchanan (1965) and Rabenau and Stahl (1974) at the initial stage in conceptualizing his ideas on the demand-surface framework.

21) One of the simplest functions for the cost-surface would be given by:

$$\$ = a + bQ^2 + 0 \times M \qquad (a, b \geqq 0).$$

22) This curve is considered in this paper as the marginal social benefit curve as explained earlier.

23) If the demand curve is given by $\$ = g(Q)$, then the $MGCS$ curve can be obtained as follows to become identical to the demand curve:

$$MGCS(Q) = d\left\{\int_0^Q g(q)dq\right\}/dQ = g(Q).$$

24) Note that, in the demand-surface framework, $GCS(Q)$ can not be obtained through integrating the equilibrium demand function $ED(q)$ with respect to q from zero to Q. Exception for this is the case where the shape of $g(Q, M)$ remains the same for all M, when we can obtain $GCS(Q)$ by simply integrating $ED(q)$ with respect to q from zero to Q in the same way as we obtain the gross consumers surplus by integrating the demand function $D(q)$ with respect to q from zero to Q in the conventional approach.

25) More precisely, "no externality" means here "neither external agglomeration economies nor external agglomeration diseconomies."

26) This kind of symmetricalness with respect to a pair of cost side curves and a pair of demand side curves does not come out in the conventional approach for the analysis of the optimal congestion tax. It is because, in the conventional approach, the curve of the marginal gross consumers surplus is usually supposed to be identical to the demand curve. Therefore no pair of curves exists for the demand side, but there exists only one curve for the demand side.

27) The companion curve for the marginal social cost curve is the average social cost curve, while that for the marginal gross consumers surplus curve is the equilibrium demand curve.

28) Note that the optimal level of congestion tax is *neither* equal to the distance between V and W which corresponds to the intersection V where the marginal social cost curve

Case 1 :	Case 3 :
$g(Q, M)=a(1-Q/b)M \quad (a, b>0)$	$g(Q, M)=-(M-1)^2+2+0\times Q$
$ED(Q)=aQ-aQ^2/b$	$ED(Q)=-(Q-1)^2+2$
$GCS(Q)=aQ^2-aQ^3/(2b)$	$GCS(Q)=-Q^3+2Q^2+Q$
$MGCS(Q)=2aQ-3aQ^2/(2b)$	$MGCS(Q)=-3Q^2+4Q+1$
Case 2 :	**Case 4 :**
$g(Q, M)=-aQ/b-aM/c+a \quad (a, b, c>0)$	$g(Q, M)=1-Q^2-(M-1/2)^2$
$ED(Q)=-aQ(1/b+1/c)+a$	$ED(Q)=-2Q^2+Q+3/4$
$GCS(Q)=\{-aQ^2(1/b+2/c)+2aQ\}/2$	$GCS(Q)=-4Q^3/3+Q^2+3Q/4$
$MGCS(Q)=-aQ(1/b+2/c)+a$	$MGCS(Q)=-4Q^2+2Q+3/4$

[Note] Q: actual demand level
M: premised demand level
$g(Q, M)$: demand-surface function
$ED(Q)$: equilibrium demand function
$GCS(Q)$: gross consumers surplus function
$MGCS(Q)$: marginal gross consumers surplus function

Table N-1 Examples of Demand-surface

cuts the equal demand curve, *nor* equal to the distance between L and N which corresponds to the intersection L where the marginal social cost curve cuts the demand curve corresponding to the premised demand level being equal to zero.

29) It is because *MSC*-curve cuts the *ED*-curve at point V.

30) Other numerical examples are shown, for reference, in Table N-1. Case 1 takes into account the exclusive existence of external agglomeration economies, while Case 2 is for the exclusive existence of agglomeration diseconomies. By Cases 3 and 4 are presented the situations in which external agglomeration economies and diseconomies are both involved. Four demand-surfaces and their associated *ED*-curves and *MGCS*-curves are diagrammatically illustrated for each of these cases by Figure N-4.

31) For the cost function side, we have a pair of curves; the *marginal* social cost curve and the *average* cost curve. A pair of curves for the demand function side are the *marginal* gross consumers surplus curve and the equal demand curve which can be regarded as *average* revenue curve from the viewpoint of suppliers.

REFERENCES

Beckmann, M., C. B. McGuire, and C. B. Winsten (1955): *Studies in the Economics of Transportation*, Cowles Commission, Yale.

Buchanan, J. M. (1956): "An Economic Theory of Clubs," *Economica*, 32.

———, (1965): "Private Ownership and Common Usage: The Road Case Re-examined," *Southern Economic Journal*, 22, pp. 305–315.

Button, K. J. and A. D. Pearman (1983): "Road Pricing —— Some of the More Neglected Theoretical and Policy Implications," *Transportation Planning and Technology*, 8, pp. 15–28.

Dewees, D. N. (1979): "Estimating the Time Costs of Highway Congestion," *Econometrica*, 47, pp. 1499–1512.

Else, P. K. (1981): "A Reformulation of the Theory of Optimal Congestion Taxes," *Journal*

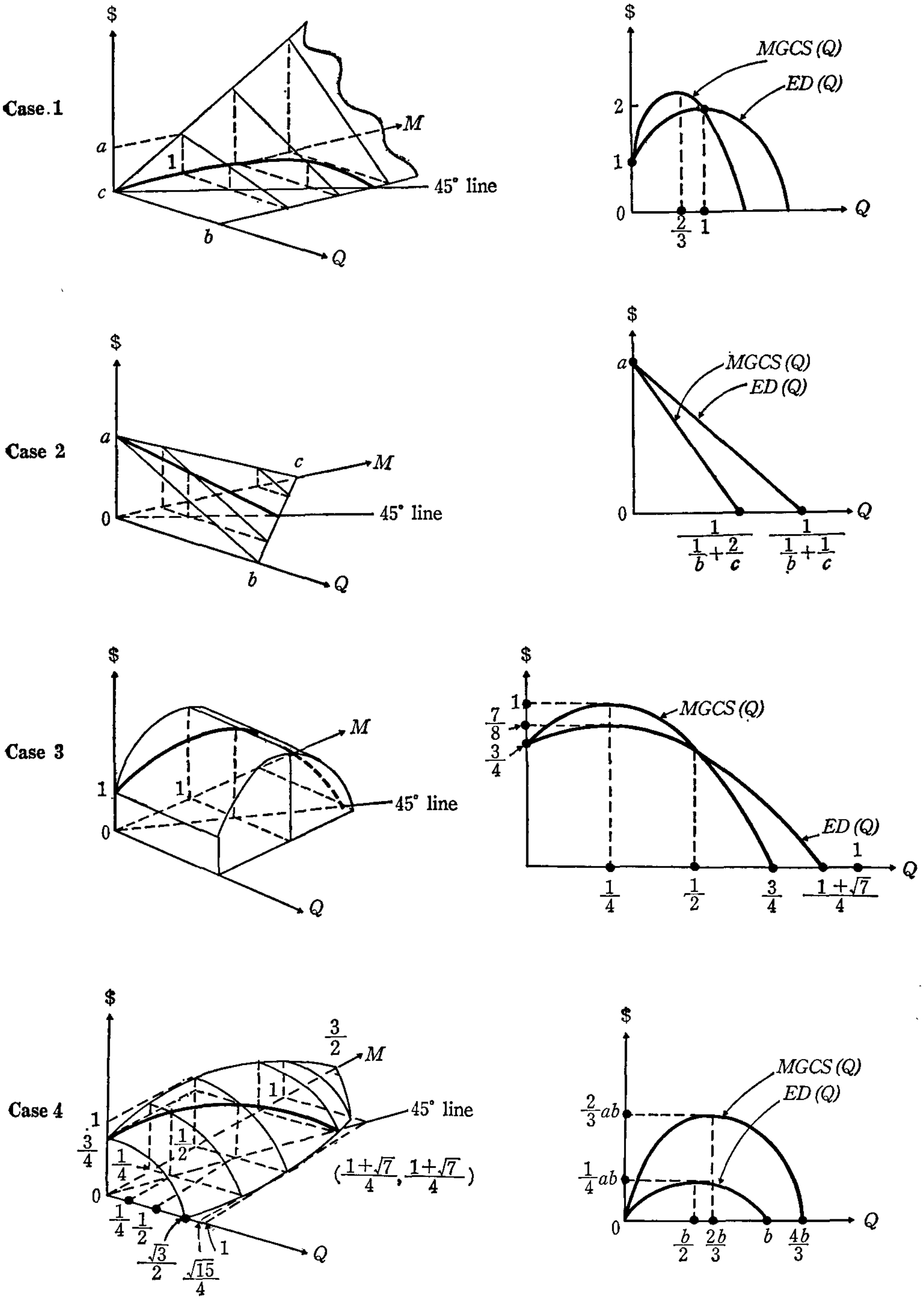


Figure N-4 Examples of Demand-surface

of Transport Economics and Policy, 15, pp. 217–232.
———, (1982): "A Reformulation of the Theory of Optimal Congestion Taxes: A Rejoinder," (in Notes and Comments of) *Journal of Transport Economics and Policy*, 16-3, pp. 299–304.
———, (1986): "No Entry for Congestion Taxes?," *Transportation Research-A*, 20A-2, pp. 99–107.
Jansson, J. O. (1969): "Optimal Congestion Tolls for Car Commutors," *Journal of Transport Economics and Policy.*
Johnson, M. B. (1964): "On the Economics of Road Congestion," *Econometrica*, 32, pp. 137–150.
Knight, F. H. (1924): "Some Fallacies in the Interpretation of Social Cost," *Quarterly Journal of Economics*, 38, pp. 582–606.
Kawashima, T. (1975): "Economics of Urban Environments," *Shin Toshi*, 29–3, Tokyo: Toshikeikaku Kyokai, pp. 4–14 (in Japanese).
———, (1981): "Urban Optimality: A Theoretical Study on the Effects of Urban Agglomeration Economies," in *Cities in Transition: Problems and Policies*, Ed. by P. Nijkamp and P. Rietveld, Alphen aan den Rijn, The Netherlands: Sijthoff & Noordhoff.
Meyer, J. R., M. J. Peck, J. Stenason, and C. Zwick (1959): *The Economics of Competition in the Transportation Industries*, Harvard.
Mohring, H. (1964): "Relation between Optimum Congestion Tolls and Present Highway User Charges," in *Traffic Congestion as a Factor in Road-user Taxation. Highway Research Record 47*, Washington, D.C.: National Academy of Science.
Nash, C. A. (1982): "A Reformulation of the Theory of Optimal Congestion Taxes: A Comment," (in Notes and Comments of) *Journal of Transport Economics and Policy*, 16-3, pp. 295–299.
Pigou, A. C. (1920): *Wealth and Welfare*, First Edition, London: Macmillan and Co.
Rabenau, B. and K. Stahl (1974): "Dynamic Aspects of Public Goods: A Further Analysis of the Telephone System," *Bell Journal of Economics and Management Science*, Fall.
Smeed, R. J. (1968): "Traffic Studies and Urban Congestion," *Journal of Transport Economics and Policy*, 2, pp. 33–70.
Vickrey, W. (1965): "Pricing as a Tool in Coordination of Local Transportation," *Transportation Research*, New York: National Bureau of Economic Research, pp. 275–291.
Walters, A. A. (1954): "Track Costs and Motor Taxation," *Journal of Industrial Economics.*
———, (1961): "The Theory and Measurement of Private and Social Cost of Highway Congestion," *Econometrica*, 29, pp. 676–699.